

Instituto de Estudios de Posgrado
Programa de Doctorado en Lenguas y Culturas



UNIVERSIDAD
DE CÓRDOBA

**El grado de especialización textual:
caracterización a partir de la percepción
sociocognitiva de la terminología
y las relaciones discursivas**

The text specialization degree:
features based on the sociocognitive perception
of terminology and discourse relations

Tesis doctoral
Para optar al título de Doctor por la Universidad de Córdoba

Sergio Rodríguez-Tapia

Director:
Dr. Alfonso Zamorano Aguilar
Profesor Titular de Lingüística General
Universidad de Córdoba

Fecha de depósito: 2 de noviembre de 2018

TITULO: *El grado de especialización textual: caracterización a partir de la percepción sociocognitiva de la terminología y las relaciones discursivas*

AUTOR: *Sergio Rodríguez Tapia*

© Edita: UCOPress. 2019
Campus de Rabanales
Ctra. Nacional IV, Km. 396 A
14071 Córdoba

<https://www.uco.es/ucopress/index.php/es/ucopress@uco.es>



TÍTULO DE LA TESIS:

El grado de especialización textual: caracterización a partir de la percepción sociocognitiva de la terminología y las relaciones discursivas

DOCTORANDO/A:

Sergio Rodríguez Tapia

INFORME RAZONADO DEL DIRECTOR DE LA TESIS

(se hará mención a la evolución y desarrollo de la tesis, así como a trabajos y publicaciones derivados de la misma).

La tesis que presenta el doctorando Sergio Rodríguez Tapia es fruto de un concienzudo trabajo de investigación y fruto de un plan de formación e investigación riguroso, serio y altamente productivo.

En un principio, la tesis pretendía ahondar en el análisis del texto semiespecializado (tipología menos investigada, frente a los textos especializados y los divulgativos); sin embargo, la rentabilidad y el rigor de la metodología empleada (teoría de la percepción sociocognitiva, aprendizaje automático y teoría comunicativa de la terminología) permitió, desde los primeros estadios de la tesis, ampliar el tipo de textos y centrar la investigación, por tanto, en el grado de especialización textual, con un foco muy especial en los textos semiespecializados. Por tanto, en esta tesis el Sr. Rodríguez Tapia lleva a cabo un preciso análisis empírico, descriptivo-interpretativo y correlacional para investigar los distintos extremos de lo que se denomina "grado de especialización" de los textos, tomando como ejes del estudio el léxico (como nivel lingüístico) y la relación emisor-destinatario (como nivel pragmático-discursivo).

Los objetivos e hipótesis se han abordado, a mi juicio, con extremo rigor científico y los resultados son concluyentes con relación al tema analizado y al avance del conocimiento en esta área. Son muchos los méritos y logros de esta tesis, sin embargo, me gustaría recalcar en dos de ellos: a) el carácter interdisciplinar y el avance no solo en terminología, sino también en lingüística textual y de corpus; b) la aplicación seria y meticulosa de instrumentos matemático-estadísticos para la demostración, cuantitativa y cualitativa, de las hipótesis que se plantean, lo que aporta un alto nivel de falsabilidad al trabajo.

A mi juicio, el Sr. Rodríguez Tapia ha sabido conjugar perfectamente el rigor científico, la metodología precisa y madura, los instrumentos teóricos y técnicos a su alcance, el trabajo experto con corpus y la precisión formal en la redacción y presentación de resultados de su tesis doctoral.

Como becario FPU desde el inicio de su etapa doctoral y siguiendo con escrupulosidad tanto el plan de investigación como el plan de formación, los

resultados a que ha llegado son muestra manifiesta de una completísima etapa doctoral. De especial interés y relieve en los resultados de su tesis son las tres estancias de investigación realizadas durante el proceso de elaboración de su trabajo doctoral:

- Institut Universitari de Lingüística Aplicada (Universitat Pompeu Fabra, Barcelona, España) (08/09/2016-08/10/2016-1 mes);
- Institut für Angewandte Linguistik und Translatologie (Universität Leipzig, Leipzig, Alemania) (01/05/2017-31/07/2017-2 meses);
- Dipartimento di Interpretazione e Traduzione (Università di Bologna, Forlì, Italia) (05/03/2018-05/06/2018-3 meses);

En los tres casos, se trata de centros especializados en la temática de la tesis, lo que le ha permitido contactos con investigadores/as y grupos de investigación de impacto en el panorama internacional, que ha redundado de forma muy positiva en su tesis. Este hecho, así como su formación en lenguas, le permite poder presentar esta tesis doctoral en el marco del programa de Mención Internacional, como se refleja en la composición de la propuesta de tribunal.

Los resultados científicos de su trabajo, bajo mi dirección y supervisión, se han concretado, de forma brillante y amplia con relación al mínimo exigido en el programa de doctorado, en la publicación 9 artículos en revistas científicas, 5 capítulos de libro, 3 colaboraciones en ediciones de libros, 3 comunicaciones en congresos nacionales, 7 comunicaciones en congresos internacionales, 1 ponencia por invitación en congreso nacional y 4 ponencias por invitación en congresos internacionales. Asimismo, su colaboración docente en materias del área de Lingüística General (como becario FPU) y su trabajo constante, como miembro colaborador, en las iniciativas del grupo de investigación que dirijo ("Pensamiento, lenguas y textos", HUM-060) y, con anterioridad en el grupo "Lenguajes" (HUM-224) son muestra más que evidente de su tesón, seriedad académica y rigor investigador.

Finalmente, los informes externos de evaluación de su CV y de la tesis doctoral (como requisito para la Mención internacional) avalan la calidad de su trabajo y la investigación doctoral realizada, con una excelente proyección en el terreno científico y académico.

Por todo ello, se autoriza la presentación de la tesis doctoral.

Córdoba, 22 de octubre de 2018

Firma del director


Prof. Dr. Alfonso Zamorano Aguilar

AGRADECIMIENTOS

Esta tesis doctoral ha sido posible gracias a la ayuda concedida para la Formación de Profesorado Universitario, de los subprogramas de Formación y de Movilidad incluidos en el Programa Estatal de Promoción del Talento y su Empleabilidad, en el marco del Plan Estatal de Investigación Científica y Técnica y de Innovación 2013-2016 por el Ministerio de Educación, Cultura y Deporte del Gobierno de España con referencia FPU14/05095.

Este proyecto, agotador y apasionante a partes iguales, ha estrechado sus lazos con multitud de personas a lo largo de los tres años en los que se ha ido desarrollando. Estas personas me han procurado consejo, apoyo académico y emocional, y, sobre todo, unas extraordinarias ganas de continuar.

En primer lugar, tengo que agradecer su inestimable apoyo a mi director, maestro y amigo, Alfonso Zamorano. Le doy gracias por saber ver más allá de cualquier párrafo, su paciente y pulcra dirección, sus consejos, sus ánimos y su conveniente exigencia, fruto de la cual es mi tímido avance académico. Ha sido y es el mejor padre académico que un doctorando pudiese tener.

A María Luisa Calero le agradezco que se embarcase en las primeras fases de esta aventura. Admiro su cariñosa atención a mi desarrollo académico y personal (gracias a la cual realicé mi primera estancia), su envidiable paciencia y su precisión microscópica. A Salvador López, por aceptar con los brazos abiertos y una sonrisa mi llegada al Departamento de Ciencias del Lenguaje. A María Martínez-Atienza, por estar siempre presente en la facultad y porque una charla y un debate acompañados de un café repone energías para seguir al pie del cañón. A Adela le agradezco ser compañera de camino y fatigas, unas más agradables que otras, pero siempre con la risa en la mochila. También, por supuesto, a mis compañeros del Departamento de Ciencias del Lenguaje de la Universidad de Córdoba, por mostrar interés por este proyecto y por hacerme sentir bienvenido.

Esta tesis no solo se desarrolló en Córdoba, sino que viajó conmigo a otros lugares. En el IULA, debo agradecer su cercanía a Mercè Lorente, que aceptó mi breve estancia y me hizo sentir como en casa. También debo agradecer su tiempo y sus comentarios, que sirvieron para mejorar y hacer madurar mi tesis, a Amor Montané, Besharat Fahti, Elisabet Llopart, Iria da Cunha, Jorge Vivaldi, Judit Freixa, Laida Vidal y María Teresa Cabré.

En el IALT de Leipzig, recuerdo especialmente los debates en los seminarios con mi tutor, Carsten Sinner, y las largas charlas con Dunia Hourani al pie de la facultad. Sus comentarios fueron más que bienvenidos y significan un antes y un después en este proyecto. También agradezco a Félix San Vicente acogerme como doctorando en Forlì.

Y como una tesis doctoral no es solo un proyecto académico, sino una aventura personal, no quisiera terminar esta sección sin nombrar a tres personas en especial. A Carmen; porque siempre está ahí, incluso en la distancia, porque es amiga y colega y porque

es agonía e ilusión a partes iguales. A Julio, por su impagable (literalmente) ayuda con las matemáticas y el aprendizaje automático, por reírse de (y con) todo y por su don didáctico. A ellos también les agradezco haberse convertido en compañeros académicos. A Javi, por tener que soportar las inclemencias de esta tesis y de mi sacrificio, por compartir éxitos, ser compañero incansable e irremplazable, y por haber sido los ojos que han visto con detalle todo el camino.

Por último, y desde luego no por eso menos importante, agradezco a mi padre, mi madre y mi hermana su interés, su apoyo incondicional y sus interminables preguntas por este mundo tan fascinante y complejo. Les doy las gracias por vivir mis alegrías como suyas y por su eterna curiosidad.

A todos ellos, mil gracias

RESUMEN

Este trabajo es un estudio empírico descriptivo correlacional que tiene por objetivo principal caracterizar el grado de especialización textual, particularmente su léxico y relación emisor-destinatario, a partir de un corpus de textos analizados por cuatro grupos de informantes, usando como base la teoría sociocognitiva de la percepción y el aprendizaje automático.

Esta tesis doctoral viene motivada por la escasez de estudios que pretenden estudiar el grado de especialización textual, en concreto, sus condicionantes y límites y, especialmente, el grado intermedio: el texto semiespecializado. Las hipótesis esbozadas están relacionadas con: (1) la posibilidad de encontrar pruebas lingüísticas que se ajusten a cada uno de los grados de especialización; (2) la relación de similitud entre texto semiespecializado y texto no especializado; (3) la relevancia del tema o el tipo textual en la discriminación del grado de especialización; (4) la similitud perceptiva de los usuarios a la hora de distinguir el grado; y (5) la relevancia de la relación emisor-destinatario y la terminología como factores de distinción del grado de especialización.

En este trabajo tomamos como marco de estudio la teoría de la comunicación para estudiar elementos diversos como son el texto, la tipología textual, el campo de la terminología, el término como objeto de estudio, la noción de valor especializado, el conocimiento especializado, los procedimientos de reformulación y las características vinculadas hasta ahora al discurso especializado, así como las propuestas de clasificación binaria y *continuum*.

Nuestro material de trabajo está compuesto por un corpus que cuenta con 315 textos y medio millón de palabras. Su representatividad cualitativa para estudiar el grado de especialización textual reside en que parte de tres temas diferentes (reconocidos socialmente, como son el derecho, la informática y la medicina). Posteriormente, hemos seleccionado siete tipos textuales posibles en los tres temas (el artículo en revista científica, la entrada de blog, la revista o web de divulgación especializada, la noticia, la entrevista, la participación en foro y la tesis doctoral). De esta combinación de tipo textual y tema se han recopilado 15 textos.

En dicho corpus se estudian diversas variables, tanto cuantitativas como cualitativas. Entre las variables cualitativas encontramos las variables nominales (el tipo de texto y el tema del texto). Entre las variables cualitativas ordinales encontramos el emisor, el destinatario, y la clase de texto. Entre las variables cuantitativas, se analizan el número

de formas léxicas, la ratio estandarizada entre tipos de formas léxicas y el total de formas léxicas, el índice de densidad terminológica, el número de procedimientos de reformulación y el índice de reformulación.

Parte de estas variables son proporcionadas por el investigador, mientras que otra parte es proporcionada por cuatro grupos de informantes (estudiante, usuario medio, traductor e investigador), lo que arroja una base de datos con 1260 patrones que es analizada mediante la correlación manual de dos o más variables y mediante aprendizaje automático. El análisis mediante técnicas de aprendizaje automático emplea los algoritmos de red bayesiana, de regresión logística, el árbol de decisión J48 y el algoritmo de agrupamiento no supervisado k-means mediante el software WEKA.

Las principales conclusiones sugieren la relevancia de las formas léxicas semicrípticas y crípticas en la discriminación de los polos opuestos del grado de especialización, así como el destinatario como factor regulador de la configuración discursiva de los textos cuando se atiende al grado. También se han identificado las similitudes del texto semiespecializado con los otros dos grados, principalmente con el texto divulgativo (desde un punto de vista formal) y con el texto especializado (desde un punto de vista semántico).

ABSTRACT

This work is an empirical descriptive correlational study whose main objective is to depict the text specialization degree, especially its lexical forms and speaker-recipient relationship, based on a corpus of texts analyzed by four groups of informants, using as a basis the sociocognitive theory of perception and machine learning.

This doctoral thesis is motivated by the scarcity of studies that intend to analyze the text specialization degree, in particular, its determining factors and limits and, especially, the intermediate degree: the semi-specialized text. The outlined hypotheses are related to: (1) the possibility of finding linguistic evidences that may correspond to each of the specialization degrees; (2) the relationship of similarity between semi-specialized text and non-specialized text; (3) the relevance of the topic or the type of text in the discrimination of specialization degree; (4) the perceptive similarity of the users when it comes to distinguishing the specialization degree; and (5) the relevance of the speaker-recipient relationship and the terminology as factors of distinction of the specialization degree.

In this work, the theory of communication is taken as a frame of reference to study diverse elements such as text, textual typology, the field of terminology, the term as an object of study, the notion of specialized value, specialized knowledge, reformulation methods and the characteristics linked up to now for the specialized discourse, as well as binary and continuum classification proposals.

Our material of study is composed of a corpus of 315 texts and half a million words. Its qualitative representativeness to study the text specialization degree lies in that it is based on three different subjects (socially recognized, such as law, computer science and medicine). Subsequently, seven possible types of texts in these three subjects have been selected (article in scientific journal, blog post, scientific-divulcation magazine or web, news, interview, participation in a forum and doctoral thesis). From this combination of type of text and subject, 15 texts have been collected.

In this corpus several variables are studied, both quantitative and qualitative. Among the qualitative variables the nominal variables are found: the type of text and the subject of the text. Among ordinal qualitative variables the speaker, the recipient, and the text class are studied. Among the quantitative variables, the number of lexical forms, the standardized ratio between types of lexical forms and the total of lexical forms, the terminological density index, the number of reformulation means and the reformulation index are analyzed.

Part of these variables are provided by the researcher, while another part is provided by four groups of informants (student, average user, translator and researcher), which yields a database with 1260 patterns that is analyzed through the manual correlation of two or more variables and through machine learning. The analysis using machine learning techniques employs Bayesian network, logistic regression algorithms, the J48 decision tree and the unsupervised k-means clustering algorithm using the WEKA software.

The main conclusions suggest the relevance of semi-cryptic and cryptic lexical forms in the discrimination of the opposite poles of the specialization degree, as well as the recipient as a regulating factor of the discursive configuration of the texts when attending the degree. The similarities of the semi-specialized text with the other two degrees have also been identified, mainly with the non-specialized text (from a formal point of view) and with the specialized text (from a semantic point of view).

Índice de contenidos

I. PRELIMINARES Y JUSTIFICACIÓN

1. SUPUESTOS DE PARTIDA Y PLANTEAMIENTO DEL PROBLEMA	15
2. OBJETO	17
3. HIPÓTESIS	18
4. OBJETIVOS	19
5. ESTRUCTURA DE LA TESIS DOCTORAL	20

II. FUNDAMENTOS TEÓRICOS

1. LA LINGÜÍSTICA APLICADA	25
1.1. Definición y objetivos de la lingüística aplicada	25
1.2. Objetos de trabajo y problemas.....	30
1.3. Lingüística (teórica), lingüística aplicada y aplicaciones de la lingüística	34
1.4. Recapitulación.....	36
2. LA COMUNICACIÓN COMO OBJETO Y PROBLEMA	37
2.1. Definiciones, propuestas y elementos definatorios de la comunicación	37
2.2. La comunicación especializada.....	41
2.3. Relación discursiva entre los interlocutores.....	44
2.4. Competencia cognitivo-comunicativa de los interlocutores	46
2.5. La función lingüística.....	50
3. EL TEXTO	52
3.1. Definiciones y propuestas	52
3.2. Los tipos textuales.....	56
3.2.1. Utilidad teórica de los tipos textuales	57
3.3. Conceptos en torno al texto: <i>clase, tipo, género y secuencia</i>	58
3.4. Las tipologías textuales	59
3.4.1. Clasificación de las tipologías textuales (I): problemas teóricos	60
3.4.1. Clasificación de las tipologías textuales (II): criterios de tipologización	61

4. LA TERMINOLOGÍA COMO CAMPO CIENTÍFICO.....	63
4.1. Teorías de corte normativo: la Teoría General de la Terminología (TGT).....	65
4.2. Teorías de corte comunicativo	67
4.2.1. La socioterminología.....	67
4.2.2. La Teoría Comunicativa de la Terminología (TCT)	67
4.3. Teorías de corte cognitivo.....	69
4.3.1. La Teoría Sociocognitiva de la Terminología (TSCT)	69
4.3.2. La Terminología basada en Marcos (TbM).....	70
5. FUNDAMENTOS PARA ANALIZAR EL TÉRMINO COMO OBJETO DE ESTUDIO	71
5.1. Punto de partida	71
5.2. La teoría de puertas, el principio de poliedricidad y el principio de variación...	74
5.3. Fundamentos discursivistas para el análisis del léxico especializado.....	77
5.4. Perspectivas de estudio del término	80
5.5. Modelos de relación entre discurso común y discurso especializado	81
5.5.1. El modelo de Baldinger (1952)	82
5.5.2. El modelo de Reindhardt (1966)	82
5.5.3. El modelo de Heller (1970).....	83
5.5.4. El modelo de Hoffmann (1985)	84
5.6. Definición y caracterización del término	85
5.7. Criterios para diferenciar <i>término</i> y <i>unidad léxica no especializada</i>	86
5.8. Variación especializada: diferencias entre situaciones de uso del léxico	88
6. EL VALOR ESPECIALIZADO (I).....	90
6.1. Nuevas aproximaciones: valor léxico y relevancia de la situación discursiva....	90
6.2. Dificultades para la identificación del término: denominación, comprensión y coste de procesamiento	95
6.3. Denominaciones de la terminología en torno al valor especializado.....	98
6.4. Fronteras del conocimiento especializado	99
6.5. Aproximación a su definición	100
6.6. Criterios para la distinción teórica del conocimiento especializado	101
7. RELACIONES DENOMINACIÓN-CONCEPTO-SITUACIÓN DISCURSIVA	104
7.1. Fundamentación en el principio de variación	104
7.2. La terminologización	106
7.3. La transterminologización.....	108
7.4. La banalización	109

7.4.1. La banalización efímera	113
7.4.2. La banalización perenne o crónica	115
7.5. La variación denominativa y conceptual	117
7.5.1. La variación conceptual a través de la polisemia parcial interdominio	118
7.5.2. La variación denominativa a través de la sinonimia intradominio.....	119
7.6. La desteterminologización	120
7.7. Esquema de relaciones	122
7.8. Los términos crípticos y délficos	123
8. EL VALOR ESPECIALIZADO (II): CONCLUSIÓN.....	126
8.1. Criterios para determinar el valor especializado.....	126
8.2. El principio de activación del valor especializado en el discurso común	128
9. EL TEXTO ESPECIALIZADO Y EL GRADO DE ESPECIALIZACIÓN.....	136
9.1. Primera definición y aproximaciones al texto especializado	136
9.2. Características del texto especializado.....	137
9.2.1. Características pragmáticas	137
9.2.2. Características cognitivas	138
9.2.3. Características lingüísticas	139
9.2.4. Características superestructurales	143
9.3. Principio de variación y eje vertical.....	144
9.4. Clasificación binaria y <i>continuum</i>	144
9.5. Modelos y métodos de clasificación del grado de especialización textual	148
10. LOS PROCEDIMIENTOS DE REFORMULACIÓN FORMAL EN TERMINOLOGÍA ...	151
10.1. La reformulación lingüística.....	152
10.1.1. La reformulación parafrástica	154
10.1.2. La reformulación no parafrástica	155
10.1.3. Los constituyentes de la reformulación.....	156
10.2. La reformulación en el texto especializado	157
10.3. Los procedimientos de reformulación formal en el texto especializado	158
10.4. Nuestra postura teórica y metodológica	159

III. FUNDAMENTOS METODOLÓGICOS

1. LA LINGÜÍSTICA DE CORPUS COMO INSTRUMENTO DE TRABAJO	165
1.1. Objeto de la lingüística de corpus	166
1.2. Fundamentos teóricos de la lingüística de corpus.....	167
2. LA PERCEPCIÓN SOCIOCOGNITIVA DEL TÉRMINO	168
2.1. Fundamentos de variación lingüística.....	168
2.2. La cognición lingüística	170
2.3. Variación no funcional vs. variación funcional y variación intraindividual vs. variación extraindividual.....	171
2.4. Conceptos relacionados: captación, atención, fijación, subjetividad, valoración y actitud.....	172
2.5. Fases y funcionamiento de la percepción	175
2.6. Tipos de percepciones y selección de información.....	177
2.7. Implicaciones teóricas y metodológicas de la percepción en nuestro estudio ..	179
3. CORPUS PARA ANALIZAR EL GRADO DE ESPECIALIZACIÓN.....	181
3.1. Criterios de confección del corpus	181
3.2. Disciplinas y campos temáticos	183
3.3. Tipos textuales	184
3.4. Representatividad, selección de muestras y problemas de compilación.....	186
3.5. Representatividad cualitativa	188
3.6. Tipo de corpus compilado.....	190
3.7. Representatividad cuantitativa: características estadísticas	193
4. TIPO DE ESTUDIO Y MÉTODO	197
5. VARIABLES DE ESTUDIO	198
5.1. Diversidad léxica: la relación tipos de formas léxicas y formas léxicas.....	201
5.2. Índice de densidad terminológica: cálculo y problemas	202
5.3. Índice de reformulación: cálculo y problemas	206
5.4. Relación discursiva emisor-destinatario	211
5.5. Las clases de los textos del corpus.....	214
6. GRUPOS DE INFORMANTES	216
6.1. Criterios de selección de los grupos de informantes.....	216
6.1.1. Género	216
6.1.2. Edad.....	216

6.1.3. Clase: nivel educativo y profesión	217
6.2. Materiales de trabajo	218
6.3. Grupo 1: estudiantes.....	220
6.4. Grupo 2: usuario medio de la lengua	221
6.5. Grupo 3: traductor profesional.....	221
6.6. Grupo 4: investigador.....	221
6.7. Datos que proporcionan los grupos de informantes.....	222
6.7.1. Tipos de términos y cómputo	223
6.7.2. Perfil de emisor y destinatario.....	224
6.7.3. Clase textual	225
6.7.4. Resumen	225
6.8. Datos que proporciona el investigador.....	226
7. APLICACIÓN DE LA INTELIGENCIA ARTIFICIAL	228
7.1. El aprendizaje automático	229
7.2. Red bayesiana	232
7.3. Regresión logística	234
7.4. Árbol de decisión J48.....	236
7.5. Agrupamiento mediante clasificación no supervisada: <i>simple k-means</i>	237

IV. ANÁLISIS Y RESULTADOS

1. ANÁLISIS CORRELACIONAL DESCRIPTIVO GLOBAL.....	244
1.1. Análisis del TEMA.....	244
1.1.1. Relación con la CLASE	244
1.1.2. Relación con el EMISOR y el SUBEMISOR	244
1.1.3. Relación con el DESTINATARIO y el SUBDESTINATARIO.....	248
1.1.4. Relación SUBEMISOR-SUBDESTINATARIO	250
1.1.5. Relación SUBEMISOR-CLASE	254
1.1.6. Relación SUBDESTINATARIO -CLASE	258
1.1.7. Relación con las FORMAS LÉXICAS ESPECIALIZADAS	265
1.1.8. Relación con los PROCEDIMIENTOS DE REFORMULACIÓN FORMAL.....	268
1.1.9. Conclusiones parciales sobre el TEMA.....	271
1.2. Análisis de la CLASE.....	272

1.2.1. Relación con el TIPO DE TEXTO.....	273
1.2.2. Relación con el EMISOR y el SUBEMISOR.....	277
1.2.3. Relación con el DESTINATARIO y el SUBDESTINATARIO.....	279
1.2.4. Relación SUBEMISOR-SUBDESTINATARIO.....	281
1.2.5. Relación con las FORMAS LÉXICAS ESPECIALIZADAS.....	284
1.2.6. Relación con los PROCEDIMIENTOS DE REFORMULACIÓN FORMAL.....	286
1.2.7. Conclusiones parciales sobre la CLASE.....	288
1.3. Análisis del TIPO DE TEXTO.....	289
1.3.1. Relación con el EMISOR y el SUBEMISOR.....	290
1.3.2. Relación con el DESTINATARIO y el SUBDESTINATARIO.....	290
1.3.3. Relación SUBEMISOR-SUBDESTINATARIO.....	293
1.3.4. Relación con las FORMAS LÉXICAS ESPECIALIZADAS.....	300
1.3.5. Relación con los PROCEDIMIENTOS DE REFORMULACIÓN FORMAL.....	303
1.3.6. Conclusiones parciales sobre el TIPO DE TEXTO.....	306
1.4. Análisis del ÍNDICE DE DENSIDAD TERMINOLÓGICA.....	307
1.4.1. Relación con el TEMA.....	308
1.4.2. Relación con la CLASE.....	310
1.4.3. Relación con el TIPO DE TEXTO.....	311
1.4.4. Relación con el EMISOR y el SUBEMISOR.....	313
1.4.5. Relación con el DESTINATARIO y el SUBDESTINATARIO.....	315
1.4.6. Relación SUBEMISOR-CLASE-IDT.....	317
1.4.7. Relación SUBDESTINATARIO-CLASE-IDT.....	318
1.4.8. Relación SUBEMISOR-SUBDESTINATARIO-IDT.....	319
1.4.9. Conclusiones parciales sobre la DENSIDAD TERMINOLÓGICA.....	323
1.5. Análisis del ÍNDICE DE REFORMULACIÓN.....	324
1.5.1. Relación con el TEMA.....	325
1.5.2. Relación con la CLASE.....	326
1.5.3. Relación con el TIPO DE TEXTO.....	327
1.5.4. Relación con el EMISOR y el SUBEMISOR.....	330
1.5.5. Relación con el DESTINATARIO y el SUBDESTINATARIO.....	331
1.5.6. Relación SUBEMISOR-CLASE-IR.....	333
1.5.7. Relación SUBDESTINATARIO-CLASE-IR.....	338
1.5.8. Relación SUBEMISOR-SUBDESTINATARIO-IR.....	344
1.5.9. Relación TIPO DE TEXTO-CLASE-IDT-IR.....	346

1.5.10. Conclusiones parciales sobre la REFORMULACIÓN	349
1.6. Análisis de la VARIEDAD LÉXICA	350
1.6.1. Según TEMA e IDT	350
1.6.2. Según TEMA e IR	351
1.6.3. Según CLASE e IDT	352
1.6.4. Según CLASE e IR	353
1.6.5. Según TIPO DE TEXTO e IDT	354
1.6.6. Según TIPO DE TEXTO e IR	355
1.6.7. Conclusiones parciales sobre la VARIEDAD LÉXICA.....	356
1.7. Resumen de análisis correlacional descriptivo global	356
2. ANÁLISIS CORRELACIONAL DESCRIPTIVO SEGÚN LOS INFORMANTES.....	358
2.1. Análisis de la CLASE	358
2.1.1. Vinculación con el TIPO DE TEXTO	361
2.2. Análisis del EMISOR y del SUBEMISOR.....	367
2.2.1. Vinculación con el TIPO DE TEXTO	369
2.3. Análisis del DESTINATARIO y del SUBDESTINATARIO	373
2.3.1. Vinculación con el TIPO DE TEXTO	375
2.4. Análisis de las FORMAS LÉXICAS ESPECIALIZADAS	379
2.4.1. Vinculación con el TIPO DE TEXTO	385
2.5. Análisis del ÍNDICE DE DENSIDAD TERMINOLÓGICA.....	386
2.5.1. Vinculación con el TEMA.....	387
2.5.2. Vinculación con el TIPO DE TEXTO	388
2.5.3. Vinculación con la CLASE.....	393
2.5.4. Vinculación con el EMISOR y el SUBEMISOR.....	394
2.5.5. Vinculación con el DESTINATARIO y SUBDESTINATARIO	398
2.6. Conclusiones parciales sobre los INFORMANTES	402
2.7. Resumen de análisis correlacional descriptivo según los informantes	404
3. ANÁLISIS CORRELACIONAL A PARTIR DE APRENDIZAJE AUTOMÁTICO	410
3.1. Probabilidades de clasificación en clases según variables.....	410
3.2. Relevancia de los atributos	420
3.3. Árboles de decisión	428
3.4. Grupos identificados según el aprendizaje automático.....	438

V. SUMMARIZATION AND CONCLUSIONS

1. SUMMARIZATION	453
2. CONCLUSIONS	457
3. CONTRIBUTION OF THIS THESIS TO THEORETICAL AND APPLIED LINGUISTICS.....	471
4. FUTURE LINES OF RESEARCH	474

VI. REFERENCIAS BIBLIOGRÁFICAS

1. REFERENCIAS BIBLIOGRÁFICAS	479
--	------------

VII. ÍNDICE DE ILUSTRACIONES, GRÁFICOS Y TABLAS

1. ÍNDICE DE ILUSTRACIONES, GRÁFICOS Y TABLAS.....	507
---	------------

Lista de abreviaturas

ART	Artículo en revista científica	IR a)	Índice de reformulación por expansión
BLOG	Entrada de blog	IR b)	Índice de reformulación por reducción
CC	Conocimiento común	IR c)	Índice de reformulación por variación
CE	Conocimiento especializado	IR	Índice de reformulación
CG	Conocimiento general	LA	Lingüística Aplicada
DC	Discurso común	Leg NF	Lego no formado
DE	Discurso especializado	MED	Medicina
DER	Derecho	NOT	Noticia
DIV	Revista o web de divulgación especializada	PRF	Procedimientos de reformulación formal
ENT	Entrevista	PRF a)	Procedimientos de reformulación formal por expansión
Esp I	Especialista investigador	PRF b)	Procedimientos de reformulación formal por reducción
Esp NI	Especialista no investigador	PRF c)	Procedimientos de reformulación formal por variación
Est	Estudiante	PRFC	Procedimientos de reformulación formal y conceptual
FL	Forma léxica	TbM	Terminología basada en Marcos
FLE a)	Formas léxicas especializadas délficas / términos délficos	TCT	Teoría Comunicativa de la Terminología
FLE b)	Formas léxicas especializadas semicrípticas / términos semicrípticos	TE	Texto especializado
FLE c)	Formas léxicas especializadas crípticas / términos crípticos	TES	Tesis doctoral
FLE	Forma léxica especializada	TGT	Teoría General de la Terminología
FOR	Participación en foro	TNE	Texto no especializado
IDT a)	Índice de densidad terminológica de formas délficas	TSCT	Teoría Sociocognitiva de la Terminología
IDT b)	Índice de densidad terminológica de formas semicrípticas	TSE	Texto semiespecializado
IDT c)	Índice de densidad terminológica de formas crípticas	Trad	Traductor
IDT	Índice de densidad terminológica	TTR	Ratio <i>types/tokens</i> estandarizada
INF	Informática	UCE	Unidad de conocimiento especializado
Ins AE	Instruido aprendiz de especialista	UL	Unidad léxica
Ins LF	Instruido lego formado	ULE	Unidad léxica especializada
Ins SE	Instruido semiespecialista	USE	Unidad de significación especializada
Inv	Investigador	Us Med	Usuario medio

I. PRELIMINARES Y JUSTIFICACIÓN

1. SUPUESTOS DE PARTIDA Y PLANTEAMIENTO DEL PROBLEMA

Las teorías terminológicas (la Teoría Comunicativa de la Terminología [TCT], Cabré, 1999; y la Teoría Sociocognitiva de la Terminología [TSCT], Temmerman, 2000 y 2002, principalmente) se han centrado en describir con detalle las características de los textos especializados y contrastarlas con las de los textos divulgativos o no especializados. Si bien todas las teorías descriptivas (las que se dedican a explicar el uso textual real de los discursos y situaciones comunicativas especializadas) contemplan un nivel intermedio o una gradación, los estudios dedicados a este perfil central del *continuum* han sido muy escasos en comparación con los enfocados en el nivel especializado, que han prestado más atención a determinar las características lingüísticas (Cabré *et al.*, 2007), en especial el nivel léxico, ya que es ampliamente reconocido por la comunidad científica que constituye la característica esencial del texto especializado (Monterde Rey, 2002b: 115) o en establecer relaciones con otros campos, como la traducción (Estopà, 2014), la terminografía (Alarcón *et al.*, 2007) o la enseñanza de español para fines específicos (Gómez de Enterría, 2001 y 2007).

Esta tesis doctoral parte de un trabajo de investigación iniciado en 2013 y titulado *Los textos semiespecializados: procedimientos de reformulación formal y conceptual y estrategias de traducción (EN-ES). Análisis cuantitativo de resultados*, dirigido por el profesor Dr. Alfonso Zamorano, cuyos resultados se publicaron en tres estudios (Rodríguez-Tapia, 2015, 2016a y 2016b). Esta investigación permitió identificar las lagunas teórico-metodológicas que existen en la bibliografía científica en torno al concepto de *grado de especialización* y, en concreto, en torno al de *texto semiespecializado*, así como demostrar la hipótesis de *continuum* y cumplir diferentes objetivos, entre los que destacan:

- a) Llevar a cabo una aproximación a las características lingüísticas y extralingüísticas del texto semiespecializado.
- b) Realizar una propuesta de análisis del texto especializado en sus distintos niveles.
- c) Caracterizar el texto semiespecializado desde el punto de vista de los procedimientos de reformulación formal y conceptual a partir de un análisis cuantitativo de un corpus que incluye textos especializados, semiespecializados y divulgativos.

El proyecto de esta tesis doctoral pretende continuar las investigaciones ya iniciadas en anteriores trabajos, donde hemos tratado el concepto de *continuum* y de *texto*

semiespecializado. Los resultados de investigaciones previas arrojaron luz sobre el comportamiento del nivel microestructural (Dijk, 1992 [1983]) del texto semiespecializado y, en concreto, sobre cómo afecta la reformulación a los términos. Las conclusiones más relevantes determinaron que la reformulación, tanto formal como conceptual, afecta de manera similar a textos divulgativos, semiespecializados y especializados, siendo los procedimientos de los dos primeros idénticos tanto en cantidad como en tipología (Rodríguez-Tapia, 2015). Asimismo, determinamos que la densidad terminológica¹ del texto semiespecializado se encuentra en un valor intermedio entre el divulgativo y el especializado (Rodríguez-Tapia, 2016b), lo que favorece nuestra hipótesis de *continuum*.

No obstante, las limitaciones encontradas en los resultados (Rodríguez-Tapia, 2016a), constituyen suficiente justificación para la propuesta de nuevas hipótesis y objetivos. Entre estas limitaciones, encontramos:

- a) Las características que recoge el modelo de análisis cualitativo propuesto no son lo suficientemente integradoras como para comprender la complejidad del texto semiespecializado, su relación con los polos opuestos y su integración en el marco del *continuum*.
- b) Los límites rígidos de carácter cuantitativo para determinar el grado de especialización, aun siendo lo suficientemente orientativos y descriptivos, resultan un instrumento demasiado complejo para su uso académico (ya que concibe los límites de forma demasiado estricta) y su uso profesional (ya que consume demasiado tiempo en el análisis).
- c) El estudio se centró en un corpus configurado en torno a un objeto concreto del campo médico (la insuficiencia cardíaca), lo que excluyó multitud de otras disciplinas o campos y supone silenciar otras realizaciones lingüísticas.
- d) Las conclusiones del estudio permitieron proponer aplicaciones desde un punto de vista teórico, si bien deben demostrar su valor instrumental en la lingüística aplicada, disciplina que, a nuestro juicio, puede explotar los resultados de diversas formas.

¹ La densidad terminológica constituye la relación entre el número de unidades léxicas especializadas por unidad léxica de un texto (Domènech, 2006: 207).

2. OBJETO

Esta tesis doctoral trata los problemas teórico-metodológicos que supone construir una tipología textual según el grado o *continuum* de especialización textual. Nos referimos al grado de abstracción que realiza el usuario sobre un objeto temático concreto en el eje vertical. Este *continuum*, como avanzamos anteriormente, es divisible en tres categorías teóricas interrelacionadas por diversos factores: el texto divulgativo, el texto semiespecializado y el texto especializado. Tanto el denominado texto especializado como el texto divulgativo han caracterizado los objetos de estudio de diversas líneas de investigación nacionales e internacionales. El centro del *continuum*, que representa un objeto, a nuestro juicio, de confluencia de elementos característicos de los polos opuestos, no ha recibido suficiente atención en la reflexión lingüística científica. Entre uno de los trabajos que se aproximó a la definición del texto semiespecializado, encontramos el que desarrollamos en Rodríguez-Tapia (2016b), de acuerdo con los fundamentos teórico-metodológicos desarrollados en Rodríguez-Tapia (2016a):

Hasta el momento, los corpus que hemos estudiado nos permiten sostener que el texto semiespecializado se define como aquel texto redactado por un especialista y dirigido indistintamente a un público lego o especialista, en el que predominan, de forma independiente o en conjunto, las funciones representativa y comunicativa, cuyo índice de densidad terminológica no es alto (0,12 de media) y que se asocia a la superestructura de los glosarios, artículos divulgativos y artículos científicos (Rodríguez-Tapia, 2016b: 246).

Entre las cuestiones que suscitan más problemas desde el punto de vista metodológico se encontrarían el empleo de un valor numérico para caracterizar el texto semiespecializado o la amplitud de la propia definición (fruto, no obstante, de la esencia teórica del texto semiespecializado).

Asimismo, debido a las exigencias sociales de la actualidad, en las que la comunicación interlingüística entre científicos y especialistas recibe cada día mayor atención, las utilidades que atribuimos a la identificación del grado de especialización de los textos se corresponden con perspectivas de la lingüística aplicada. Entre ellas se encuentran las siguientes (las dos primeras reciben especial atención en nuestra tesis doctoral):

- a) La confección de criterios teórico-metodológicos para la compilación de corpus lingüísticos que sirvan a multitud de usuarios: lingüistas profesionales o especialistas en otras disciplinas que necesiten realizar consultas lingüísticas sobre sus materias, como médicos, ingenieros, etc.

- b) El desarrollo de criterios para la creación de recursos terminográficos y lexicográficos y la extracción de contextos que sirvan de ejemplos reales para estos recursos (como manuales de enseñanza de lenguas o diccionarios, por ejemplo).
- c) La predicción de los potenciales problemas de comprensión o de traducción de un texto, lo que resulta de utilidad en el campo de la docencia, así como en el profesional a la hora de evaluar las necesidades de un proyecto de traducción.
- d) La predicción de la dificultad de comprensión y adquisición de conocimiento y vocabulario por parte de los aprendices de lenguas para fines específicos.

3. HIPÓTESIS

Adoptando la idea de *continuum* y del grado de especialización propuesto por diversas teorías terminológicas, nuestra tesis doctoral parte de las siguientes hipótesis:

- a) **Hipótesis 1:** Es posible emplear las variables relacionadas con la percepción del léxico y con el perfil cognitivo-comunicativo de los emisores y destinatarios de los textos para percibir el objeto texto desde la consideración de objeto marcado por el grado de especialización en las categorías de texto no especializado, texto semiespecializado y texto especializado, es decir, por el *continuum*.
- b) **Hipótesis 2:** Existen indicios o pruebas suficientes en la configuración de un texto para clasificarlo como texto semiespecializado de igual forma que existen (y han sido investigadas) para el texto especializado.
- c) **Hipótesis 3:** El texto semiespecializado se acerca más a las características formales del texto no especializado. Solo comparte criterios semánticos con el texto especializado.
- d) **Hipótesis 4:** Los tipos textuales que adopta el texto semiespecializado son idénticos a los que puede adoptar tanto el no especializado como el especializado.
- e) **Hipótesis 5:** La configuración terminológico-conceptual de las diferentes disciplinas afecta directamente a la catalogación de los textos según su grado de especialización a partir de la percepción del destinatario.
- f) **Hipótesis 6:** No existen diferencias perceptivas entre diferentes grupos sociales a la hora de clasificar los textos según el grado de especialización, si bien pueden

presentar divergencias en el criterio que aplican para distinguir los tipos de términos.

- g) **Hipótesis 7:** La relación discursiva entre emisor y destinatario es el factor que regula el resto de las variables, y, por tanto, la configuración discursiva del texto. Así, dicha relación puede condicionar el grado de especialización mediante la combinación con otras variables.
- h) **Hipótesis 8:** El factor más relevante para distinguir el grado de especialización textual es la densidad terminológica de las unidades léxicas que se perciben como opacas semánticamente o como poco frecuentes en el discurso común.
- i) **Hipótesis 9:** La reformulación terminológica de un texto va ligada inevitablemente a la densidad terminológica de dicho texto y sus valores aumentan cuando incrementa el grado de especialización textual.

4. OBJETIVOS

Para demostrar estas hipótesis, proponemos la consecución de los siguientes objetivos:

- a) Proponer una relación de criterios para determinar el valor especializado de las unidades léxicas, a fin de poder ser analizadas posteriormente mediante el empleo de un modelo teórico para el grado de especialización de los textos.
- b) Proponer un modelo integrador para analizar el grado de especialización de los textos. Este modelo aunará las propuestas de análisis del texto y la perspectiva de las teorías terminológicas antes mencionadas, centrándonos en el nivel microestructural para justificar parámetros relevantes como la relación emisor-destinatario, la superestructura o la reformulación, y adoptando un enfoque orientado a la lingüística aplicada.
- c) Reunir un grupo de diferentes subconjuntos de informantes que aporten información sobre su percepción pragmático-discursiva, léxica y clasificatoria de los textos.
- d) Compilar un corpus representativo y multidisciplinar de áreas biomédicas, jurídicas y técnicas que contemplen diferentes tipos textuales y grados de especialización de acuerdo con las características determinadas por la TCT y las propuestas que adoptemos de la lingüística del texto.

- e) Identificar la relevancia de cada uno de los parámetros que contribuyen a la distinción del grado de especialización y distinguir los posibles patrones que existan entre dichos parámetros.
- f) Evaluar los parámetros que intervienen en la identificación del texto semiespecializado y que resultan de posible aplicación tanto en la configuración del texto especializado como en la del divulgativo.
- g) Identificar las particularidades formales y discursivas que caracterizan al texto semiespecializado, los tipos textuales que adopta y las posibles diferencias entre disciplinas.
- h) Proponer usos para identificar y explotar el texto semiespecializado en la lingüística aplicada, principalmente en lingüística de corpus, ya que estos campos representan los sectores profesionales de los servicios lingüísticos más demandados en el mercado, a los que afecta constantemente el dinamismo de la ciencia y su divulgación, desde la perspectiva práctica y metodológica respectivamente².

5. ESTRUCTURA DE LA TESIS DOCTORAL

Esta tesis doctoral se estructura en seis grandes bloques. El primero, en el que se inserta esta sección se centra en la exposición de argumentos que motivan el estudio, la descripción del objeto, así como la enumeración de hipótesis, objetivos e instrumentos.

El bloque II de esta tesis doctoral se destina a reflexionar sobre el estado de la cuestión en torno al grado de especialización textual y a determinar aquellos aspectos teóricos (tanto conceptuales como denominativos) que adoptamos como base para el posterior análisis de datos. Entre estos fundamentos teóricos destacan:

- a) La concepción de la lingüística aplicada como dimensión de la lingüística orientada a la resolución de problemas comunicativos de la sociedad.

² En relación con este objetivo Cabré *et al.* (2007: 853) exponen: «el valor de este estudio para la lingüística y para el análisis del discurso radica en la caracterización misma del discurso especializado, y para la lingüística aplicada se puede concretar en utilidades tan diversas como el reconocimiento automático de los textos especializados en Internet, la enseñanza de la lengua para fines específicos -es decir, de los discursos especializados- e, incluso, la valoración y la adaptación de la dificultad de los textos en los ámbitos académico, profesional y de los medios de masas (p. ej., divulgación periodística de la ciencia)».

- b) Los diferentes perfiles que pueden encontrarse en la comunicación especializada y su gradación, con atención a la subclasificación del especialista, el instruido y el lego.
- c) La noción de texto y de tipo textual y su repercusión teórico-metodológica en el diseño de criterios de compilación de corpus.
- d) La adopción de la Teoría Comunicativa de la Terminología como marco de configuración del acto discursivo especializado y su gradación.
- e) La adopción de la noción de valor y de sentido situado como determinante para discriminar una forma léxica especializada de una forma léxica no especializada, así como la propuesta teórica para la discriminación de este valor cuando las condiciones discursivas no son las prototípicas.
- f) La descripción de las relaciones que establece la denominación con el concepto y la situación discursiva.
- g) La reformulación y sus diferentes procedimientos, así como una propuesta para identificarlos.
- h) La revisión de la noción de texto especializado, sus características y los diferentes modelos que analizan la existencia del grado de especialización textual.

El bloque III se encarga de la exposición de los fundamentos metodológicos empleados. Partimos de una sección dedicada a la lingüística de corpus como objeto e instrumento de trabajo, así como de otra sección en la que se debaten los conceptos más relevantes de la teoría sociocognitiva de la percepción, de la cual tomamos tanto la noción de *percepto*, *ocurrencia* o *prominencia*, como de *percepción analítica* y *percepción sintética*. El bloque III también se encarga de justificar la compilación de nuestro corpus desde el punto de vista de la representatividad cuantitativa y cualitativa, de determinar el tipo de estudio y de clasificar las variables de estudio, entre las que se encuentran la diversidad léxica, el índice de densidad terminológica, el índice de reformulación, la relación discursiva emisor-destinatario y las clases de textos según el grado de especialización. Asimismo, también se incluye la descripción de los grupos de informantes, de los datos que proporcionan estos últimos en contraste con el investigador y de la exposición de los algoritmos de aprendizaje automático empleados en esta tesis doctoral, centrándonos en su función y utilidad.

El bloque IV se divide en tres partes: (i) en la primera parte se realiza un análisis correlacional descriptivo con el conjunto de datos analizados por los cuatro grupos de informantes, de forma que estudiamos el corpus en su totalidad; (ii) en la segunda

centramos nuestra atención en identificar las similitudes y diferencias perceptivas que se observan en los resultados analizados por los cuatro grupos de informantes; y (iii) en la tercera se aplican los cuatro algoritmos de aprendizaje automático para extraer información relacionada con la probabilidad, la relevancia de las variables, las reglas de clasificación o los centroides o patrones más representativos de una clase.

El bloque V se centra en la recapitulación de lo desarrollado en la tesis, la exposición de conclusiones, en las que se comprueban o falsan las hipótesis planteadas, así como se propone una relación de futuras líneas de investigación. El bloque VI recoge la bibliografía empleada para la redacción de este trabajo.

En cuanto a la distribución de la información contenida en la tesis, deben señalarse dos cuestiones:

- a) Esta tesis doctoral está acompañada de varios ficheros PDF en un CD adjunto, que constituyen los anexos a esta investigación. Los archivos contenidos son los siguientes:
 - Anexo 1A y 1B: base de datos con la relación de información cualitativa y cuantitativa tras analizar las variables del corpus.
 - Anexo 2: hoja de instrucciones para informantes.
 - Anexo 3A y 3B: relación de tablas y gráficos con el cruce de variables, por datos globales y según informantes.
 - Anexo 4: datos bibliográficos de las fuentes del corpus.
 - Anexo 5: ejemplo de consentimiento para informantes.
- b) Junto con estos anexos, con objetivo de facilitar la lectura, se incluye una ficha en soporte físico que relaciona la leyenda de colores empleada en la tesis, así como el listado de abreviaturas más empleadas. Recibe el nombre de Anexo 6.
- c) Las referencias bibliográficas de las fuentes y citas textuales que se emplean como ejemplos reales, ya provengan del corpus o no, se relacionan como nota al pie en la misma página en la que aparece el ejemplo y no se relacionan en la sección de referencias bibliográficas final, con el objetivo de que el lector cuente con la información sobre el tipo textual, el autor o el año de forma inmediata.
- d) Toda referencia contenida en esta tesis doctoral a personas de sexo masculino se entenderá hecha igualmente a personas de sexo femenino y de forma inversa salvo que se explicita o el contexto indique lo contrario.

II. FUNDAMENTOS TEÓRICOS

1. LA LINGÜÍSTICA APLICADA

1.1. Definición y objetivos de la lingüística aplicada

Esta tesis doctoral se inscribe dentro del marco de la lingüística aplicada, dimensión de la lingüística que tiene por objetivo resolver los problemas sociales que involucran la lengua o el lenguaje.

Los dominios y la definición de la lingüística aplicada (en adelante LA) constituyen una cuestión difícil de delimitar y acotar actualmente. Desde sus inicios a finales de los años 50, y hasta su desarrollo exponencial a partir de los 90, la LA ha experimentado un cambio de concepción epistemológica, un debate (todavía vigente) sobre su estatus disciplinario y una constante reflexión sobre las taxonomías que permite ordenar, acotar o delimitar los problemas que intenta solucionar.

Como mostramos a continuación, las definiciones anteriores a los años 90 abogan por una concepción de la LA en la que se propone la *aplicación* de los estudios de la lingüística³:

The creation of applied linguistics as a discipline represents an attempt to find **practical applications** for ‘modern scientific linguistics’ (Mackey, 1966: 197).

The **application of linguistic knowledge** to some object —or applied linguistics, as its name implies— is an activity. It is not a theoretical study. **It makes use of the findings of theoretical studies** (Corder, 1973: 10).

A branch of linguistics where the primary concern is **the application of linguistic theories**, methods and findings to the elucidation of language problems which have arisen in other areas of expertise (Crystal, 1980: 28-29).

Applied linguistics has been considered a subset of linguistics for several decades, and it has been interpreted to mean the **applications of linguistics principles** to certain more or less practical matters (Brown, 1987: 147).

Tanto Mackey como Corder, Crystal, Brown y Bright (este ya en los 90) hacen hincapié en esa *aplicación* o *uso práctico* de los resultados de la lingüística. Entendemos que esta vinculación con el uso práctico pretende complementarse o contrastar con la llamada *lingüística teórica*. De hecho, la definición de Corder (1973: 10) explicita que considera la LA una actividad y que «it is not a theoretical study».

³ La negrita es del autor de la tesis.

A partir de los 90 se aprecia la atención a la interdisciplinariedad y no al estatus primordial de la lingüística como disciplina progenitora. También comienzan a observarse los conceptos de *problema* y de *comunicación real* (efectivamente, Bright, en la definición anterior, avanza algunos de los campos en los que existe ese *problema*)⁴.

[...] una serie de propiedades de la LA que podríamos enumerar como: **utilidad**, **valor práctico**, **sentido heteroteleológico**, **interdisciplinariedad** y variedad [...]. (Calvo, 1990: 322-323).

Applied linguistics is a **multidisciplinary approach** to the solution of **language-related problems** [in language education, language policy and planning, speech-communication research, specialized occupational languages, speech therapy, lexicography and dictionary-making, translation and interpreting, and language and the professions] (Stevens, 1992: 17-22).

Whenever knowledge about language is used to solve a **basic language-related problem**, one may say that A[pplied] L[inguistics] is being practiced (Kaplan y Widdowson, 1992: 76).

Applied linguistics, as is widely recognized, is a field the purpose of which is to solve **real-world language-based problems**. In order to do this, the field draws upon linguistics, but not exclusively; it draws on knowledge from a variety of other fields as well (Kaplan y Grabe, 1992: 3).

La lingüística aplicada puede entenderse como una orientación o una dimensión de la investigación lingüística en el conjunto de campos de estudio de las ciencias del lenguaje. Con esta orientación, y desde un marco teórico **interdisciplinario**, la lingüística aplicada tiene una **finalidad práctica** consistente en **resolver los problemas** provocados por los procesos comunicativos propios de las **sociedades actuales** (Payrató, 1998: 33).

[Applied linguistics is] the theoretical and empirical **investigation of real-world problems** in which language is a central issue (Brumfit, 1997: 93).

[...] the focus of applied linguistics is on trying to resolve **language-based problems** that people encounter in the **real world**, whether they be learners, teachers, supervisors, academics, lawyers service providers, those who need social services, test takers, policy developers dictionary makers, translators, or a whole range of business clients (Grabe, 2002: 9).

A practice-driven discipline that addresses **language-based problems** in **real-world** contexts (Grabe, 2002: 10; Grabe. 2010: 42).

⁴ La negrita y el uso de colores son del autor de la tesis. Los colores que se emplean en las citas pretenden representar de forma gráfica las similitudes en las definiciones: el azul para el valor práctico, el rojo para la interdisciplinariedad y el verde hace referencia al estudio del discurso real.

“Applied Linguistics” is using what we know about (a) language, (b) how it is learned, and (c) how it is used, in order to achieve some purpose or **solve some problem** in the **real world** (Schmitt y Celce-Murcia, 2002: 1).

Applied linguistics can be described as a broad **interdisciplinary field** of study concerned with **solutions to problems** or the improvement of situations involving language and its users and uses (Berns y Matsuda, 2006: 394).

Applied linguistics activity focuses on a range of language-related issues and often **draws upon other disciplines** in studying **language problems**. Engaging expertise from other professions (e.g., medicine and law) or fields of study (e.g., psychology, communication studies, or sociology) presupposes precise identification of the problem to determine its scope (Berns y Matsuda, 2006: 396).

En efecto, las nociones de comunicación real, problema, utilidad e interdisciplinarietà están estrechamente relacionadas y, de hecho, desde nuestra perspectiva, establecen la mencionada relación en ese orden. Partiendo de la sincronía y la *parole* saussureana y en el contexto de desarrollo teórico y empírico de la sociolingüística, la pragmática y la lingüística textual y de la comunicación, la relevancia de analizar la relación entre usuarios, la comunicación no artificializada, permite identificar problemas desde diferentes perspectivas teórico-metodológicas (que coinciden con las perspectivas desde las que las diferentes disciplinas analizan sus respectivos objetos de estudio).

En nuestro caso, el problema de identificar el grado de especialización textual puede determinarse desde diferentes perspectivas del poliedro (principio de la poliedricidad; Cabré, 2008; Zamorano, 2013). Determinar el grado de especialización textual constituye un problema:

- a) metodológico para la lingüística de corpus y la documentación,
- b) ontológico para la terminología,
- c) ontológico y metodológico para la traducción,
- d) etc.

El interés por solucionar estos problemas obliga a que exista una relación de coordinación y cooperación entre la lingüística y otras disciplinas muy diversas (tanta diversidad como problemas se identifiquen), que van desde la neurología a las ciencias jurídicas, pasando por la medicina o la ingeniería informática, por esbozar solo algunos ejemplos. Estas relaciones pueden dilucidarse a partir de las definiciones anteriores de Kaplan y Grabe (1992: 3), Grabe (2002: 9) y Berns y Matsuda (2006: 396).

En cuanto a esta interdisciplinariedad, parece que existe un claro consenso⁵ por parte de la comunidad científica en reconocer la actividad de la LA en los límites de contacto entre disciplinas. Estos límites se perciben como resultado de la cooperación entre disciplinas y revelan la complejidad y dinamismo del conocimiento que intervienen en la resolución de los problemas reales de la comunicación. Esta cooperación a la que hacemos referencia puede percibirse en las palabras de los editores de la revista *Language Learning*, que en 1993 defendían la aceptación de múltiples contribuciones científicas de esta forma:

- [...] encourage the submission of more manuscripts from
- (a) diverse disciplines, including applications of methods and theories from linguistics, psycholinguistics, cognitive science, ethnography, ethnomethodology, sociolinguistics, sociology, semiotics, educational inquiry, and cultural or historical studies, to address:
 - (b) fundamental issues in language learning, such as bilingualism, language acquisition, second and foreign language education, literacy, culture, cognition, pragmatics, and intergroup relations.

Por tanto, siguiendo a Grabe (2010: 41), es defendible considerar que la LA se comporta como una disciplina que hace uso de diferentes herramientas interdisciplinares⁶ (incluidas las herramientas de la lingüística) para abordar los problemas de comunicación real.

Pero ¿qué requisitos son necesarios para definir e intentar acotar (que no delimitar o limitar) las fronteras difusas de la LA como disciplina? Kaplan y Grabe (1992: 3) parten de dos preguntas:

First, what part of linguistics can be applied to the real-world language-based problems that applied linguistics presumes to mediate? Second. What kinds of problems can be solved through the mediation of applied linguistics?

En principio, parecería que las respuestas a estas cuestiones resultan demasiado amplias, vagas o ambiguas, poco esclarecedoras incluso, para proponer una definición.

⁵ Widdowson (2000a, 2004 y 2005) no contempla la interdisciplinariedad de la LA, sino que la percibe como un campo de mediación entre el plano teórico de la lingüística y la aplicación a los problemas anteriormente mencionados. No obstante, esto contrasta con su trabajo de 1992, en el que comenta: «it is necessarily interdisciplinary because the real-world problems which AL has to tackle are not idealized so as to fit the specialist and partial interest of particular disciplines» (Kaplan y Widdowson, 1992: 80); si bien, mantiene la idea de mediación: «it mediates between theory and practice» (ibid.: 76).

⁶ Manel Lacorte (2007: 21) sostiene que uno de los rasgos distintivos de la LA es la *multidisciplinariedad* (no sabemos si establece una distinción consciente con la *interdisciplinariedad* o si usa el primer término como sinónimo del segundo). Desde nuestro punto de vista, la multidisciplinariedad fundamenta su trabajo en el estudio de un mismo objeto, que es descrito desde diferentes disciplinas (por ejemplo, el objeto de estudio *cáncer* puede estudiarse desde la medicina, la economía, la sociología, la política o la terminología; en cambio, la interdisciplinariedad es resultado de la contribución de diferentes disciplinas en el estudio (o creación) de un (nuevo) objeto con un (nuevo) método de trabajo. Es el caso de la terminología, la lingüística de corpus, la lingüística computacional, la sociología del lenguaje, la neurolingüística, etc.

Stevens (1992: 17) ese mismo año, señala seis elementos definatorios fundamentales para la LA:

1. Applied linguistics is based in intellectual inquiry, gives rise to and makes use of research, and is discipline-related;
2. Linguistics is essential to applied linguistics, but is not the only discipline that contributes to it;
3. The choice of which disciplines are involved in particular applied linguistic circumstances, and which parts of those disciplines, is contingent: it depends on what the circumstances are;
4. The multidisciplinary nature of applied linguistics requires that its practical operations be realized in a number of different domains of human activity;
5. Applied linguistics is typically concerned with achieving an end, with improving existing language-related operations, and with solving language-related problems;
6. Linguists are not exempt from being socially accountable, from displaying a social conscience, and therefore, when possible, they should use their knowledge and understanding in the service of humanity.

Estos seis puntos podrían resumirse en seis palabras clave:

- Investigación.
- Lingüística.
- Interdisciplinariedad.
- Condicionamientos particulares.
- Resolución de problemas.
- Fin social.

Por otra parte, Grabe (2010: 42-44) aumenta el número de elementos definatorios con el objetivo de defender el estatus disciplinario de la LA, la cual:

- a) Posee diferentes muestras de las disciplinas académicas: revistas científicas, asociaciones profesionales, reconocimiento internacional y autorreconocimiento individual de sus profesionales y estudiantes como lingüistas aplicados.
- b) Cuenta con congresos con áreas y subáreas científicas bien diferenciadas muy similares a los problemas que la LA tiende a resolver.
- c) Reconoce la relevancia de la lingüística en la formación de nuevos lingüistas aplicados, pero defiende la necesidad de contar con otras disciplinas.
- d) Se fundamenta en los problemas que existen en la comunicación real (principalmente con cuestiones relacionadas con el uso de la lengua, la evaluación lingüística, el contacto lingüístico y el multilingüismo, las políticas lingüísticas y la enseñanza-aprendizaje de lenguas), que son amplias y variadas,

lo que provoca que no se identifique un fin común claro entre aquellos lingüistas aplicados.

- e) Incorpora conocimientos de otras disciplinas para cumplir sus objetivos.
- f) Es un campo necesariamente interdisciplinar, puesto que los problemas a los que se enfrenta pocas veces pueden solucionarse exclusivamente desde la lingüística.
- g) Incluye un conjunto de prácticas que son resultado del trabajo llevado a cabo por diferentes lingüistas aplicados.
- h) Incorpora o incluye con regularidad multitud de subcampos: lingüística de corpus, lingüística forense, lexicografía o traducción e interpretación, entre otros.
- i) A veces, pretende definirse desde una perspectiva abarcadora para incluir cuestiones de otros campos relacionados con la lengua (alfabetización en la lengua materna, patologías del lenguaje y procesamiento del lenguaje natural).

A pesar de la mayor o menor acogida del estatus interdisciplinario y del resto de sus elementos definitorios, Grabe (2010: 41) afirma que el debate sobre la definición de la LA (y, por tanto, sobre su interdisciplinariedad), «will surely continue for at least another decade».

1.2. Objetos de trabajo y problemas

En primer lugar, hay que reconocer que casi todas las áreas de estudio del lenguaje contribuyen a la labor de la LA. Su aparición en el panorama científico de los 50 y su desarrollo hasta la actualidad es resultado de la reflexión de los científicos sobre la obligatoriedad de responder ante las necesidades del mundo real, del mundo no académico-científico (Kaplan y Widdowson, 1992: 76), idea que se resume en las siguientes palabras: «AL is a technology which makes abstract ideas and research findings accessible and relevant to the real world; it mediates between theory and practice» (idem).

Aunque ya hemos mencionado en las diferentes definiciones el objeto de trabajo de la LA, merece la pena detenerse en la siguiente cita de Kaplan y Widdowson (1992: 78) (la negrita es nuestra):

It concerns itself with how real people in real situations **achieve communicative objectives**, and why they do so. It seeks to establish the relationship between what is said and the **social roles** of the sayers, so that the **manipulations** and motivations of language users may be more clearly understood. It concerns itself with what can reasonably be taught; the circumstances under which it can be taught; the activities through which it can be taught; the political, social and economic structures that

will permit the teaching and learning to occur; and the real costs and benefits of the entire enterprise.

El problema comunicativo es el elemento central que aborda la LA, siendo los perfiles sociales de los interactantes, junto con la manipulación y la motivación de los usuarios de la lengua, otro vector relevante en las investigaciones de la LA. En efecto, nuestro objetivo tiene que ver con los problemas de comunicación (inter e intralingüística) que derivan del cambio de grado de especialización, en el que influyen notablemente la relación social entre emisor y destinatario y sus motivaciones (cognitivas, sociales y comunicativas) para manipular (no en sentido negativo) o modificar el texto que es objeto del cambio en el grado de especialización.

Por otra parte, la concepción de Kaplan (2002: 514) es que la LA opta por trabajar con nuevos datos, más que continuar con el debate o reflexión de nuevas teorías. Sin embargo, la reflexión de la LA no permanece inmanente y estática en la mediación entre teoría y práctica, sino que «is also capable of opening up avenues of theoretical study» (Kaplan y Widdowson, 1992: 80).

En resumen, la LA contribuye a estudiar (nuevos) objetos a través de nuevos datos y sus resultados pueden (aunque no es *conditio sine qua non*) arrojar luz sobre los fundamentos de nuevas teorías lingüísticas. A propósito de la forma en la que la LA trata sus problemas, Strevens establece una analogía muy ilustradora con la ingeniería:

The relation between disciplines and domains in applied linguistics is rather parallel to the relation in, for example, engineering between, on one hand, the disciplines of mathematics, physics, chemistry, and so forth, and on the other hand, the engineer's aims in given practical circumstances. [...] Applied linguistics creates a similar dynamic; depending on the purposes of the application, the disciplines relevant to those purposes will differ (Strevens, 1992: 17-18).

En definitiva, los problemas establecen los condicionantes en los instrumentos de trabajo de la LA. Pero ¿a qué problemas del mundo real se enfrenta la LA? Lacorte (2007: 20) considera que «el compromiso de analizar y solventar aspectos prácticos y reales asociados al lenguaje» se origina en el contexto de la enseñanza-aprendizaje de una segunda lengua. Aunque sus inicios estuvieron estrechamente vinculados con la enseñanza de lenguas extranjeras y los estudios sobre lingüística aplicada desde los años 50 han girado en torno a dicho tema, actualmente la lingüística aplicada abarca multitud de objetos, tantos como problemas existen. Estos problemas pueden adquirir forma de pregunta, estrategia que siguen Davies y Elder (2004: 1) para esbozar una panorámica:

How can we teach language better? How can we diagnose speech pathologies better? How can we improve the training of translators and interpreters? How can we write a valid language examination? How can we evaluate a school bilingual program? How can we determine the literacy levels of a whole population? How can we helpfully discuss the language of a text? What advice can we offer a Ministry of Education on a proposal to introduce a new medium of instruction? How can we compare the acquisition of a European and an Asian language? What advice should we give a defense lawyer on the authenticity of a police transcript of an interview with a suspect?

Otra de las formas de representar los problemas que pretende solucionar la LA es acudir a las áreas temáticas de la Association for Applied Linguistics. En palabras de Berns y Matsuda (2006: 394) (la negrita es nuestra):

The inclusiveness of AILA [Association for Applied Linguistics] is represented in a set of 25 topic areas called Scientific Commissions, a list that often serves as a ready-made definition of the field. Topics may emphasize **learner and user groups**, as in Child Language, Adult Language Learning, and Sign Language, or **analytical procedures or methods**, as in Contrastive Linguistics and Error Analysis, Discourse Analysis, and Rhetoric and Stylistics. **Issues in information transfer and interpersonal communication** are highlighted in Language and Ecology, Language and Gender, and Language and the Media. Psycholinguistics and Second Language Acquisition focus on **cognitive processes**. **Language problems in professional contexts** are associated with Forensic Linguistics, Interpreting and Translating, and Communication in the Professions, whereas other areas explicitly emphasize **instructional issues**: Literacy, Immersion Education, Learner Autonomy in Language Learning, Foreign Language Teaching Methodology and Teacher Education, Educational Technology and Language Learning, and Mother Tongue Education. Language Planning, Language and Education in Multilingual Settings, and Language Contact and Language Change represent research on the **interplay of language use, learning, and development**.

Esta lista también era uno de los criterios que usaban Kaplan y Widdowson en su trabajo de 1992 (p. 77), junto con los índices de la revista *Applied Linguistics*, la *Annual Review of Applied Linguistics* y los descriptores del diccionario usado para acceder al sistema ERIC. El contenido de estos criterios, según los mencionados autores, describiría las áreas de interés de la LA, resumibles en cinco categorías:

- Logopedia (*speech therapy*).
- Interacciones comunicativas (*communicative interactions*).
- Planificación lingüística y política lingüística (*language planning and language policy*).
- Lengua en la educación (*language in education*).
- Enseñanza y aprendizaje de la lengua (*language teaching and learning*).

Cada una de estas áreas mostraría, por tanto, diferentes aspectos del estudio del lenguaje, diferentes perspectivas, y sería fruto del trabajo interdisciplinar⁷. Calvo (1990: 324-325), por su parte, propone la siguiente división de categorías:

Aplicaciones internas al lenguaje

A1): Aplicaciones inmediatas:

- Confección de gramáticas teóricas, aplicadas y contrastivas.
- Procesos de adquisición del lenguaje por el niño (Psicolingüística y Sociolingüística).
- Aprendizaje de segundas lenguas (Didáctica y Metodología lingüísticas).

A2): Aplicaciones mediatas:

- Análisis y crítica del Lenguaje según variedades o registros: oratoria, publicidad y marketing, lenguaje coloquial, etc.
- Diseño y realización de diccionarios.
- Traducción e interpretación.

Aplicaciones externas al lenguaje

B1): Técnicas:

- Lingüística computacional (Análisis lingüístico).
- Fabricación de aparatos fonadores y sintetizadores de voz.
- Inteligencia artificial (IA) y traducción automática (TA).

B2) Terapéuticas:

- Aportaciones lingüísticas a los trastornos del lenguaje: afasia, etc. (Lenguaje y Cerebro).
- Ideación de métodos de lectura y escritura para sordomudos, etc. (Logopedia).
- Terapéutica mediante el lenguaje (Psicoanálisis).

B3): Otras aplicaciones:

- Lingüística aplicada a la confección de tests (de inteligencia, orientación profesional, etc.).
- Lingüística aplicada a idiomas artificiales (esperanto, volapuk).
- Lenguaje animal (Zoolingüística).
- Desciframiento de la escritura (bronce de Botorrita, kípús del quechua, etc.).
- Estudios antropológicos (campos del parentesco, del color...).
- Lógica y Filosofía.

En nuestro caso, el área que nos interesa, las interacciones comunicativas, queda desarrollada según Kaplan y Widdowson (1992: 77) de la siguiente forma:

b) Communicative interactions. Applied linguistics have turned their attention to the social environment in which language is used, and to the ways in which it is used by various constituencies. Careful investigations have been undertaken of the

⁷ A propósito de esta postura, Kaplan y Widdowson (1992: 77) sostienen que «[t]he very problems that the field approaches require the application of knowledge derived from linguistic theory, but also and importantly from anthropology, learning theory, psychology, sociology, and a variety of other disciplines, depending on the sort of problem being solved and including economics, as one attempts to solve problems in a cost-effective manner. Kaplan 1980 takes the view that AL is the stage upon which all of the human (social) sciences coalesce in the solution of human problems based in language».

interactions between segments of the population, interactions between physicians and their patients, between teachers and their students, interactions in the courts of law, and uses of language by advertisers, governments, and politicians [...]. These studies go well beyond providing lists of special terms used in the contexts described; rather, they examine the syntactic and rhetorical structure of different discourse types and the interaction strategies employed, particularly when the relative roles of the participants are unequal.

Cabe destacar la idea de que, según Berns y Matsuda (2006: 396), la LA no ofrece soluciones definitivas a los diferentes problemas, sino que sus resultados proporcionarían información necesaria [¿y suficiente?] para entender los problemas y sugerir opciones de resolución (por esta postura, entendemos la perspectiva de mediación entre teoría y práctica de Widdowson, 2000a, 2004 y 2005). Grabe, ya en 2010 (p. 40), relacionaba algunos problemas que pretende solucionar la LA, diferenciados en nueve subcategorías:

- Language assessment problems (validity, reliability, usability, responsibility, fairness).
- Language contact problems (bilingualism, shift, spread, loss, maintenance, social and cultural interactions).
- Language inequality problems (ethnicity, class, region, gender, and age).
- Language learning problems (emergence of skills, awareness, rules, use, context, automaticity, attitudes, expertise).
- Language pathology problems (aphasias, dyslexias, physical disabilities).
- Language policy and planning problems (status planning, corpus planning, acquisition planning, ecology of language, multilingualism, political factors).
- Language teaching problems (resources, training, practice, interaction, understanding, use, contexts, inequalities, motivations, outcomes).
- Language and technology problems (learning, assessment, access, use).
- Language translation problems (access, effectiveness, technologies).

1.3. Lingüística (teórica), lingüística aplicada y aplicaciones de la lingüística

El debate sobre la interdisciplinariedad y el estatus disciplinario de la LA ha sido una cuestión muy repetida en las páginas anteriores y la falta de consenso en los dominios y límites de la LA constituye, por lo general, la norma. Kaplan y Grabe (2000: 5-6) parten de la dificultad de acotar qué es *lingüística* como base para proponer una delimitación clara de la LA.

Entre aquellos que consideran la LA una subdisciplina o simplemente un campo de aplicación (Brown, 1987: 147) (perspectiva de inclusión) y aquellos que defienden su disciplinariedad (Kaplan, y Grabe, 1992: 3; Kaplan y Widdowson, 1992: 78; Kaplan y

Grabe, 2000; y Grabe, 2000 y 2002; entre otros) (perspectiva de autonomía), podríamos citar el esquema de Ellis (1990), que distingue entre dos modelos de investigación:

- a) El modelo teoría e investigación; un modelo deductivo.
- b) El modelo investigación y teoría; un modelo inductivo.

Ante esta distinción, Davies y Elder (2004: 10) presuponen que el modelo *teoría e investigación* se corresponde con la lingüística y que el modelo *investigación y teoría* es el modelo seguido por la lingüística aplicada. No obstante, esta división bipolar no se halla como la única opción de división. Según Berns y Matsuda (2006: 395), existen tres perspectivas en la relación entre lingüística y lingüística aplicada.

- a) En la primera, la lingüística sería la **disciplina progenitora de la lingüística aplicada**:

The process or activity of applied linguistics is carried out by taking the known research and theory of linguistics and applying a linguistic analysis to specific contexts outside linguistics proper (e.g., language teaching, interpreting and translating, or lexicography). [...] This position is taken by those whose work is influenced by a functional view of language in the tradition of Roman Jakobson, Michael Halliday, and Dell Hymes. This view assumes that only linguists can participate in applied linguistic work (ídem).

- b) En la segunda, la LA se percibe desde una **perspectiva de semiautonomía** respecto de la lingüística o cualquier otra disciplina:

[...] and allows that anyone can be an applied linguist. While acknowledging that linguistics may be part of applied linguistics, practitioners do not rely exclusively on linguistics (ídem).

- c) En la tercera, la LA se basa en la **aplicación de la lingüística**:

It is distinguished from other views in its recognition that the knowledge and skills of a linguist are inadequate to the task of solving problems related to the uses and users of language. To address this inadequacy, the applied linguist calls upon the skills and knowledge of other professionals both inside and outside the academic world (ídem).

En todas estas perspectivas, se suele recurrir a los estudios sobre sociolingüística de Halliday y Hymes en relación con la lengua en uso, así como a la importancia de la noción de *situación* de Firth, Malinowski y Halliday, factores que pueden explicar de forma acertada la razón de (inter)dependencia.

De este debate sobre la entidad disciplinaria de la LA surge el concepto de *aplicaciones de la lingüística* (*linguistics-applied* en inglés), que ha recibido más atención

en los últimos 20 años. De hecho, el manual de Davies y Elder (*The Handbook of Applied Linguistics*, 2004) se estructura en dos grandes bloques: uno para *Linguistics-Applied* (pp. 17-420) y otro para *Applied-Linguistics* (pp. 421-807), aunque Davies, ya en 1999 (p. 6) defendía una gradación entre estas *aplicaciones de la lingüística* (al mismo nivel que la lingüística descriptiva, la lingüística histórica, etc.) y la *lingüística aplicada* como disciplina autónoma. En este par, Widdowson (2000b: 5) distingue:

- a) **Aplicaciones de la lingüística:** el problema se puede reformular a través de la aplicación directa y unilateral de conceptos y términos que derivan de la propia investigación lingüística; en definitiva, es la perspectiva que sostiene que los problemas lingüísticos se pueden resolver con soluciones [¿instrumentos?] lingüísticas.
- b) **Lingüística aplicada:** disciplina que pretende conciliar diferentes perspectivas de la realidad para abordar los problemas lingüísticos.

Sobre la autopercepción del científico, Davies sintetiza su postura de la siguiente forma (si bien subraya que los límites no son claros y que pueden llegar a confluir los dos perfiles):

If they regard themselves as linguists applying linguistics because they wish to validate a theory, that is linguistics applied (L-A). If they see themselves as applied linguists because they seek a practical answer to a language problem, that is applied linguistics (A-L) (Davies, 2004: 19).

1.4. Recapitulación

En resumen, en este apartado justificamos por qué consideramos que nuestra tesis sienta sus bases sobre la LA. Desde nuestro punto de vista, esta dimensión de la disciplina lingüística, como cualquier otra, queda caracterizada por las nociones de objeto, objetivo y método, por lo que consideramos:

- a) Que nuestro **objeto de estudio** corresponde con un problema de la comunicación real. Desde una perspectiva teórica, el grado de especialización constituye un problema no solo de concepción y clasificación, sino de comunicación, puesto que la gradación no es un constructo artificial: surge por causas cognitivo-discursivas y por necesidades comunicativas en las que se debe conciliar la abstracción y la comunicación. Algunas preguntas que surgen del objeto son:

¿qué caracteriza a cada tipo de grado?, ¿por qué?, ¿somos conscientes de ello al redactar textos cuando debemos modificar el grado de especialización?

- b) En cuanto al **objetivo**, como mencionamos anteriormente en palabras de Berns y Matsuda (2006: 396), la LA no ofrece soluciones definitivas, sino información necesaria para explicar y proponer soluciones. En nuestra investigación, la clasificación de los textos según su grado de especialización, la elaboración de un nuevo modelo de análisis y la relación de características atribuibles a cada grado no ofrece soluciones definitivas al problema comunicativo que supone la gradación de la especialización, pero sí es útil a la hora de explicar su naturaleza y de elaborar estrategias de resolución de problemas.
- c) Nuestro **método** de trabajo pasa por la lingüística de corpus, la ciencia de la computación, la teoría sociocognitiva, la terminología y la lingüística del texto y los resultados pretenden servir, asimismo, como instrumento de trabajo para la enseñanza de lenguas extranjeras, la traducción e interpretación o la propia lingüística de corpus.

2. LA COMUNICACIÓN COMO OBJETO Y PROBLEMA

Debido a que nuestro trabajo se centra en el análisis del grado de especialización, es fundamental que estudiemos el concepto de *texto* y su correspondiente marco, la comunicación, desde diferentes perspectivas para poder caracterizar nuestro objeto de estudio y, posteriormente, centrarnos en ciertos elementos. En primer lugar, antes de analizar la perspectiva del texto como unidad comunicativa, nos centraremos en algunos aspectos de la comunicación.

2.1. Definiciones, propuestas y elementos definatorios de la comunicación

Tal y como señalan Alcaraz Varó y Martínez Linares (2004: 140), desde el punto de vista funcional, la comunicación es la meta última del lenguaje (Bühler, 1979 [1934]; Austin, 1962; Jakobson, 1970; y Halliday, 1985; entre otros), no en cambio desde los modelos de caja translúcida como el generativismo, que se centra en la «concepción del lenguaje como sistema idealizado en detrimento de la comunicación». Con los estudios

funcionalistas, la comunicación se convierte en un objeto de trabajo, un producto que es posible analizar.

Con esta perspectiva se superaban las teorías de la comunicación propuestas por Saussure (2007 [1916]: 60-65), en la que la comunicación constituía parte del proceso psico-físico-fisiológico del lenguaje basado en la audición-fonación de un concepto asociado a una imagen acústica; Bloomfield (1984 [1933]: 22-23), cuyo modelo reducía la comunicación a un análisis entre momentos del habla; y Shannon y Weaver (1948 y 1949), cuya teoría matemática de la comunicación introducía el concepto de ruido mas dejaba atrás factores externos.

Entre las aportaciones más importantes del funcionalismo encontramos las contribuciones a las funciones del lenguaje de Bühler, Jakobson y Halliday, que se concretan en el acto comunicativo a través del sistema (verbal o no verbal desde nuestro punto de vista). Bühler (1979 [1934]: 44-50) parte de las consideraciones de Platón en su *Cratilo*, en las que el filósofo griego expone el lenguaje como *ὄργανον* ('herramienta') «para comunicar uno a otro algo sobre las cosas» (ibíd.: 44). A partir de este esquema triangular, Bühler esboza las funciones del lenguaje (función representativa, expresiva y apelativa) que se corresponden con los objetos y relaciones que participan en la comunicación (estado de las cosas, emisor y receptor, respectivamente) (ibíd.: 48):

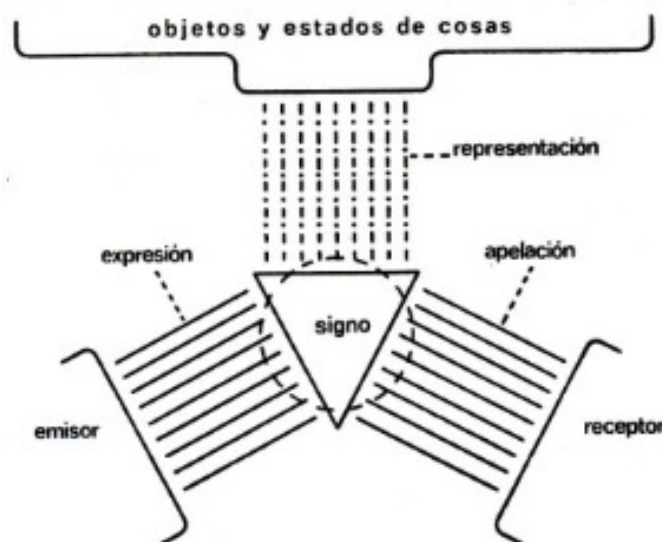


Ilustración 1: Funciones del lenguaje según Bühler (1979 [1934]: 48)

Por su parte, Jakobson (1983 [1958]: 39) completa el modelo de Bühler a partir de los factores que constituyen el proceso lingüístico: emisor, receptor, contexto, canal, mensaje y código. Con ello, deduce la existencia de seis funciones del lenguaje que se

corresponden con los factores anteriores: función expresiva, conativa, referencial, poética, fática y metalingüística. Como veremos posteriormente en nuestro trabajo, la función referencial será primordial en los textos semiespecializados, mientras que la función poética apenas tendrá protagonismo.

En tercer y último lugar, Halliday plantea la discusión de la dicotomía saussureana entre lengua y habla, desarrolla la lingüística sistémica funcional y propone un modelo que integra el componente social y cultural. En *El lenguaje como semiótica social* (1982 [1979]: 154 y 155) expone las metafunciones ideativa, que conecta al hablante con el mundo real; interpersonal, que permite el establecimiento y mantenimiento de relaciones sociales; y textual que permite la adecuación de un texto a una situación concreta.

Volviendo sobre nuestro punto de partida, según recoge Lewandowski en su *Diccionario de Lingüística* (1982: 66), «[e]n el sentido de la teoría de la información, comunicación es el intercambio de información entre sistemas retroalimentados y dinámicos/elaboradores de información, constituido por el emisor, el canal y el receptor en el proceso de la comunicación». La consideración de comunicación como sistema dinámico también se recoge en Bernárdez (1995: 137) cuando comenta las nuevas tendencias que consideran el texto como un objeto dinámico (Ballmer, 1985; Ballmer y Wildgen, 1987; Beaugrande, 1988) y señala los nuevos métodos de la ciencia para estudiar el comportamiento de sistemas naturales complejos (formados por multitud de subsistemas interactivos), dinámicos (que varían con el tiempo) y abiertos (que reciben influencia del entorno que altera su propia dinámica interna) como la comunicación⁸. Esta última característica hace imposible «predecir o describir de forma totalmente precisa» los sistemas (ibíd.: 139). Según Bernárdez (1995: 139-141), los procesos de producción y recepción del texto se pueden considerar sistemas abiertos influidos por el contexto y por el/los otro(s) participante(s) (más aún por los propios procesos de producción y de recepción). Según este autor, el proceso de producción del texto depende de factores internos y externos.

⁸ Bernárdez (1995) atribuye los adjetivos *complejo*, *dinámico* y *abierto* al lenguaje según explica en la página 138: «*Complejo*, porque está formado por la interacción de numerosos subsistemas (que a su vez son en su mayoría dinámicos y abiertos). En el lenguaje no se trata solamente de los “(sub)sistemas gramaticales” (morfología, sintaxis, pragmática, etc.), sino que abarca también al lenguaje “como herramienta” para conseguir algo, “como medio de cooperación social”, etc. *Dinámico* porque el factor “tiempo” es fundamental: tanto la producción como la recepción del texto tienen lugar en el tiempo y desde el “estado mental” de P hasta el “estado mental” de R no se produce solamente una codificación/decodificación atemporal, sino una serie de procesos sucesivos, esto es, temporales y que pueden irse modificando cronológicamente. *Abierto* porque la comunicación lingüística depende siempre de factores externos: como es sabido, la estructura de un texto no depende solamente de la gramática de la lengua, sino también, y fundamentalmente, de las características del productor, el receptor, el medio, la situación comunicativa, etc.». Estos atributos también se recogen en Zamorano (2012), junto con los términos de *atractor*, *fractal*, *oscilación*, *no linealidad* e *iteración*, que representan conceptos propios de las teorías del caos.

Los factores internos, que definen la dinámica del proceso de producción del texto, están constituidos por:

- a) El conocimiento del mundo del productor.
- b) El conocimiento de las estrategias comunicativas del productor.
- c) El conocimiento lingüístico del productor.
- d) El contenido del mensaje que el productor desee transmitir.
- e) La intención del productor al transmitir el mensaje.
- f) El mensaje como suma de contenido y de intención.

Por otra parte, los factores externos representan las influencias sobre el proceso de producción por el *sistema contexto* y el *sistema receptor* y están constituidos por:

- a) Las expectativas del productor acerca del conocimiento del mundo del receptor.
- b) Las expectativas del productor acerca del conocimiento de las estrategias comunicativas por el receptor.
- c) Las expectativas del productor acerca de la distorsión de su mensaje por el contexto, el medio, etc., lo que se asemeja a las ideas referidas anteriormente de la teoría matemática de la comunicación de Shannon y Weaver (1949).
- d) Las expectativas del productor acerca de la capacidad del receptor de salvar las distorsiones producidas por el contexto, el medio, etc.

De igual forma podríamos referir los factores internos y externos en el proceso de recepción. Con todo, este proceso de producción-recepción del texto implica cierta pérdida, por lo que, según Bernárdez, el texto que percibe el receptor no es idéntico al que emite el productor, debido a las interferencias y al aumento o disminución de precisión en los mensajes (ibíd.: 140 y 141). No obstante, desde nuestro punto de vista, y aun teniendo en cuenta que el esquema propuesto constituye un diseño muy amplio que asegura abarcar todas las posibilidades de producción-recepción, consideramos que hay ocasiones en las que sí se puede garantizar un mensaje de recepción idéntico al mensaje de producción, al menos desde el punto de vista comunicativo, aunque no desde la perspectiva puramente cognitiva.

El autor concluye con un apunte muy interesante, que aplicaremos posteriormente a nuestro trabajo: «[e]l “problema de la comunicación lingüística” consiste [...] en cómo puede saber P [productor] que un texto garantiza, en la medida de lo posible, que R [receptor] comprenderá M [mensaje] a partir de T [texto]» (Bernárdez, 1995: 142). La aportación del autor se percibe incluso más reveladora cuando se aplica a la comunicación

en contextos especializados, donde la relación entre P-R-M-T se puede llegar a ver enmascarada por la variación denominativa o conceptual. Esto es, las variaciones denominativas o conceptuales que afecten a T y, por lo tanto, a M, pueden suponer que R perciba el texto de una manera diferente a como P configuró T, es decir el conocimiento especializado codificado. Más complejo, si cabe, es intentar explicar este fenómeno en el grado de especialización.

Por último, Zamorano (2012: 690) caracteriza la comunicación lingüística basándose en el modelo de Bühler-Jakobson y propone un modelo integrado en el que la comunicación lingüística constituye un proceso no lineal y dinámico bi(multi)direccional, cuyos elementos componen una red de conexiones, cuyas unidades llevan a cabo una retroalimentación, que tiene en cuenta factores externos y cambios por la presencia de elementos específicos, que posee rasgos sustanciales de impredecibilidad y que puede generar múltiples interpretaciones. En el marco de la aproximación caológica a la comunicación verbal humana de Zamorano (2012), son relevantes los conceptos de impredecibilidad y de fractal.

En relación con la impredecibilidad, esta deviene del carácter no lineal de la comunicación verbal humana. Uno de nuestros objetivos en esta investigación consiste en aproximarnos al carácter predictivo de los textos a partir de la caracterización de los elementos más representativos del grado de especialización textual. Esta impredecibilidad está relacionada con las «alteraciones en el comportamiento estable de los elementos que intervienen en el proceso comunicativo» (Zamorano, 2012: 698), que, en nuestro caso, pueden vincularse con la relación entre el uso del léxico especializado en las situaciones de comunicación no especializadas. Este mismo fenómeno tiene relación con la noción de fractal empleada por Zamorano, que adopta de Mandelbrot (1987 [1975]: 168), que asocia la comunicación verbal humana a una combinación de irregularidad y estructura (Zamorano, 2012: 697), lo que también puede aplicarse al grado de especialización textual.

2.2. La comunicación especializada

La comunicación especializada, entendida a grandes rasgos como la comunicación que tiene lugar entre especialistas que tratan objetos temáticos concretos, surge para satisfacer las necesidades específicas de comunicación de los interlocutores protagonistas, que divergen de las observables en la comunicación no especializada.

El modelo teórico de la comunicación en el que se basaba la Teoría General de la Terminología (TGT), que puede resumirse en los rasgos de bipolaridad, linealidad, unidireccionalidad, jerarquicidad y acontextualidad (Díaz Rojo, 2000: 73) perseguía constituir un modelo comunicativo ideal para la comunicación profesional internacional, que evitase los obstáculos que causan la imprecisión, diversificación y polisemia de la comunicación no especializada (Contreras Izquierdo, 2008: 62), lo que provocó que apareciesen teorías descriptivas del modelo comunicativo real.

Así, las teorías actuales consideran que la comunicación especializada se caracteriza por el empleo de lenguas especializadas (a las que nosotros denominamos *discursos especializados*), que transmiten conocimiento especializado, definidas como registros funcionales o variedades de lengua condicionadas por criterios pragmáticos concretos, como son los interlocutores, la función o intención que pretenden cumplir, etc.; en definitiva, condicionadas por la situación discursiva. De esta forma, es posible hallar tanto semejanzas como diferencias entre la comunicación especializada y la no especializada «en los modos de significar, en la frecuencia de las funciones lingüísticas que seleccionan, en la organización y tipo del discurso, etc. (cf. Cabré, 2000: 87)» (Contreras Izquierdo, 2008: 33).

Dado que el principio fundamental en la comunicación humana es el principio de variación, cabe recurrir a la división de ejes propuesta por Hoffmann (1998 [1987]: 68) en eje vertical y eje horizontal para advertir las posibles actualizaciones discursivas de la comunicación especializada a través de la lengua, que se trata de un «macrosistema abstracto que comprende el conjunto de recursos lingüísticos de que disponen los hablantes de un idioma [...] compuesta por diversos subconjuntos que se actualizan en función de las necesidades comunicativas» (Contreras Izquierdo, 2008: 38).

a) **Variación en el eje horizontal:** hace referencia a las disciplinas en las que se puede inscribir la actualización de un discurso; en definitiva, a la temática que trata un texto. La dificultad de parcelar el eje horizontal es debida a los límites difusos entre disciplinas y la interdisciplinariedad de la que es producto el conocimiento especializado actual. De esta forma, es posible hablar, en mayor o menor medida: del discurso económico, del discurso médico, del discurso científico⁹, etc., que guardan entre sí semejanzas y diferencias.

⁹ Para hacer referencia a los tipos de discurso según el eje horizontal suele hablarse de lenguas especiales, lenguas de especialización, lenguas especializadas, tecnolectos, lenguas para fines específicos, etc. En esta investigación empleamos la denominación *discurso* para hacer referencia a cada subconjunto actualizable.

b) **Variación en el eje vertical:** hace referencia a los grados de abstracción en los que se actualiza el discurso dependiendo de la situación discursiva, que normalmente se dividen en alto, medio y bajo (véase el apartado II. 9.3., p. 144, en el que se profundiza en esta cuestión). Debido a las necesidades comunicativas especiales (Hoffmann, 1998 [1987]: 69) condicionadas normalmente por los interlocutores y la función comunicativa, es posible identificar dos ámbitos de funcionamiento diametralmente opuestos (pero entre los que es posible identificar una gradación): i) el que tiene lugar entre especialistas y ii) el que tiene lugar entre especialista y sociedad (niveles de investigación y divulgación respectivamente según Díaz Rojo, 2000: 70-72).

Según el modelo ecolingüístico de la comunicación, sería posible interpretar el eje vertical de la comunicación especializada como un *ecosistema*, «en que los distintos elementos mantienen una relación equilibrada entre sí y con el resto de elementos contextuales de carácter extralingüístico» (Díaz Rojo, 2000: 71). Los argumentos del modelo ecolingüístico sirven para cuestionar uno de los criterios teóricos de la TCT en su metodología de compilación de corpus especializados para distinguir textos especializados y no especializados¹⁰ (en relación con nuestro objeto de estudio), puesto que defiende:

El intento de definir la comunicación tomando como criterio el carácter de los participantes resulta difícil, pues no está claro si para considerar a una comunicación como especializada se debe exigir que sean expertos en la materia tanto el emisor como el receptor, solo uno de los dos, o ninguno de ellos (Díaz Rojo, 2000: 76).

Para ilustrar dicho planteamiento, Díaz Rojo (2000: 76) se sirve del siguiente ejemplo:

Por ejemplo, en un congreso o reunión sobre alimentación de niños diabéticos, el intercambio de opiniones entre un pediatra, un diabetólogo y un especialista en nutrición, daría lugar a una comunicación especializada, pues los tres pertenecen al campo de la medicina. Sin embargo, además de términos comunes, cada disciplina de las tres mencionadas tiene su propia terminología y su propio lenguaje. Por otro lado, convendría también analizar si el discurso de una persona afectada por la diabetes infantil —un padre, por ejemplo— no experta pero bien informada, es especializado o no lo es. A pesar de no ser un investigador en el tema, puede utilizar el lenguaje propio de la materia con precisión y rigor, y, por tanto, ser un usuario que domina adecuadamente la terminología específica. Y si aceptamos que su discurso es especializado, habrá que considerar si necesariamente debemos

¹⁰ Como se mencionará en apartados posteriores (apartados II. 5.4. y II. 9.3.), el principal criterio para determinar si un texto es especializado o no especializado es que el emisor de dicho texto sea un especialista.

considerarlo como un tipo de comunicación de nivel inferior, es decir, una simple vulgarización por el hecho de que el emisor no es profesional de la medicina, aunque el grado de científicidad de su lenguaje sea equivalente al de un médico.

Relacionado con esta fluctuación en el perfil de los interlocutores, Domènech (2006: 50) sostiene: «des d'un punt de vista operatiu, doncs, no és fàcil respondre a la pregunta de quins són els factors que permeten determinar que l'emissor d'un text especialitzat és efectivament un especialista en l'àmbit temàtic en què s'inscriu el text».

Así, Díaz Rojo (2000: 76-77) se inclina hacia la comunicación especializada con un planteamiento de competencia comunicativa y no de competencia profesional y descarta los criterios «formales o externos, como ocupar un puesto de trabajo docente o investigador» (ídem) con los que clasificar un texto como especializado. Se sirve, pues, de la gradualidad, la reticularidad, la no jerarquicidad, la multidireccionalidad y la contextualidad (ibíd.: 77-78) del modelo ecolingüístico para defender desde la categorización gramatical del cognitivismo que:

Cada género no se definiría en términos de un conjunto de rasgos o condiciones suficientes y necesarios, sino más bien de unos rasgos prototípicos. De esta manera, la pertenencia de un elemento a una categoría se establece a partir del grado de similitud con el prototipo, sin necesidad de que todos los miembros de una clase posean todas las características de aquel (ibíd.: 78).

Así, a nuestro juicio, además de los rasgos lingüísticos (véase apartado II. 9.2.3., p. 139), los principales condicionantes del grado de especialización son la relación discursiva entre los interlocutores y la función textual que cumple el texto.

2.3. Relación discursiva entre los interlocutores

Cuando se pretende estudiar el grado de especialización textual, es necesario recurrir a todas las posibilidades de actualización discursiva, lo que obliga a evitar la clasificación binaria de especialista y no especialista y recurrir, de nuevo, a la gradación de los interlocutores según su competencia cognitiva-comunicativa. Dicho de otra forma, usando como criterio las condiciones de conceptualización discursiva (desde el emisor o desde el destinatario), podemos aproximarnos a la clasificación textual según el grado de especialización.

Bajo nuestro punto de vista, el destinatario del texto constituye el punto de partida de la configuración discursiva. Así, puede afirmarse que el grado de especialización de un

texto depende del perfil de su destinatario (Domènech, 2006: 72), pues, en definitiva, es el eje principal del rasgo de aceptabilidad (Beaugrande y Dressler, 1997: 35-45): en la comunicación especializada, el destinatario espera información relevante relacionada con el conocimiento especializado, si bien esta información se ajusta al conocimiento del destinatario.

De esta forma, desde una perspectiva cognitiva, la complejidad (cognitiva) del texto se configura según la información que comparten emisor y destinatario, lo que implica la aparición de necesidades de adaptación discursiva por parte del emisor (Domènech, 2006: 65), que debe adaptar el producto textual al nivel de conocimiento del destinatario con el fin de garantizar la constitución de un texto comunicativo. En palabras de Lundquist (2003: 17), existiría un proceso en el que participa «people knowing more conveying knowledge to people knowing less».

Con todo, es posible recurrir una vez más a las palabras de Bernárdez (1995: 142) (*vid. supra* apartado II. 2., p. 37) que plantea que «[e]l “problema de la comunicación lingüística” consiste [...] en cómo puede saber P [productor] que un texto garantiza, en la medida de lo posible, que R [receptor] comprenderá M [mensaje] a partir de T [texto]».

Domènech (2006: 72) se sirve de la teoría cognitiva (en concreto, de los modelos idealizados) para intentar resolver dicho problema:

Des del punt de vista de les seva caracterització cognitiva, és un fet prou fonamentat que l'emissor adapta el contingut del text que produeix al domini cognitiu idealitzat especialitzat que atribueix al seu receptor, de manera que podem afirmar que la densitat de coneixement del text és directament proporcional al seu nivell d'especialització. Aquesta densitat de coneixement l'entendem tant des del punt de vista quantitatiu (un text és més dens com més quantitat de coneixement transmet, quantitat que es pot entendre, per exemple, com el nombre de conceptes o nusos especialitzats que transmet), com des del punt de vista qualitatiu (un text és més dens com més estructurada i pertinent és la informació que transmet), perquè de fet, les dues dimensions, qualitativa i quantitativa, es condicionen recíprocament.

De esta forma, es posible matizar que el punto de partida del grado de especialización no es el receptor en sí, sino el grado de conocimiento que el emisor del texto presupone que posee el receptor. Así, el criterio que guía nuestra clasificación es la teoría sobre descripción sociolingüística de Hymes en torno a la competencia comunicativa.

2.4. Competencia cognitivo-comunicativa de los interlocutores

Hymes (1996 [1972]) incorpora el componente de la interacción social a la teoría sobre competencia y actuación de Chomsky para defender una postura que se centre en la competencia comunicativa de los hablantes, y no tanto en la competencia lingüística como conocimiento tácito e inconsciente de las estructuras de una lengua. Esta competencia se constituiría como un *continuum* de variedades, de las más restringidas a las menos restringidas (1996 [1972]: 19-20), e incluiría «habilidades y criterios que son relativos a, e interdependientes con, factores socioculturales» (1996 [1972]: 21-22).

De esta forma, para poder construir correcta y adecuadamente los discursos especializados (y sus variantes), los usuarios deben contar con una competencia que depende «del conocimiento (tácito) y del uso (habilidad para éste)» (Hymes, 1996 [1972]: 26-27). Este conocimiento se extiende a cuatro parámetros de comunicación: posibilidad de formación, factibilidad en el discurso, adecuación en relación con la situación y efectos de la ejecución de la comunicación (Hymes, 1996 [1972]: 20-21). En este sentido, un usuario cuenta con el conocimiento y habilidades necesarias sobre los elementos de los sistemas comunicativos de que puede disponer y es capaz de interpretar o evaluar la conducta discursiva de los usuarios según los parámetros mencionados (Hymes, 1996 [1972]: 26). A nuestro juicio, el conocimiento no solo tiene carácter discursivo-gramatical sino también enciclopédico, es decir, cognitivo, en el sentido de que en los discursos especializados y sus variantes es necesario controlar los sistemas conceptuales que se tratan.

La TCT, en su posición teórica para analizar y compilar textos especializados, señala la necesidad de que el emisor de un texto siempre sea especialista¹¹, de forma que el campo conceptual esté controlado. Por tanto, los niveles de gradación del texto especializado dependerían del perfil del receptor (Domènech, 2006: 67), normalmente dividido en especialista, público formado, pero no especialista o aprendiz de especialista y público lego en general¹² (Freixa, 2002b: 181). No obstante, nuestro objetivo en esta tesis doctoral es estudiar el grado de especialización, incluidos los textos no especializados, por lo que, en principio, los perfiles según la competencia cognitivo-comunicativa de los interlocutores son aplicables tanto al emisor como al destinatario del texto.

¹¹ Domènech (2006: 50) defiende que el perfil del emisor no es la única característica pragmática de los textos especializados, pero sí la más determinante.

¹² Pérez (2002) por ejemplo, distingue las categorías de especialista, iniciado, semiespecialista, lego, profesor y alumno.

Así, los usuarios configuran los discursos atendiendo a dichos parámetros y habilidades propias y de sus destinatarios, y activan diferencias (Hymes, 1996 [1972]: 35) según la competencia comunicativa (y también cognitiva) del destinatario como pueden ser a) la presencia o ausencia de un rasgo; b) el valor semántico asignado a un rasgo, c) la distribución de rasgos entre contextos, o d) las interrelaciones entre estos y cada uno de ellos en los usos marcados y no marcados. Ejemplos de esto podrían ser la presencia o no de reformulaciones en un texto, que podría (pero no necesariamente) estar relacionada con la falta de conocimientos sobre el tema de los que dispone el destinatario, o el valor que adquiere una unidad léxica como *fuerte* en un discurso especializado sobre informática vinculado a la unidad *variable*.

De esta forma, desde nuestro punto de vista, se distinguen tres categorías principales de interlocutores según su competencia comunicativa y cognitiva, subdivididos cada uno en perfiles concretos.

a) **Especialista:** interlocutor con motivación profesional que, normalmente, posee formación académica y profesional en la materia y es capaz de llevar a cabo reflexión y argumentación sobre el conocimiento que domina, lo que le permite (aunque no siempre) producir conocimiento especializado nuevo (Cabré, 2002b). Dicho conocimiento no forma parte del conocimiento común que comparte el resto de hablantes y ha sido adquirido a través de un proceso de aprendizaje sistemático y voluntario, lo que implica un uso de la terminología más preciso que el empleado por un no especialista. Es el perfil de los investigadores y profesionales en general: por ejemplo, un abogado en un artículo científico:

- (1) Por su parte, el estudio atento de la jurisprudencia y de la doctrina administrativa pone de manifiesto que, al margen del distinto origen del evento que produce el desequilibrio económico del contrato, los requisitos para la aplicación de la doctrina del factor principis no son muy diferentes a los que son exigidos en relación con la aplicación doctrina del riesgo imprevisible (Gallego, 2016: 119)¹³.

Cabe defender que pueden existir especialistas que hayan adquirido conocimiento especializado de forma consciente y voluntaria y, aun así, no posean formación académica y profesional (reglada, es decir, que cuente con una titulación universitaria en el tema), así como usuarios que habiendo sido formados académica o profesionalmente no cuenten con las habilidades para

¹³ Como se mencionó en el apartado I. 5., la información bibliográfica de las fuentes que se empleen como ejemplo para ilustrar casos reales se relaciona como nota al pie para facilitar la identificación de datos como el autor, el soporte, el año, el tipo de texto en el que aparece la información, etc. Fuente: GALLEGO, Isabel (2016): «Circunstancias externas y reequilibrio económico del contrato público. ¿Nuevas perspectivas?», *Revista digital de derecho administrativo*, p. 119. Disponible en: <http://revistas.uexternado.edu.co/index.php/Deradm/article/view/4590/5278>.

crear conocimiento especializado nuevo (y, por tanto, no sean especialistas). El perfil de especialista puede subdividirse a su vez en a) el **especialista investigador** (que genera, en mayor o menor medida, conocimiento especializado nuevo) y en b) el **especialista no investigador** (que no genera conocimiento especializado nuevo).

b) **Instruido**¹⁴: interlocutor con motivación profesional o personal que domina en gran medida la materia que trata, pero no tiene formación académica o profesional en ella, sino que ha sido fruto de procesos de documentación e información complementarias, por lo que no suele ser capaz de crear conocimiento especializado novedoso ni de usar la terminología con total precisión. Es el perfil que podría adquirir un mediador lingüístico como un periodista o un traductor o aquel individuo que por razones individuales o sociales muestra un interés activo por el tema en cuestión: por ejemplo, un periodista especializado en nutrición en una entrevista para un blog de divulgación especializada:

(2) La primera vez que oí hablar de la dieta Dukan reconozco que lo primero hice fue ponerme a la defensiva y pensar que no debía tratarse de nada bueno –nunca he sido muy amigo de las dietas y los “milagros”– y por supuesto soy un escéptico de cuidado. Pasé del tema y ahí se quedó la cosa. Mi opinión sobre la dieta Dukan quedó fundamentada en un prejuicio y en lo que mis colegas de la comunidad médica iban diciendo por ahí, datos que no se me ocurrió contrastar con la bibliografía¹⁵.

Los instruidos pueden clasificarse, a su vez, en a) **semiespecialistas**, es decir, profesionales de ramas relacionadas; b) **aprendices de especialista**, normalmente alumnos universitarios o egresados en formación; y c) **legos formados**, generalmente individuos que, por motivación personal, su interés por el tema les ha llevado a poder comprender el sistema conceptual. Es posible que existan instruidos que se conviertan en especialistas, según el nivel de competencia cognitivo-comunicativa que adquieran.

c) **Lego**: interlocutor con o sin motivación personal que no posee formación académico-profesional en la materia ni tampoco es capaz de crear conocimiento especializado nuevo. Usa la terminología con imprecisión, a veces como meras

¹⁴ El instruido no especialista recibía el nombre de *experto* en las publicaciones de Rodríguez-Tapia (2015, 2016a y 2016b) y cambió su denominación, más descriptiva a nuestro juicio, en Rodríguez-Tapia y Camacho-Cañamón (2018).

¹⁵ Fuente: Miguel MC (3 de noviembre de 2012): «Toda la verdad sobre la dieta Dukan (entrevista al doctor Álvaro Campillo)», *PARADIGMedical. Curiosa forma de ver la medicina*. Disponible en: <https://paradigmmedical.wordpress.com/2012/11/03/toda-la-verdad-sobre-la-dieta-dukan-entrevista-al-doctor-alvaro-campillo/>.

formas denominativas sin conocimiento de su carga semántica. Suele ser el perfil que se vincula con el público en general, que es objeto de la divulgación de masas y suelen ser variados en competencia cognitiva y comunicativa: por ejemplo, un estudiante de secundaria en un trabajo de clase o una persona que realiza una consulta en un foro.

- (3) Hola buenas tardes el día 2 de agosto sufrí una caída de escaleras golpeándome con el filo la región lumbar muy cerca de la columna, inmediatamente se me inflamó el área y desde entonces tengo dolor al tacto y sigo teniendo inflamado esa misma área como un bulto que es doloroso al toque, fui al médico general y me dijo que es un esguince de cadera me mandó 3 inyecciones de B12 con lidocaína y dexametazona, meloxicam con metocarbamol y aplicar una crema de diclofenaco pero ya llevo más de una semana y aunque eh visto algo de mejoría la inflamación no cede, sigo teniendo la bolita. No me molesta al caminar ni tengo debilidad en manos ni pies solo la espalda alta siento como cansancio y hormigueo como ya dije solo el área me molesta al tacto, al recargarme o acostarme boca arriba. Será algo serio? tengo 31 años soy varón no eh tenido temperatura alta ni nada de eso¹⁶ [se han mantenido los errores ortográficos y gramaticales del original].

En síntesis, la siguiente tabla resume las categorías y subperfiles de los interlocutores aquí expuestos:

Especialista	Investigador
	No investigador
Instruido	Semiespecialista de otra disciplina
	Aprendiz de especialista
	Lego formado
Lego	Lego no formado

Tabla 1: Perfiles y subperfiles de tipos de interlocutores según su competencia cognitivo-comunicativa

La propuesta de división y subdivisión de los perfiles de emisores y destinatarios según la competencia cognitivo-comunicativa toma como marco la teoría de prototipos y del nivel básico de la lingüística cognitiva (Rosch, 1978). En este sentido, las categorías¹⁷ que proponemos se organizan de dos formas: (i) de forma horizontal (según la teoría de prototipos), en la que una categoría está formada por entidades más o menos prototípicas¹⁸ o ideales y más o menos periféricas en torno al prototipo, con límites claramente difusos y no delimitados; y (ii) de forma vertical (según el nivel básico), de lo más general a lo más específico.

¹⁶ Fuente: Emmanuelortega (14 de agosto de 2016): «Caída de escalera bulto en zona lumbar», *Portalesmedicos.com*. Disponible en: http://www.portalesmedicos.com/foros_medicina_salud_enfermeria/ubbthreads.php/posts/195227/Caida_de_escalera_bulto_en_zon.

¹⁷ Las categorías serían los resultados de la categorización, un mecanismo de organización de la información fruto de la experiencia que permite simplificar la realidad a partir de los procedimientos de generalización y discriminación (Cuenca y Hilferty, 1999: 32).

¹⁸ El prototipo sería el objeto más representativo de la categoría, que guarda similitudes con el resto de miembros de la categoría (Rosch y Mervis, 1975).

Desde este punto de vista, los límites que existen entre los diferentes subperfiles no están acotados *a priori* por un conjunto de principios que rijan el inicio o el fin de la categoría, sino que es posible que cada subperfil comparta características con sus vecinos. En nuestro caso, cada perfil corresponde con una categoría compuesta por características actualizables en el discurso prototípicas y periféricas, que dependen de la percepción de los hablantes. Por lo que con esta propuesta de perfiles en subcategorías (las que aparecen en la columna de la derecha de la Tabla 1) pretendemos conocer el elemento prototípico para la supercategoría de la columna de la izquierda y dónde el conjunto de informantes sitúa los límites.

2.5. La función lingüística

De acuerdo con la función lingüística que adquieren los textos del grado de especialización, puede afirmarse que no suelen materializarse todas las funciones lingüísticas propuestas por Jakobson (Domènech, 2006: 50), sino que, a nuestro juicio, la función referencial es la que tiene un mayor protagonismo.

De hecho, Ciapuscio y Kuguel (2002: 45) proponen un nivel funcional en su modelo de análisis tipológico textual. Este nivel funcional se construye según la noción de función textual de Halliday y «se concibe como el efecto de los textos en el contexto de la interacción social» y se articula en torno a cuatro funciones básicas: informar, expresarse, dirigir y contactar. A nuestro juicio, dichas funciones se asimilan al modelo de Jakobson con las respectivas funciones: referencial, expresiva, apelativa y fática. En el corpus del estudio que estas investigadoras realizan, identifican una función principal (la de informar) a partir de indicadores lingüísticos diversos y otra complementaria (la de dirigir) a partir de indicadores léxicos.

Estos resultados pueden avalar la afirmación de la que partíamos al principio: la función referencial es la que ostenta mayor protagonismo (*vid. supra* apartado II. 2.1.). Ahora bien, es posible defender que la función referencial de Jakobson está presente en casi todos los textos, puesto que relaciona el texto con la representación de la realidad. Por ende, debemos encontrar una forma de diferenciar los textos según sus funciones principales y complementarias. Si usamos para ello los resultados del estudio de Ciapuscio y Kuguel (2002), es posible moverse entre dos funciones principalmente: la referencial y la apelativa. Desde nuestra perspectiva, esta función referencial en el grado de especialización puede

graduarse de la misma forma, y distinguir tres funciones principalmente, según si el texto tiende a la representación del objeto que pretende comunicar o si el texto se inclina hacia la comunicación del conocimiento sobre ese objeto. De esta forma, el grado funcional de referencialidad a la realidad está sujeto a cómo el contenido se aproxima en mayor o menor medida a la función apelativa:

- a) Una **función representativa**: resultado de la mayor abstracción en el tema tratado. Con esta función, el texto centra su objetivo en dibujar un esquema conceptual que exponga el objeto que trata. La comprensión de un texto con función principalmente representativa queda fuera del usuario medio de la lengua. Suele ser la función que cumplen los artículos científicos o las tesis doctorales, o la que se activa en textos entre especialista y especialista.
- b) Una **función mixta** representativa-comunicativa, que puede recibir el nombre de **función didáctica** o **función instructiva**: el tema es tratado desde un enfoque especialista pero el discurso sufre una recontextualización con estrategia comunicativa (reformulaciones, explicaciones, aclaraciones, etc.). El usuario medio de la lengua es capaz de comprender el texto, aunque, a veces, con serias dificultades por desconocimiento del sistema conceptual del ámbito temático. Es la función que cumplen ciertos textos orientados a la iniciación en alguna disciplina o a la enseñanza-aprendizaje de cuestiones relacionadas con las mismas.
- c) Una **función comunicativa**: que presta atención a la transmisión de información, más que a la elaboración de un esquema conceptual riguroso o exhaustivo. Los textos con función primordialmente comunicativa son comprensibles para el usuario medio de la lengua, si bien algunos conceptos pueden tratarse con cierta laxitud conceptual. Suele ser la función observable entre un especialista y un lego o en guías y folletos para el usuario no especialista.

La TCT defiende las funciones representacional o representativa y comunicativa de la terminología (Cabré, 1999: 70, 80, 274), si bien hay que resaltar que en todos los textos del grado de especialización aparecen de forma simultánea las mencionadas funciones lingüísticas, por lo que es posible distinguir la interrelación entre ellas: en todos los textos se pretende representar un objeto de la realidad, comunicar información sobre ese objeto y, en definitiva, instruir mediante ambas funciones. Por tanto, hablamos de que, clasificar los textos según su función lingüística puede hacerse atendiendo a la función que *prima* sobre

el resto y no necesariamente a la *única* función observable. En definitiva, los límites para la distinción de estas tres funciones no son discretos.

Estas funciones pueden distinguirse a nivel lingüístico por el tipo de procedimientos de reformulación que se observan (lo que indica la estrategia discursiva del emisor según el conocimiento que presupone del destinatario; aunque no todos los procedimientos de reformulación suponen una estrategia para solucionar la falta de conocimiento del destinatario) y a nivel extralingüístico por el tipo de texto y los interlocutores que participan en la comunicación.

Con todo, en investigaciones anteriores en las que se aplicó aprendizaje automático a una base de datos similar, pero más reducida, a la que explotamos en esta tesis doctoral, se advirtió que los valores de la función lingüística no resultan relevantes para la discriminación de clases (Rodríguez-Tapia y Camacho-Cañamón, 2017, 2018), especialmente si se atiende a las complicaciones teórico-metodológicas que entrañan.

3. EL TEXTO

3.1. Definiciones y propuestas

Una vez hemos analizado el concepto de comunicación, de la perspectiva de comunicación como sistema abierto y de las funciones del lenguaje, procederemos a estudiar los trabajos realizados por Loureda (2002), Bernárdez (1982), Chareaudeau (1992), Beaugrande y Dressler (1997), Dijk (1992 [1983]) y Zamorano (2012) acerca del concepto de *texto*, ya que consideramos que estos autores aportan una perspectiva integradora, comunicativa y funcional del mismo, cuyos aspectos más relevantes sintetizaremos al final de este apartado. La necesidad de detenernos en el análisis del objeto textual se halla en la descripción explicativa que proporciona el contexto supraoracional sobre la estructura y función del texto en las interacciones comunicativas (Ciapuscio, Kuguel y Otañi, 2005: 95).

Loureda (2002: 158) parte del enfoque dimensional coseriano para proponer que los textos cuentan con un aspecto individual («aquello por lo que un texto es ese texto y no otro») y una dimensión universal, que consiste en «unas propiedades comunes a todos ellos; entre otras, el ser dichos (o escritos) a alguien mediante una lengua por un emisor en una circunstancia dada».

Desde nuestro punto de vista, la concepción de texto de Bernárdez es una de las más abarcadoras del objeto. Según este autor, el texto:

[...] es la unidad lingüística comunicativa¹⁹ fundamental, producto de la actividad verbal humana, que posee siempre carácter social; está caracterizado por su cierre semántico y comunicativo, así como por su coherencia profunda y superficial, debida a la intención del hablante de crear un texto íntegro, y su estructuración mediante dos conjuntos de reglas: las propias del nivel textual y las del sistema de la lengua (1982: 85).

La definición que propone integra factores comunicativos, pragmáticos (que tienen relación con la intención del hablante y la situación) y estructurales (que se refieren a la parte interna del texto). A partir de esta definición podemos determinar que los puntos que nos interesa analizar en nuestro trabajo son, principalmente, el carácter social del texto (de hecho, como veremos más adelante, muchas clasificaciones parten de la relación emisor-destinatario), la autonomía semántica y comunicativa y la coherencia profunda y superficial. A esto hace referencia Cabré (2002b: 18) cuando defiende que un texto «alcanza su meta en unas determinadas condiciones de producción, transmisión y recepción cuyo conjunto puede determinar su adecuación». No obstante, cuando Bernárdez indica que se trata de un «producto de la actividad verbal humana», consideramos que deja fuera de su definición el aspecto de texto como manifestación semiológica (o al menos no aparece en la definición de forma patente) que sí está presente en la propuesta del autor francés Charaudeau: «[l]e texte est la manifestation matérielle (verbale et sémiologique: orale/graphique, gestuelle, iconique, etc.) de la mise en scène d'un acte de communication, dans une situation donnée, pour servir de projet de parole d'un locuteur donné» (1992: 645).

Es por esta perspectiva integradora por lo que dejamos fuera de nuestro trabajo la contribución de Álvarez (2004 [1996]: 1) que considera que «el texto es una configuración lingüística. Es un conjunto de elementos lingüísticos (palabras, oraciones...) organizados según reglas estrictas de construcción». Además de esto, el autor distingue, desde el punto de vista terminológico, entre *texto* y *discurso* al afirmar que «el discurso es la emisión concreta de un texto, por un enunciador determinado, en una situación de comunicación determinada» (ibíd.: 1) y que «el texto funciona como discurso en una situación

¹⁹ La concepción de texto como unidad comunicativa ya se tenía en cuenta en Hartmann (1971: 17) o Kallmeyer *et al.* (1974), así como en Hoffmann (1990: 6): «Der Text ist eine thematisch relativ geschlossene, gegliederte, strukturierte und in sich kohärente komplexe sprachliche Einheit, die im Rahmen sozialer Handlungszusammenhänge eine oder mehrere kommunikative Funktionen zu erfüllen hat».

determinada» (ibíd.: 2), al igual que distingue Dijk (1980: 32) al sostener que el texto es un concepto abstracto que se manifiesta en discursos concretos (Bernárdez, 1982: 87): «este término [texto²⁰] se usará aquí para denotar la construcción teórica abstracta que subyace a lo que normalmente se llama un discurso. Aquellas expresiones a las que puede asignarse estructura textual son, pues, discursos aceptables de la lengua»²¹.

A las definiciones anteriormente mencionadas se añaden los criterios²² desarrollados por Beaugrande y Dressler (1997: 35), por los cuales «un texto es un acontecimiento comunicativo que cumple siete normas de textualidad. Si un texto no satisface alguna de estas normas no puede considerarse que ese texto es comunicativo». Es importante este último apunte, que especifica que el texto no puede considerarse *comunicativo*, lo que no implica que deje de considerarse *texto*. Beaugrande y Dressler estructuran estas normas de textualidad en dos grandes bloques. El primer bloque reúne los principios reguladores, entre los que se incluyen la eficacia, la efectividad y la adecuación (ibíd.: 46):

- a) La eficacia «depende de que dos participantes empleen o no un mínimo de esfuerzo en una intención comunicativa».
- b) La efectividad «depende de si genera o no una fuerte impresión en el receptor y si crea o no las condiciones más favorables para que el emisor pueda alcanzar la meta comunicativa que se había propuesto».
- c) La adecuación «depende de si se establece o no un equilibrio entre el uso que se hace de un texto en una situación determinada y el modo en el que se respetan las normas de textualidad».

El segundo bloque reúne los principios constitutivos (ibíd.: 35-45), que describimos en los siguientes apartados:

- a) La cohesión «establece las diferentes posibilidades en que pueden conectarse entre sí los distintos componentes de una superficie textual dentro de una secuencia determinada».

²⁰ Para ahondar en la investigación del concepto de texto, véase el capítulo 3 «El concepto de texto» en Bernárdez (1982: 75-100).

²¹ No obstante, en publicaciones posteriores del autor neerlandés, se sostiene: «[e]n el trabajo previo, en consecuencia, propusimos distinguir sistemáticamente entre el texto como objeto formal y lo que realmente ocurre en el (los) discurso(s). Aquí no deseamos realizar esta distinción, sino simplemente utilizar indistintamente el texto y el discurso. No obstante, más adelante se da por entendido que el texto o el discurso puede tener propiedades generales, abstractas o libres de contexto, que pueden explicarse mediante cierto tipo de gramáticas del discurso, y propiedades que varían a través de los diferentes contextos (las situaciones, los hablantes, etc.) en la misma cultura» (Dijk, 1990: 45).

²² El modelo de Beaugrande y Dressler no constituye el único modelo de textualidad aunque ha servido de referencia para construir otros, como por ejemplo los cinco módulos que propone Adam (2001: 21).

- b) La coherencia «regula la posibilidad de que sean accesibles entre sí e interactúen de un modo relevante los componentes del mundo textual, es decir, la configuración de los conceptos y de las relaciones que subyacen bajo la superficie del texto». Este principio debe respetar el saber elocutivo del lenguaje (Coseriu, 1992 [1982]: 106 y ss.).
- c) La intencionalidad está íntimamente ligada con los dos principios anteriores. Se refiere a la actitud del emisor del texto. Tiene que ver con la transmisión del conocimiento y el objetivo de alcanzar una meta dentro de un plan.
- d) La aceptabilidad «se refiere a la aptitud del receptor». Un texto es aceptable si el destinatario percibe que tiene alguna relevancia, entendida como adquisición de conocimiento nuevo o cooperación con su interlocutor en la consecución de una meta discursiva determinada.
- e) La informatividad «sirve para evaluar hasta qué punto las secuencias de un texto son predecibles o inesperadas, si transmiten una información conocida o novedosa».
- f) La situacionalidad se refiere a los factores que hacen que un texto sea relevante en el contexto en que aparece.
- g) La intertextualidad «se refiere a los factores que hacen depender la utilización adecuada de un texto del conocimiento que se tengan de otros textos anteriores».

Por otra parte, Dijk, en *La ciencia del texto* (1992 [1983]) comenta la estructura textual en tres niveles.

El primer nivel, la **superestructura**, constituye el esqueleto formal y la estructura global que caracteriza el tipo de un texto. En palabras del autor neerlandés «una superestructura es un tipo de forma del texto, cuyo objeto, el tema, [...] es el contenido del texto» (ibíd.: 142). De igual forma, la superestructura determina el orden global del texto (ibíd.: 143) (en la superestructura encontraríamos la explicación teórica de las recomendaciones de encabezamientos descritas por las normas de Vancouver para la publicación de trabajos científicos sobre medicina, por las que reconoceríamos que el artículo se corresponde con un tipo de forma de texto y no con otro (ICMJE, 2010: 11-15)). En otras palabras, la superestructura reúne todas las características de la organización de la información en el texto (que a su vez viene determinada por el tipo de texto, que impone el tipo de estructura, regida por la forma o por la presentación de la información).

En segundo lugar, encontramos la **macroestructura**, que se manifiesta mediante el tema y los subtemas del texto y que constituye la unidad semántico-conceptual del texto

(de hecho, Dijk se refiere a ella como «macroestructura semántica» y se encarga de estudiar este nivel desde la semántica). Desde este nivel podemos, por ejemplo, determinar que un texto constituye una narración mediante el análisis de su progresión temática. En otro orden de cosas, tanto la superestructura como la macroestructura configuran los textos y comparten la característica de no poder ser definidas con relación a secuencias aisladas del texto, sino que se definen en su conjunto o según determinados fragmentos del mismo (ibíd.: 142). No obstante, la constitución de la superestructura depende directamente de la organización de la macroestructura (y viceversa; ibíd.: 70).

En último lugar, Dijk distingue el nivel de la **microestructura** (ibíd.: 143), establecida en el nivel de las oraciones, los sintagmas y las palabras. En este nivel son analizables las relaciones de cohesión, las variedades funcionales y la modalidad que corresponda al valor informativo de las proposiciones. Mediante el análisis del contenido de este nivel se puede llegar a establecer la macroestructura (ibíd.: 59).

Aunque desde nuestra perspectiva consideramos el texto como un todo, como un bloque intercondicionado y multidimensional, el nivel microestructural y superestructural (como en los trabajos de Rodríguez Tapia, 2015 y 2016a) tendrán gran relevancia en la metodología y discusión de datos de esta investigación.

Una vez hemos analizado las diferentes propuestas de estudios del texto, determinamos que en el presente trabajo adoptaremos la posición de texto definida por Bernárdez (1982: 85), matizada por la manifestación semiológica del texto que contempla Charaudeau (1992: 675), enmarcada en torno a las normas de textualidad de Beaugrande y Dressler (1997: 35) y en torno a los niveles de estructuración textual propuestos por Dijk (1992 [1983]) y caracterizada por los rasgos de impredecibilidad y diversa interpretación de Zamorano (2012) referidos a la comunicación verbal humana (*vid. supra*) (Rodríguez-Tapia, 2016a: 988), así como acotadas por las dimensiones coserianas que propone Loureda (2002).

3.2. Los tipos textuales

Los textos constituyen unidades informativas complejas y abiertas que están condicionadas de forma discursiva y social (Cabré, 2002a: 92 y 93). Al agrupar los textos en conjuntos se siguen diferentes criterios, a veces guiados por el estereotipo social, funcional, conceptual o estructural (ídem), que constituyen los denominados *tipos textuales*

o *tipos de texto* establecidos «a partir de matrices de criterios discursivos y textuales, la mayoría graduales o graduables» (ídem).

Los tipos textuales no son establecidos universal ni permanentemente (Cabré, 2002a: 93), sino que se trata de construcciones socioculturales (*culturales* según Loureda, 2002: 159), que dependen de las condiciones de producción-transmisión-recepción de la información (Cabré, 2002a: 93), lo que lleva al surgimiento de nuevos tipos textuales (a través de adaptación o recomposición de tipos ya establecidos²³). Así, «representan modelos estructurales y estilísticos para las necesidades comunicativas de los hablantes» (Loureda, 2002: 158) y, por tanto, son respuesta a necesidades expresivas de ambos interlocutores (ibíd.: 159). De esta forma, «se presenta como horizonte de expectativas para el receptor o destinatario» con el objetivo de garantizar la funcionalidad del hablar (ídem).

3.2.1. Utilidad teórica de los tipos textuales

Dentro de la lingüística textual, uno de los objetivos más perseguidos ha sido la elaboración de propuestas de clasificación de los tipos de texto. La necesidad de diseñar las tipologías textuales pretende, según Isenberg (1987: 98):

- Solucionar problemas básicos de la lingüística del texto.
- Solucionar problemas de la gramática.
- Determinar el ámbito de validez de las regularidades (principios, máximas, reglas, normas), que son fundamento de la producción y recepción de textos.
- Analizar textos concretos.
- Describir la intertextualidad.
- Determinar la intertextualidad de los textos.

Además, Isenberg (1987: 96-97) distingue dos grandes bloques de trabajos dentro del marco de la lingüística del texto. En el primer grupo se incluyen aquellos trabajos que tratan las propiedades generales de los textos, y en el segundo, aquellos estudios que tratan tipos concretos seleccionados de textos (sobre todo, la narración). En este segundo grupo, según el mencionado autor (1987: 97):

[e]xiste una relación directa entre lo que se quiere describir (una clase determinada de texto) y la cuestión sobre la tipología textual, dentro de cuyo marco esta clase

²³ Loureda considera, por tanto, que los tipos textuales poseen su propia historia: «los tipos de texto nacen (se crean), crecen (se modifican o evolucionan), se reproducen (se generan unos a partir de otros) y finalmente mueren o desaparecen debido a la modificación extrema o al agotamiento del modelo» (2002: 159).

de texto se define como tipo de texto y se delimita frente a otros tipos de texto, o se ordena como subtipo dentro de un tipo de texto.

Loureda (2002: 161) por su parte, considera que la clasificación de los géneros permite conocer la aplicabilidad de las reglas para la producción y recepción de los discursos a los (o algunos) actos de habla y, a nuestro juicio el aspecto más productivo en nuestro estudio (en el sentido de la automatización) «hacer predicciones sobre la organización y el contenido de los textos a partir de sus estructuras».

La motivación principal para la división en tipos textuales (es decir, la propuesta de tipologías textuales) se encuentra precisamente en la instrumentalización que permite la predicción. A través de la abstracción de cualidades textuales («los tipos de texto equivalen a representaciones ideales formadas por el común de una serie de discursos»; Loureda, 2002: 158) se logra conseguir un atributo abstracto vinculado con el texto (el tipo textual), que permite servir de modelo para otros objetos textuales reales, puesto que estos últimos guardan similitud en sus atributos (idem)²⁴.

3.3. Conceptos en torno al texto: *clase, tipo, género y secuencia*

La terminología de la lingüística textual está sujeta a variación denominativa y conceptual. Esto afecta a conceptos como *género, tipo, texto, clase*, etc., que sufren constante restricción y ampliación semántica según la escuela o tradición investigadora de los autores. Aunque la estructuración y descripción sistemática de dicho campo conceptual parece una tarea harto ardua actualmente, debe reconocerse que sí existen algunas propuestas o sistematizaciones sobre dichos conceptos, ampliamente reconocidos por la comunidad investigadora.

La distinción teórica de estos cuatro términos resulta relevante en un trabajo como el nuestro, puesto que contribuye a delimitar el objeto de estudio. Isenberg (1987: 101), por ejemplo, parte del concepto de **clase de texto**, que define como:

[...] designación, conscientemente vaga, para toda forma de texto, cuyas características pueden fijarse mediante la descripción de determinadas propiedades, no válidas para todos los textos, independientemente de si estas propiedades pueden concebirse teóricamente y de qué manera, dentro del marco de una tipología textual.

²⁴ En otras palabras, a través del análisis de las cualidades de un texto, se puede llegar a determinar que un texto se clasifica como tipo de texto A, B o C, dado que dichas cualidades se asemejan a otras cualidades de textos ya clasificados como A, B o C.

Es decir, a primera vista, la clase de texto sería una categoría superior, no delimitada de forma clara teóricamente. Por el contrario, el **tipo de texto** constituye una categoría más precisa, una «designación teórica de una forma específica de texto, descrita y definida dentro del marco de una tipología textual» (Isenberg, 1987: 101) y la tipología textual se entiende como «una serie de regularidades dentro de un grupo de textos muy similares» (Corbacho, 2006: 83).

En relación con la cuestión del **género textual**, Calsamiglia y Tusón (2007: 242) afirman que su uso se orienta a la clasificación de gran parte de «los productos culturales en los que el uso de la palabra es un elemento fundamental». A esto se suman los condicionantes que esboza Bajtín (1999 [1979]: 252): la estructura interna, el registro y la estabilidad de estos elementos. En definitiva, el género discursivo constituye una categoría más amplia y abarcadora.

Por último, el concepto de **secuencia textual**, que surge de las investigaciones de Adam (1990, 2001 [1992]), defiende que el texto suele presentarse como un conjunto de secuencias heterogéneas que representan la unidad constituyente del texto, por lo que los textos no se presentan como un objeto homogéneo. En otras palabras: los textos pueden presentar diferentes tipos textuales en un mismo objeto textual. De esta forma, las secuencias son partes de un texto que cuentan con autonomía propia.

3.4. Las tipologías textuales

Desde nuestra perspectiva, la propuesta de tipos textuales parte de la estructuración de tipologías textuales. Para Isenberg (1987: 102), la tipología textual constituye un complejo de enunciados sobre textos que contiene los siguientes elementos:

- a) Una determinación general de su *campo de aplicación* [qué textos hay que estudiar].
- b) Una *base de tipologización*, esto es, un criterio (a ser posible complejo), de acuerdo con el cual puedan diferenciarse los tipos de texto que han de ser determinados.
- c) Un conjunto manejable y limitado de *tipos de texto*, definidos sobre la base de la tipologización.
- d) Para cada tipo de texto definido *t*, una *especificación* exacta de *t*, esto es, una descripción de todas las regularidades textuales específicas de *t*, no contenidas en la definición de *t* y que, por lo general, no son válidas para ninguno de los demás tipos de texto.
- e) Un conjunto de *principios de aplicación* que evidencian de qué manera pueden relacionarse los tipos de texto con los textos que pueden observarse de manera

aislada, o de qué manera los textos aislados han de ordenarse dentro de los tipos de texto definidos en la tipología (en Bernárdez, 1987: 102).

3.4.1. Clasificación de las tipologías textuales (I): problemas teóricos

Asimismo, existen diferentes formas de clasificación de las tipologías textuales. A este respecto, Isenberg (1987: 122) atiende a la propuesta de Beaugrande (1980: 197), quien se basa, en parte, en Schmidt (1978: 55):

- Se parte de las clases de textos tradicionales y se intenta determinar las peculiaridades características de cada clase de texto.
- Primero se desarrolla una teoría del texto y, luego, se comprueba si de ella resulta una tipología textual adecuada.
- Al elaborar una teoría del texto se intenta aplicarla a una tipología textual, de manera que se puedan definir las clases de texto tradicionales.
- Se desarrolla una tipología textual dentro del marco de una teoría textual e independientemente de las clases de texto tradicionales.

Como se observa, las posibilidades para clasificar los tipos de textos son diversas y parten de diferentes puntos de inicio. Sin embargo, es perceptible una falta de coherencia en las sistematizaciones de las tipologías, que conviven con diferentes puntos de vista (lógico, por otra parte, dada la multidimensionalidad textual), lo que suele derivar en la negación de la validez de las tipologías textuales, puesto que no son explicativas de ciertos objetos. Por ejemplo, atender exclusivamente a la función, al número de hablantes o a la temática, y no contar con criterios adicionales, puede derivar en que una tipología textual se encuentre con objetos textuales *inclasificables*, dadas sus particularidades. Como señala Loureda (2002: 161): «la tarea sistematizadora halla más dificultades en la propia naturaleza de los discursos: su casi infinita variedad, su mutabilidad y la inexistencia de formas materiales específicas para cada clase son trabas para alcanzar una tipología textual definitiva».

De hecho, ciñéndonos a las propiedades de los textos antes mencionadas, su carácter abierto y dinámico impide que las tipologías sean categorías estancas, estáticas, puesto que están sujetas a los cambios que se producen en los objetos textuales. De esta forma, cabe preguntar si la propuesta de tipologías textuales definitivas tiene en cuenta dicha dinamicidad. A nuestro juicio, no resulta práctico buscar una solución teórica definitiva a un problema cambiante, sino que, de igual forma en la que se trabaja en lingüística de

corpus, la tipología textual propuesta debe responder a unas necesidades de clasificación de acuerdo con los objetivos de los estudios (índice parcial según Loureda (2002: 162) en el que se tiene en cuenta un ámbito, como el periodístico, el literario, el administrativo...), y no pretender responder a todas las necesidades teóricas al mismo tiempo (índice total, ídem). Así, las tipologías textuales se diseñarían *a posteriori*.

3.4.1. Clasificación de las tipologías textuales (II): criterios de tipologización

Loureda (2002: 162) propone un método de elaboración de tipologías textuales basándose en el modelo de Isenberg (1987: 102) y considera el punto de inicio de la tipologización la selección de diversos individuos pertenecientes a una forma textual «intuitivamente reconocidos como tales»²⁵. El problema de dicho punto de partida se halla en el criterio de carácter subjetivo²⁶, así como en la posibilidad de que ciertos individuos no se ajusten a los tipos seleccionados para la tipología (Loureda, 2002: 163) o que un individuo clasificado dentro de un tipo textual cuente con secuencias textuales (por ejemplo, un tipo de texto *narración* puede contener una secuencia *descripción* o *ensayo*).

Dado que el objetivo de trabajar con tipos textuales es dar cuenta de una base de tipologización (que permite sistematizarlos naturalmente; Loureda, 2002: 164), cabe recurrir a la selección de rasgos o atributos de los textos, que permitan concluir sobre analogías y similitudes entre individuos. Para ello, es recomendable (con el fin de favorecer la homogeneidad de la tipología) partir de un número limitado de rasgos o jerarquizar los mismos (ibíd.: 162), puesto que «las circunstancias engloban una amplia serie de hechos que pueden ser relevantes en el hablar» (ibíd.: 166). Estos hechos son los que constituyen los criterios (o dimensiones) a los que pueden reducirse los rasgos de caracterización, de forma que existen dimensiones que son más relevantes que otras en la tipología textual, dependiendo del tipo de texto. Así, puede hablarse de rasgos esenciales (definen el tipo de texto) y rasgos complementarios (*concomitantes* según Loureda, 2002: 166; que no son definitorios del texto, pero acompañan a los esenciales). Por ejemplo, en el tipo de texto

²⁵ Esta intuición también se identifica en el trabajo de Bernárdez (1982: 212) cuando afirma que «no existe aún una teoría general de la tipología de los textos, y mucho menos una tipología o modelo tipológico ya elaborado. Se trabaja todavía, fundamentalmente, en un plano intuitivo».

²⁶ Al hablar del valor especializado de las unidades léxicas, también recurrimos a un punto de partida subjetivo.

artículo científico o *sentencia* es esencial la forma, la estructura, y es complementario el hecho de que incorporen imágenes.

Loureda (2002: 172) se sirve de la siguiente tabla para caracterizar las relaciones opositivas entre los tipos de *conversación*, usando la cursiva para los rasgos esenciales propios del tipo textual y la redonda para los rasgos esenciales comunes a la tipología. Para estos rasgos comunes ha tenido en cuenta, sobre todo, la finalidad, la relación hablante-oyente y la forma del discurso:

dimensiones tipos de texto	hablante- oyente	discurso		circunstancias	medio de comunicación	finalidad
		forma	tema			
conversación	plural	alternancia				interactiva
aparte	plural	alternancia		<i>al margen de otros</i>		interactiva
charla	plural	alternancia <i>no formal</i>				interactiva
chat	plural	alternancia <i>por escrito</i>			<i>la red internet</i>	interactiva
coloquio	plural	alternancia <i>organizado</i>		<i>tras una exposición en público</i>		interactiva
debate	plural <i>con puntos de vista diferentes sobre un tema</i>	alternancia <i>organizado</i>	<i>predeterminado</i>			interactiva
discusión	plural <i>con puntos de vista diferentes sobre un tema</i>	alternancia	<i>polémico</i>			interactiva
entrevista	plural <i>entrevistado y entrevistador</i>	alternancia <i>organizado</i>	<i>una persona (o más) y su modo de verse o de ver el mundo</i>			interactiva
murmuración	plural	alternancia	<i>la crítica de un tercero</i>			interactiva
parrafeo	plural	alternancia <i>no breve no formal</i>	<i>no determinado</i>			interactiva
regateo	<i>comprador y vendedor</i>	alternancia		<i>una transacción</i>		interactiva <i>negociadora</i>
tertulia	plural	alternancia		<i>en una reunión</i>		interactiva

Tabla 2: Relaciones opositivas entre los tipos de conversación (Loureda, 2002: 172)

Cabe resaltar que, desde nuestro punto de vista, resulta práctico jerarquizar los rasgos en esenciales y complementarios y, mucho más práctico, si hacemos referencia a *rasgos* o *criterios universales*, lo que resulta más explicativo en la descripción discursiva. Esto es lo que defienden Vilarnovo y Sánchez (1994: 34), que consideran, siguiendo a Coseriu (1998), que «los discursos se definen sobre todo por su finalidad y, en virtud del fin que se quiere conseguir, se seleccionan y articulan todos los elementos textuales». Conviene defender dicha postura, puesto que todos los textos cumplen con una función, en tanto que se articulan como objetos comunicativos. Este mismo criterio es el que defiende Cabré (2002a: 92): «los textos suelen (aunque no obligatoriamente) asociarse prioritariamente a un tipo textual de acuerdo con su función más prominente». Así, el

cumplimiento de una función por parte de los textos garantiza el éxito comunicativo; transcribiendo a Vilarnovo y Sánchez (1994: 43):

los textos tienen unas características pragmáticas [...] para cumplir con una determinada función, para realizar un tipo de interacción social, y con relación a esta finalidad [...], se constituyen de una manera determinada los diversos textos. Al mismo tiempo, si los textos consiguen alcanzar su finalidad [...], es precisamente porque han sido dotados de ciertas características, porque están constituidos de tal modo que guardan un tipo concreto de relaciones con respecto al emisor, receptor y demás elementos de comunicación (primera dimensión pragmática).

Así, Loureda, en otra publicación, distingue tres tipos de clasificaciones de las tipologías (2003: 57):

- a) Las que parten de características internas al texto, como la estructura informativa, el tema, los elementos de (cor)referencia o las estructuras sintácticas, entre otros. En esta clasificación se incluyen las de Weinrich, Biber y Posner.
- b) Las que parten de características externas al texto, como la intención del hablante, la relación entre los interlocutores, la función global del acto de habla, etc. En esta clasificación se incluye la de Glinz.
- c) Las que combinan las dos tipologías anteriores, como la clasificación de carácter binario de Sandig, la cognitiva de Werlich o las funcionales de Grosse o Adams. Este tipo de clasificación mixta sería la empleada en este trabajo.

4. LA TERMINOLOGÍA COMO CAMPO CIENTÍFICO

El presente trabajo se enmarca en el campo científico de la terminología. Guerrero Ramos (1999: 880) la define como «una materia lingüística de carácter interdisciplinario, cuyo objeto es el estudio y definición de los términos pertenecientes a las lenguas de especialización». Desde nuestro punto de vista, debe añadirse a esta definición el estudio de los textos en los que se inscriben estos términos. De igual forma, en contraste con disciplinas muy relacionadas (o de las que se sirve la terminología), Cabré (1993: 82 y ss.) apunta ciertas relaciones y diferencias con respecto a la caracterización que realiza Wüster que tienen que ver con el campo de trabajo, la unidad de base, los objetivos aplicados o el método de trabajo.

Cabe mencionar que cuando hablamos de *terminología* podemos hacer referencia a diferentes perspectivas del término (Monterde Rey, 2002a; Nkwenti-Azeh, 1998: 157, en Montero y Faber, 2008: 19):

- a) La primera considera la terminología como la **teoría** que explica las relaciones entre los conceptos y los términos. En el siglo XX existen tres posturas principales: i) la que considera la terminología como una disciplina autónoma y autosuficiente (es la postura de la Teoría General de la Terminología); ii) la que no percibe la terminología como una disciplina; y iii) la que considera la terminología un campo científico autónomo pero interdisciplinar que estudia un objeto multidimensional y transdisciplinar (es la postura de la Teoría Comunicativa de la Terminología y la que defenderemos en nuestro trabajo).
- b) La segunda concepción percibe la terminología como el **trabajo** de compilar, describir y presentar los términos, conocida también como *terminografía* o *gestión terminológica*.
- c) La tercera concepción se refiere al **vocabulario** empleado en los textos especializados, es decir, a los términos.

En otro orden de cosas, resaltaremos que la concepción de la terminología, en su acepción de vocabulario especializado, cumple dos funciones principales (Cabré, 1993: 1999): la función de representación (el término codifica el conocimiento especializado de forma sistemática) y la función de comunicación (el término transfiere el conocimiento en los actos de comunicación), que se corresponden con la función ideativa e interpersonal propuestas por Halliday (1982 [1979]) (*vid. infra*).

En último lugar, mencionaremos algunas de las razones del auge del estudio de la terminología. Entre ellas se encuentran las que recoge Rondeau (1984: 1-5, en Monterde Rey, 2002b: 214 y 215). Desde nuestro punto de vista, las siguientes son las que consideramos de mayor importancia en su clasificación y las que enmarcan nuestro trabajo:

- El avance de la ciencia y la técnica.
- El espectacular desarrollo de la tecnología.
- El desarrollo de los medios de comunicación.
- El desarrollo de las relaciones políticas y comerciales internacionales.
- La intervención de los gobiernos en materia lingüística.

4.1. Teorías de corte normativo: la Teoría General de la Terminología (TGT)

Los enfoques teóricos de la terminología se inician con las reflexiones sobre metodología aplicada al estudio y organización del vocabulario técnico de la electrotecnia realizadas por Wüster en 1931. No obstante, no es hasta 1979 (1998 en español²⁷) cuando aparece la obra teórica de conjunto de la Escuela de Viena denominada *Teoría General de la Terminología* (TGT), a la que se unieron otras escuelas.

Esta teoría, que se orienta a la normalización tanto lingüística como conceptual, tiene como objetivo mejorar la comunicación especializada planteada como el ideal de univocidad entre término y concepto, y alcanzar la monorreferencialidad de la relación signo-objeto. Para ello, Wüster parte de la idea de que los conceptos existen independientes del sistema lingüístico (perspectiva de autonomía del concepto) y de que se trata de ideas atemporales, ahistóricas, asociales e independientes de cualquier ideología (Montero y Faber, 2008: 32). Aunque el objeto de estudio de la TGT es el término, el punto de partida de esta teoría es el concepto independiente de la denominación (perspectiva onomasiológica). La concepción de los términos de Wüster se orienta a las situaciones de comunicación profesional estandarizada que tengan fines políticos, comerciales o económicos, nacionales o internacionales y descontextualizados cultural, comunicativa y discursivamente (se observa, pues, que idealiza y restringe la comunicación). Asimismo, la TGT defiende un sistema conceptual previo e inamovible y establece una clara diferencia entre la lengua natural y el código de especialización, sosteniendo que en la primera se observa la polisemia y la sinonimia, mientras que en la segunda tan solo existe monorreferencialidad y homonimia entre disciplinas. Según Wüster, los términos siempre son independientes del contexto, por lo que se niega la variación terminológica.

Con todo, y como afirma Cabré (1999: 105), Wüster concibe la terminología como una herramienta idealizada que silencia la variación debida a la dimensión social y comunicativa, la dinamicidad de la realidad y la relación interdisciplinar del conocimiento. Wüster construye una teoría idealista y objetivista orientada a la planificación y a la normalización lingüística, que fue apoyada por organismos internacionales como la ISO en 1947 y años posteriores. Consideraba la terminología como una disciplina autónoma

²⁷ A lo largo de esta tesis, para hacer referencia a los fundamentos de la TGT usaremos las fuentes de Wüster de 1985 y de 1998, editado por Cabré y traducido por Nokerman.

centrada en el estudio de los sistemas de denominación a partir de sistemas de conceptos universales.

En conclusión, la TGT se resume en los siguientes apartados:

-
- a) La perspectiva onomasiológica: prima el concepto.
 - b) El objeto de estudio es el término, cuyo significado es el concepto.
 - c) La autonomía del concepto.
 - d) Teoría prescriptiva.
 - e) La orientación normalizadora.
 - f) Silencia la variación.
 - g) La comunicación exclusiva entre especialistas.
 - h) La división entre lengua natural y lengua especializada.
-

Tabla 3: Características principales de la TGT

A partir de los fundamentos de esta teoría, Adelstein (2004: 74), basándose en el trabajo de Cabré (1999: 96), sintetiza las razones que justifican la falta de adecuación explicativa y descriptiva de los términos reales de la TGT:

La teoría de Wüster [...] no da cuenta de:

- a) la multidisciplinariedad de abordaje de las unidades terminológicas (denominativa, cognitiva y funcional)
- b) la poliedricidad de todas y cada una de estas unidades
- c) su doble función en la realidad del discurso especializado (representativa y comunicativa)
- d) la distinción entre su valor descriptivo y su valor prescriptivo, discriminado por las situaciones de comunicación
- e) la variación conceptual inherente a toda unidad de conocimiento, indisolublemente ligada a una cultura específica (general y científica) que determina una visión del mundo
- f) la dependencia lingüística de las unidades terminológicas realizadas en las lenguas particulares
- g) la variación denominativa inherente al discurso y a la comunicación, tanto general como especializada, en función de las características pragmáticas del discurso.

La ausencia de respuestas a dichas causas origina diferentes enfoques de estudio de la terminología, que se posicionan a favor de una perspectiva descriptiva. Por una parte, analizaremos las teorías de orientación comunicativa, entre las que se incluyen la socioterminología de Gaudin y la teoría desarrollada por Cabré. Por otra, analizaremos las teorías de orientación cognitiva, entre las que encontramos las propuestas de Temmerman y Faber y Jiménez.

4.2. Teorías de corte comunicativo

4.2.1. La socioterminología

A finales del siglo XX, comienzan a surgir críticas a la teoría desarrollada por Wüster y la Escuela de Viena que reclaman la incorporación de todos los factores que influyen en la comunicación, ya que la terminología forma parte de ella. En primer lugar, las críticas demandan sustituir la perspectiva prescriptiva por una descriptiva que analice los contextos y situaciones reales de aparición de los términos (Montero y Faber, 2008: 34). Una de las primeras reacciones teóricas que cuestionan la propuesta de la TGT surge con Gaudin (1993) y la socioterminología. Una de las aportaciones más innovadoras de esta teoría para nuestro estudio es la existencia de límites difusos entre dominios de especialización y, por tanto, entre los especialistas, el vocabulario empleado y toda actividad discursiva en general.

Este enfoque rechaza el estudio sincrónico de los códigos de especialización (Gaudin, 1995) y de la monosemia como fenómeno ideal y se posiciona a favor del análisis de la sinonimia y la polisemia, así como de la concepción de la realidad y la ciencia en campos sin límites discretos (Gaudin, 1993). No obstante, la postura de la socioterminología ha prestado más atención a subrayar los puntos débiles de la teoría de Wüster más que a defender una teoría propia de la terminología que sirviese para compilar diccionarios especializados (Fuertes-Olivera y Tarp, 2014: 112).

4.2.2. La Teoría Comunicativa de la Terminología (TCT)

La llegada de los estudios empíricos sobre material textual real reveló fenómenos contrarios a lo promulgado por la TGT: el ideal de la biunivocidad es inexistente en la comunicación real; los conceptos no son universales y no existen fronteras bien definidas entre disciplinas (Cabré, Feliu y Tebé, 2001: 303). La teoría de Wüster, por tanto, encontraba numerosas limitaciones en el plano discursivo, ya que negaba la evidente variación discursiva al rechazar el valor pragmático del término (en parte por considerar que el término constituía un tipo de unidad diferente a las unidades léxicas del discurso común (Cabré y Estopà, 2002: 142), basándose en el argumento de la monosemia y biunivocidad del término).

El afán de idealismo de la TGT y su interés por los aspectos prescriptivos de la terminología llevaron a Cabré a desarrollar la Teoría Comunicativa de la Terminología

(TCT) (1993, 1998 y 1999, entre otros), que parte de la idea de que los términos no son unidades aisladas con un sistema propio, «sino unidades que se incorporan en el léxico de un hablante en cuanto adquiere el rol de especialista por el aprendizaje de conocimientos especializados» (Cabré, 1999: 131). Además, defiende la asociación de la variación a las dimensiones de estudio del término, lo que implica la existencia de «niveles de especialización, grados de opacidad semántica, índices de comprensión o densidad cognitiva, etc.» (Cabré, 1999: 140).

Para hacer frente a esta complejidad discursiva, Cabré (1999: 149) propone una macroteoría lingüística constituida por tres teorías: una teoría de la gramática, una teoría de la adquisición y una teoría de la actuación. Esta última es la que más nos interesa en nuestro trabajo, dado que recoge «en un solo modelo el uso general y el especializado en toda su variada amplitud (temática, perspectiva, nivel de especialización, propósito comunicativo, propósito funcional, tipo de texto, tipo de discurso, etc.)» (ídem).

A esto se suma la consideración del estudio de la terminología *in vivo*, siendo el comportamiento de esta fruto de su situación y condiciones discursivas, que implican:

[...] el emisor y el receptor (tipo de emisor y receptor, intenciones y nivel de conocimiento que ambos poseen sobre el tema), la situación (el medio en que se produce la comunicación y el sistema de transferencia utilizado), el propósito y las expectativas de emisor y receptor con relación a su interacción. Cada uno de estos factores posee un determinado valor en cada acto comunicativo y su conjunto explica la configuración de un tipo de texto que pretende ser adecuado a las circunstancias en las que se produce (Cabré, 2002a: 89).

Todos estos factores son identificables de forma explícita en cada texto, de manera que «el texto contiene siempre las huellas de su proceso de construcción» (Cabré, 2002a: 89). Esta postura también se refleja en Cabré *et al.* (2001: 184), basándose en el estudio de Domènech (1998):

Desde una perspectiva cognitiva [...] la condición y el nivel de especialización de un texto, determinados a partir de criterios externos o pragmáticos, pueden validarse sobre la base de características internas al texto, relacionadas con las unidades portadores de conocimiento especializado (UCE) que el texto contiene.

La TCT integra la semántica y la pragmática en su teoría, va más allá de los aspectos formales y supera la separación que fijaba la teoría clásica entre la lingüística y la terminología (Cabré, Feliu y Tebé, 2001: 302). El objetivo de la terminología para la TCT como campo de conocimiento se centraría en la descripción, dejando a un lado la tendencia prescriptivista de la TGT:

El objetivo de la terminología teórica es el de describir formal, semántica y funcionalmente las unidades que pueden adquirir valor terminológico, dar cuenta de cómo lo activan y explicar sus relaciones con otros tipos de signos del mismo o distinto sistema, para hacer progresar el conocimiento sobre la comunicación especializada y las unidades que se usan en ella. El objetivo de la terminología aplicada es el de recopilar las unidades de valor terminológico en un tema y situación determinados y establecer sus características de acuerdo con esta situación. Dentro de sus características puede figurar su condición de unidad normalizada (Cabré, 1999: 132).

Por último, cabe mencionar la renovación terminológica que relaciona los usos especializados y no especializados del léxico en el discurso. Alejándose de la concepción estática de *lengua de especialización* (o *lenguaje de especialización* según otras fuentes), que dominaba la bibliografía sobre terminología desde antes de los años 90, Cabré *et al.* (2007: 852) proponen la denominación *discurso especializado*, que se ubica «en una perspectiva más procesual y holística, cognitiva y social», lo que coincide con la propuesta de Cabré (2002a: 89), que expone: «los términos constituyen en nuestra propuesta unidades insertas en el discurso, entendido el discurso en un sentido amplio: el texto y el contexto, o dicho de otro modo, el texto y sus condicionantes de producción, transmisión y recepción».

De forma complementaria, por tanto, podría emplearse el término *discurso común*, *discurso general* o simplemente *discurso no especializado*, para hacer referencia al polo opuesto. No obstante, de nuevo, existiría el problema de identificar y denominar el *continuum* intermedio: el discurso semiespecializado (*vid. infra* apartado II. 9.4., p. 144).

4.3. Teorías de corte cognitivo

4.3.1. La Teoría Sociocognitiva de la Terminología (TSCT)

Temmerman se basa en la semántica cognitiva, concretamente en el Modelo Cognitivo Idealizado (Lakoff, 1987), para construir la Teoría Sociocognitiva de la Terminología (TSCT) (2000). En ella, acaba con la hipótesis objetivista de la TGT y propone nuevos presupuestos teóricos que lleven a una nueva metodología que se centre en la descripción real de los significados de los términos que aparecen en los textos.

Los fundamentos de la TSCT parten del supuesto de que los conceptos no existen de forma objetiva como proponía la TGT, sino que se trata de unidades de comprensión, muchas de las cuales poseen una estructura prototípica, por lo que pueden clasificarse como categorías en constante cambio, que se organizan en un núcleo prototípico y en límites difusos que recogen información intercategorial e intracategorial y que se enmarca en

modelos cognitivos dependientes de la lengua (Temmerman, 2000: 73 y ss.; Montero y Faber, 2008: 44). A partir de estos supuestos, Temmerman sostiene que una metodología alternativa para la descripción terminológica debe basarse en «el estudio de las ocurrencias de los términos en un discurso especializado concreto y nunca como entidades aisladas e independientes» (Temmerman, 2002: 78), así como en función del contenido de los dominios de especialización y del perfil del usuario potencial del trabajo terminográfico, lo que constituye la base fundamental de nuestro estudio. Asimismo, sostiene que la aplicación de los modelos cognitivos, sobre todo la metáfora cognitiva, implica la perspectiva de la motivación del término (ídem).

El principio básico de la TSCT aplicado a la traducción y la interpretación defiende que:

Los términos en el discurso especializado forman parte de modelos cognitivos concretos que pueden variar a lo largo del tiempo y de la situación comunicativa concreta, lo que obliga al traductor y al intérprete a analizar los modelos cognitivos subyacentes [es decir, la disciplina en la que se emplea la unidad], tanto en el texto/discurso origen como en el texto/discurso meta, para detectar posibles problemas terminológicos y proponer soluciones factibles con el contexto de recepción (Montero y Faber, 2008: 47)²⁸.

En resumen, la TSCT se articula según los siguientes rasgos:

-
- a) La concepción cognitiva del término.
 - b) Los aspectos pragmáticos y sociolingüísticos fundamentales en el trabajo terminográfico.
 - c) El estudio descriptivo terminográfico basado en ocurrencias.
 - d) El Modelo Cognitivo Idealizado y las categorías.
 - e) La evolución constante del término.
 - f) La motivación del término.
-

Tabla 4: Características principales de la TSCT

4.3.2. La Terminología basada en Marcos (TbM)

La teoría propuesta por Faber y Jiménez (2002) se basa en el cognitivismo y toma como fuente el modelo lexemático funcional (Martín Mingorance, 1998) y la semántica de marcos (Fillmore, 1985; Fillmore y Atkins, 1994). En su teoría, Faber y Jiménez afirman que:

²⁸ Los corchetes son del autor de la tesis.

[...] el lexicón mental de un individuo no se encuentra dividido en unidades especializadas y no especializadas²⁹ sino que se trata como un *continuum* basado en una compleja red que refleja las creencias y perspectivas del sujeto acerca del mundo que le rodea (Montero y Faber, 2008: 49).

La premisa de la Teoría basada en Marcos (TbM) se corresponde con la idea de que «las expresiones lingüísticas codifican la perspectiva desde la que se percibe una situación y, además, reflejan los mecanismos cognitivos del hablante» (idea que defiende Lee; 2001) (ídem).

Los objetivos principales de la TbM se corresponden con el estudio y representación de la organización conceptual de los dominios de especialización, por una parte, y de la naturaleza multidimensional de los términos por otra, así como con la extracción y análisis de información semántica y sintáctica a través de corpus multilingües (Faber, 2009, cit. en Montero y Faber, 2008: 50).

Para lograr estos objetivos, la TbM parte de un análisis semasiológico y utiliza los patrones de la estructura léxica como puerta de entrada a la mente (ídem) para describir conceptual, léxica y pragmáticamente los distintos dominios y subdominios de especialización y construir un tesoro visual tridimensional (ibíd.: 53).

En resumen, la TbM se estructura en torno a las siguientes premisas:

-
- a) La semántica de marcos.
 - b) Las expresiones lingüísticas reflejan los mecanismos cognitivos del hablante.
 - c) La descripción de las unidades definidas en el texto.
-

Tabla 5: Características principales de la TbM

5. FUNDAMENTOS PARA ANALIZAR EL TÉRMINO COMO OBJETO DE ESTUDIO

5.1. Punto de partida

Fruto de la divulgación científica y de la democratización del conocimiento, el trasvase del uso de la terminología (en su sentido léxico) en y entre los campos de conocimiento y su relación entre el discurso común (DC) y el discurso especializado (DE) ha provocado, principalmente, la aparición de un nuevo objeto textual que denominamos *texto semiespecializado*. Para la dimensión de la lingüística aplicada, trabajar con este objeto textual es complicado, en tanto que necesita hacer uso de los criterios teórico-metodológicos de la lingüística teórica para determinar los usos especializados o

²⁹ Coincide con el modelo teórico de Lerat (1997) de la lexicografía terminológica.

terminológicos que tienen ciertas unidades léxicas y, de esta forma, proponer soluciones a los problemas comunicativos.

De acuerdo con diferentes teorías y autores (*vid. infra*), el elemento que caracteriza al texto especializado es el léxico especializado (la terminología). Entendemos, por tanto, que es un elemento determinante en la distinción del grado de especialización textual (junto con otros atributos como la relación emisor- destinatario o el tipo textual, entre otros). Para hallar criterios teórico-metodológicos que nos ayuden a determinar el valor especializado del léxico, es esencial describir el comportamiento del mismo *in vivo*³⁰, a través de los textos de la comunicación real (Cabré y Estopà, 2002: 147), es decir, siguiendo una terminología comunicacional (Cabré, Feliu y Tebé, 2001: 307).

La dificultad de determinar el valor especializado se halla en el interés económico y social de la ciencia, que provoca la divulgación de la misma y la difusión del conocimiento. Consideramos que el carácter terminológico del léxico en ciertos tipos de texto (semiespecializado y divulgativo; aunque también los especializados cuya terminología se ha banalizado a causa de su constante divulgación) depende de la percepción del usuario, puesto que es posible que las diferencias no se hallen en la denominación del término (la parte formal, el significante).

Cuando disponemos de los mencionados criterios (que van más allá de la propia percepción del usuario o que convierten esta percepción en un criterio de partida o un criterio sujeto a filtrado), podemos usarlos como herramientas para identificar los términos de un texto y establecer su densidad terminológica³¹ (atributo que consideramos relevante en la distinción del grado de especialización textual) e incluso su densidad cognitiva.

Este estudio tiene por principal objetivo contribuir a nuestra tesis doctoral con criterios fiables, fundamentados sobre una base teórica consistente y coherente. Para poder determinar los grados de especialización, debemos contribuir a delimitar los grados del valor de especialización (basándonos principalmente en el léxico). Seguimos a Cabré (2002b: 34), cuyas palabras ya usamos en la sección de hipótesis (*vid. supra* apartado I. 3., p. 18):

1) ¿Existen indicios explícitos que caracterizan un texto como especializado?

³⁰ La postura contrapuesta pretende estudiar los términos *in vitro*, «es decir fuera de su contexto de uso habitual, aisladamente a través de las unidades recogidas en los diccionarios léxicos, bases de datos o tesauros» (Cabré y Estopà, 2002: 147). También recibe el nombre de *terminología representacional* (Cabré, Feliu y Tebé, 2001: 307).

³¹ La densidad terminológica es la relación entre el número de unidades léxicas especializadas por unidad léxica de un texto (Domènech, 2006: 207).

- 2) ¿Qué propiedades tienen estos indicios? ¿son de carácter discreto o representan un *continuum* sin divisiones tajantes entre los valores que adquieren?
- 3) ¿Cómo el léxico contribuye a la tipologización de los textos especializados?

El punto 3) representa una de las preguntas a las que pretendemos responder. Si bien existen unidades léxicas cuyo valor especializado es distinguible a simple vista (es decir, a través de su denominación; ejemplo (4)), existen otras que hacen necesario ir más allá del significante de la unidad léxica y acudir a factores situacionales (pragmáticos y semántico-cognitivos; ejemplo (5)).

- (4) La implementación de un **rotor de ondas** a un **turborreactor**, conllevaría [*sic*] a estudios más avanzados de **fluido inestable**, materiales de fabricación, construcción y análisis estructural, dado a las temperaturas y presiones generadas por la **onda de choque de compresión y expansión** causada por los gases calientes que entran al **rotor de ondas** (Escobar *et al.*, 2011: 65)³².
- (5) Concurren, en verdad, como asesores a fin de evitar avenencias o transacciones precipitadas o suicidas para los intereses de la parte, (v. el art. || de la ley cit.) Sobre el acto de conciliación, tanto el demandante como el demandado deben concurrir acompañados cada uno por un **hombre bueno**, que han de ser españoles y gozar plenamente de sus derechos civiles³³.

Ya en 1999, Estopà (1999: 28) exponía este problema:

El problema principal dels SEACAT actuals de base lingüística és que detecten les UT a partir de patrons morfològics i/o morfosintàctics que no són exclusius d'aquestes unitats. Per reconèixer i delimitar una unitat de significació especialitzada (USE) no es pot recórrer només als aspectes morfosintàctics, sinó que cal tenir en compte també els elements lèxics, morfològics, sintàctics, semàntics i pragmàtics.

La autora defiende la existencia de recursos morfológicos, sintácticos, semánticos y pragmáticos para reconocer las USE (unidades de significación especializada); es decir, defiende la idea que más tarde expondría Cabré (2002; *vid. supra*). A su juicio la información semántica es la única manera de reconocer y delimitar de forma exhaustiva las unidades terminológicas³⁴ de un texto especializado (Estopà, 1999: 42-43). Este criterio aparece en el extractor automático YATE (Yet Another Term Extractor) elaborado por Vivaldi (2003: 1).

³² Fuente: ESCOBAR GARZÓN, Arnold; CERPA BERNAL, Rafael Mauricio; PIECHNA, Janusz y RICO RODRÍGUEZ, María Camila (2011): «Cálculos términos y de dinámica de gases de un turbojet “power generation x-01”», *Ingenium* 12(24), p. 56-65.

³³ Fuente: [s. a.] (6 mayo 2015): «Hombre bueno», *UniversoJus.com*. Disponible en: <http://universojus.com/definicion/hombre-bueno>.

³⁴ Para Estopà (1999), las unidades terminológicas (UT) son las unidades de significación especializada nominales.

La conjunción de estos recursos permitiría evitar tanto el silencio³⁵ como el ruido³⁶ en la extracción (Estopà, 1999: 160), que son fruto de los problemas de la lingüística formal:

Así, la llamada lingüística formal, fundamentada en la lógica y aplicada a la inteligencia artificial de manera exclusiva hasta hace poco, solo ha podido dar cuenta de los fenómenos lingüísticos susceptibles de ser explicados por formalismos. En consecuencia, los fenómenos lingüísticos no estrictamente formales o formalizables han sido relegados del punto central de su estudio para poder explicar más profundamente los fenómenos regularizables (Cabré y Estopà, 2002: 142).

Partimos de la hipótesis de que el valor especializado es el criterio determinante para la discriminación de unidades léxicas especializadas y no especializadas más allá de criterios formalistas. Se trata de un criterio pragmático-cognitivo dependiente del usuario; por lo que puede defenderse que se trata de un factor subjetivo de partida, al cual es necesario aplicar diferentes filtros (Vega Moreno, 2015: 252-295).

5.2. La teoría de puertas, el principio de poliedricidad y el principio de variación

La terminología, como campo de estudio³⁷, es multidisciplinar (Cabré, 1999: 94) porque su objeto de estudio es analizable desde diferentes puntos de vista y desde diferentes disciplinas. El léxico terminológico constituye un objeto poliédrico (Cabré, 2002a: 93; 2002b: 16; 2008: 33). A este principio de poliedricidad se suma la teoría de las puertas (Cabré, 2000b: 14):

La conception de la terminologie comme domaine de connaissance nécessairement interdisciplinaire, qui traite des termes et qui intègre les aspects cognitifs, linguistiques, sémiotiques et communicatifs des unités terminologiques, nous conduit à proposer ce que nous appelons 'la théorie des portes'. Il s'agit d'une théorie qui rend possible le traitement multidimensionnel des termes.

³⁵ El silencio es el obstáculo que se produce cuando los resultados de una extracción (identificación) terminológica automática no incluyen algunas unidades léxicas especializadas (que poseen valor especializado) a causa de que estas unidades no cumplen con los parámetros de identificación (generalmente parámetros formales) del extractor automático.

³⁶ El ruido es el bstáculo que se produce cuando los resultados de una extracción (identificación) terminológica automática incluyen tanto las unidades léxicas especializadas como otras unidades léxicas no especializadas que también cumplen con los patrones de identificación formal del extractor automático pero que carecen de valor especializado.

³⁷ Cabré (2007) se inclina hacia el término *campo de conocimiento* y rechaza considerar la terminología como una disciplina, puesto que dicha denominación «remite a materias que se han institucionalizado socialmente» (sin página).

La TCT entra por la puerta de la lingüística para estudiar las unidades especializadas, lo que no impide que se sirva de la teoría cognitiva y de la teoría comunicativa para describir su objeto de estudio. De hecho, Cabré (2002a: 91) defiende que «los textos no se conciben como meras unidades de información, sino que constituyen unidades de funcionamiento cognitivo y social».

La teoría de las puertas permite distinguir las perspectivas del término, que «es una unidad al mismo tiempo semiótica, lingüística, cognitiva y social, que puede ser tratable desde y por diferentes disciplinas» (Cabré y Estopà, 2002: 144). Así, en relación con esta colaboración disciplinaria, necesaria, pero a veces insuficiente para explicar la complejidad de nuestro objeto de estudio, pueden distinguirse tres conceptos clave:

- a) El de **transversalidad**: se trata de la relación que existe entre el objeto de un campo de conocimiento y su presencia en otros campos o disciplinas. Dado que la terminología, entendida en su sentido de léxico de una disciplina, constituye el instrumento de representación y transferencia de conocimiento, es lógico pensar que todas las disciplinas partan de su propia terminología para hacer su comunicación efectiva.
- b) El de **multidisciplinaria**: siguiendo la teoría de las puertas, el término sería observable desde sus dimensiones semiótica, lingüística, cognitiva y social (*vid. supra*), y, por tanto, explorable desde diferentes disciplinas. Un ejemplo podría ser el objeto *enfermedad*, tratable desde la patología en el ámbito médico, desde sus efectos en la economía, desde la perspectiva sociológica o desde la propia terminología.
- c) El de **interdisciplinaria**: se refiere a las dimensiones de estudio que un campo adopta para estudiar un objeto. Es decir, toma elementos de diferentes disciplinas para crear un nuevo método de trabajo. En el caso de la terminología, adoptaría fundamentos de la teoría lingüística, de la teoría cognitiva y de la teoría comunicativa.

En resumen, la terminología se percibe como: (i) un campo científico (ii) interdisciplinaria (iii) que estudia un objeto multidisciplinaria y poliédrico, (iv) que es transversal al resto de disciplinas. Estas relaciones podrían explicar, en principio, la dificultad de estudiar el comportamiento terminológico.

Condicionado, por tanto, por dichas características y por la diversidad de situaciones sociales y cognitivas en las que puede tener lugar el proceso comunicativo, así como sus perspectivas de análisis, no es de extrañar que Cabré, ya en 1998, proponga el

principio de variación, determinante en nuestro trabajo y que afecta tanto a la configuración discursiva, por una parte, como a la configuración del léxico, por otra:

Todo proceso de comunicación comporta inherentemente variación, explicitada en formas alternativas de denominación del mismo concepto (sinonimia) o en una apertura significativa de una misma forma (polisemia). Este principio es universal para las unidades terminológicas, si bien admite diferentes grados según las condiciones de cada tipo de situación comunicativa (Cabré, 1998: 72).

[...]

El grado máximo de variación de la terminología lo cumplirán los términos de las áreas más banalizadas del saber y los que se utilizarían en el discurso de registro comunicativo de divulgación de la ciencia y de la técnica; el grado mínimo de la variación sería propio de la terminología normalizada por comisiones de expertos; el grado intermedio, la terminología usada en la comunicación natural entre especialistas (Cabré, 1998: 63).

En este trabajo, nos centraremos en la segunda concepción de la terminología, en el plano léxico, y puede observarse en casos como los que siguen en los pares (6), (7) y (8), por una parte, y (9) y (10), por otra:

- (6) Estas **cirugías** tan largas son frecuentes porque atendemos patologías oncológicas complejas como tumores en páncreas, hígado, esófago, etc. (Martín García, 2015)³⁸.
- (7) Es un criterio que he conservado en todas las ocasiones el que exista una probada justificación médica para proceder a la **operación**. En el caso de la joven italiana existían múltiples problemas de tipo psíquico y puedo asegurarle que me he negado a efectuar la operación por motivos de estética, como en el caso de una maniquí francesa que quería alargar unos centímetros sus pantorrillas (s. a., 1979)³⁹.
- (8) La preparación psicológica ante la **intervención quirúrgica** en una muestra de niños sevillanos (Cullen, 1992)⁴⁰.
- (9) Su característica principal es una lesión cutánea denominada eritema migratorio. Las fiebres recurrentes producidas por diversas especies de *Borrelia* se caracterizan desde el punto de vista clínico por ciclos de fiebre y **apirexia** (Escudero-Nieto y Guerrero-Espejo, 2005)⁴¹.

³⁸ Fuente: MARTÍN GARCÍA, Cristina (24 de febrero de 2015): «¿Cómo es la residencia de Cirugía General y del Aparato Digestivo?», *DMédico*. Disponible en: <http://www.diariomedico.com/medico-joven/estudiar/mir/noticias/especialización-cirugia-general-aparato-digestivo>.

³⁹ Fuente: (s. a.) (30 de enero de 1979): «Con una operación quirúrgica se puede aumentar de estatura», *El País*. Disponible en: http://elpais.com/diario/1979/01/30/sociedad/286498818_850215.html.

⁴⁰ Fuente: CULLEN, Kathleen (1992): *La preparación psicológica ante la intervención quirúrgica en una muestra de niños sevillanos*. Tesis doctoral. Sevilla: Universidad de Sevilla. Disponible en: <https://dialnet.unirioja.es/servlet/tesis?codigo=46223&orden=0&info=link>.

⁴¹ Fuente: ESCUDERO-NIETO ÁRBOL, Raquel y GUERRERO-ESPEJO ÁRBOL, Antonio (2005): «Enfermedades producidas por *Borrelia*», *Localización: Enfermedades infecciosas y microbiología clínica*, 23(4), pp. 232-240. Disponible en: <http://www.elsevier.es/es-revista-enfermedades-infecciosas-microbiologia-clinica-28-articulo-enfermedades-producidas-por-iborrelia-i-13073150>.

- (10) Del Pozo explica que la **ausencia de fiebre** y el empeoramiento solo puede explicarse por dos causas: «O los antitérmicos con los que le están tratando camuflan la temperatura real o bien es un síntoma de que su organismo empieza a claudicar» (N. R. C, 2014)⁴².

Como puede comprobarse, los ejemplos (6), (8) y (9) optan por los formantes grecolatinos para hacer referencia al mismo concepto que en (7) y (10), respectivamente, por lo que hablamos de sinonimia entre dichas combinaciones. El grado de familiaridad o frecuencia subjetiva (por ejemplo, al oponer *apirexia* y *ausencia de fiebre*) es otra cuestión que se debate en el apartado II. 7.4. (p. 109).

5.3. Fundamentos discursivistas para el análisis del léxico especializado

Según Cabré *et al.* (2007: 851), las teorías terminológicas han evolucionado desde posturas lexicalistas, cuya atención se centra en la pieza léxica (el término) y sus características, a posiciones discursivistas, que prestan atención al marco comunicativo especializado. Las dos posturas estudian la terminología desde la puerta de la lingüística, pero adoptando fundamentos diferentes o intereses distintos. Lo que sí parece claro es la relevancia del léxico especializado, o terminología, en el texto especializado. Parece ser que se trata de la característica más distintiva, como defienden los siguientes investigadores⁴³:

In special communication terms and standardised terms make a **critical contribution** to achieving complete and effective communication (Sager, 1990: 105).

el discurso especializado está **determinado de modo decisivo** por la dimensión léxica de los textos especializados, o sea, por su terminología (Arntz y Picht, 1995: 45).

las llamadas unidades terminológicas son una **pieza imprescindible** para dar cuenta de la competencia y actuación de los hablantes (Cabré, Feliu y Tebé, 2001: 302).

El léxico tiene como **función básica** la de representar el conocimiento que constituye el esqueleto cognitivo del texto (Cabré, 2002a: 94).

⁴² Fuente: N. R. C. (10 de octubre de 2014): «La ausencia de fiebre puede ser síntoma de que el organismo claudica», *ABC*. Disponible en: <http://www.abc.es/sociedad/20141010/abci-fiebre-sintoma-empeoramiento-201410092350.html>.

⁴³ La negrita es del autor de la tesis.

[s]i eliminamos la terminología de un texto especializado, solo encontramos conectores (preposiciones, conjunciones) y otras estructuras de la lengua común (verbos, sustantivos, etc.) que permiten anclar la terminología en el texto. **Sin su terminología**, un texto de especialización **es totalmente incoherente**, pues es ella la que porta toda la carga semántica (Monterde Rey, 2002b: 115).

En el nivel léxico, que es el nivel que, junto con el textual, **más diferencia el discurso especializado del no especializado** (Cabré y Estopà, 2005: 75).

[...] la terminologia. Efectivament, és evident que les unitats portadores de coneixement especialitzat per excel·lència són les unitats terminològiques i que **una de les característiques lingüístiques més destacables** dels textos especialitzats és la presència d'aquestes unitats, específiques d'un àmbit determinat del saber (Domènech, 2007: 250).

Some researches about the differences between specialized and non-specialized texts exist. The majority of these works consider the lexicon as **the most visible and differentiator aspect**, in order to distinguish between these two kinds of texts (Cabré *et al.*, 2014: 224).

Algunas de estas definiciones sirven de justificación de la importancia del léxico especializado en el discurso especializado: su relación con la efectividad de la comunicación (Sager, 1990: 105), con la descripción de la competencia y la actuación (Cabré, Feliu y Tebé, 2001: 302), y con el hecho de que los términos son los encargados de la representación del conocimiento (Cabré, 2002a: 94) y portan la información semántica (Monterde Rey, 2002b: 115).

Según Cabré *et al.* (2007: 852), el estudio del discurso especializado se ha centrado durante décadas en emplear el punto de vista lexicalista, es decir, priorizaban «la descripción de los términos como unidades de códigos especializados» (ídem).

La investigación lingüística actual se inclina más hacia una posición discursivista. En este trabajo nos centramos en la descripción del funcionamiento del léxico desde sus condicionantes comunicativos, es decir, desde una posición discursivista. Dado que vamos a centrarnos en la descripción del léxico, conviene detenerse en acotar el metalenguaje que emplearemos en relación con este nivel.

A pesar de que en la bibliografía suelen oponerse *palabra* y *término* como unidad léxica de la lengua común y unidad terminológica o especializada respectivamente, lo cierto es que existen diferencias notables. Desde nuestro punto de vista, las *palabras* constituyen la unidad hiperonímica bajo la cual se hallan las *unidades léxicas* y las *unidades no léxicas*

(todas aquellas con significado gramatical, pero sin significado léxico)⁴⁴. Así, las *palabras* pueden servir de sinónimo de *unidad lingüística*. Bajo el paraguas de la unidad léxica se encontrarían aquellas unidades no especializadas (que la bibliografía crítica suele denominar simplemente *palabras*) y las unidades léxicas especializadas (los *términos*). El siguiente esquema ordena las ideas que acabamos de exponer:

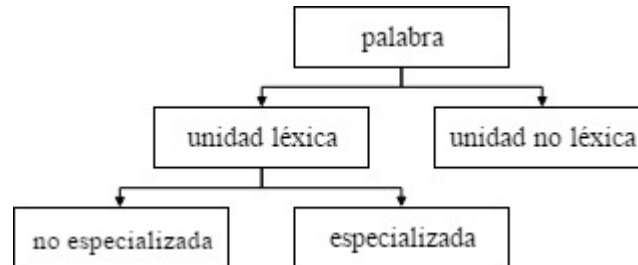


Ilustración 2: Esquema básico unidad léxica especializada-unidad léxica no especializada

Con el objetivo de aclarar las decisiones denominativo-conceptuales que afectan a nuestra tesis doctoral, emplearemos la oposición *unidad léxica especializada-unidad léxica no especializada* para referirnos a la relación entre unidad léxica sin valor especializado y la unidad léxica con valor especializado respectivamente, por lo que divergimos ligeramente de la distinción de Adelstein (2004: 55), que opta por la oposición *término-palabra*. En esta investigación, reservamos la denominación *palabra* para hacer referencia a toda aquella unidad lingüística que se encuentra entre dos espacios en blanco en el texto. El interés en realizar esta distinción denominativa reside en la relevancia con la que cuenta el término *palabra* desde el punto de vista estadístico en la lingüística aplicada (que constituye uno de los fundamentos de esta investigación), pues los gestores de corpus trabajan con *tokens*, es decir, cadena de caracteres entre espacios en blanco.

La unidad léxica no es *per se* palabra o término. Depende de su actualización en el discurso según las condiciones pragmáticas de forma que:

La información semántica estaría organizada en rasgos, componentes, módulos o paquetes de información, que se activarían de maneras diversas, según cada situación, dando lugar a los distintos sentidos posibles de la UL [unidad léxica] (Adelstein, 2004: 80 y 81).

Esta distinción basada en condiciones pragmáticas lleva a Cabré (1999, 2000b) a rechazar que los términos formen un lexicón distinto. El componente léxico estaría formado por las unidades léxicas en general (unidad y términos), de forma que integraría de la misma

⁴⁴ Respecto a dicha diferenciación, cabe recurrir a las posiciones de la lingüística generativa, que defienden una gradación entre las unidades léxicas y las gramaticales, en el sentido de que a veces no es posible distinguir de forma tajante qué unidades poseen significado propio y qué unidades activan significado.

forma el conocimiento lingüístico general y el conocimiento lingüístico especializado. En concreto, se trataría de un *continuum*, que oscilaría entre las unidades léxicas claramente especializadas y las claramente no especializadas, siendo las unidades léxicas del intervalo aquellas que se activan o no por cuestiones pragmáticas. A nuestro juicio, en este fenómeno se puede observar el carácter fractal de la comunicación especializada (Zamorano, 2012: 697; *vid. supra* apartado II. 2.1., p. 37), en el que el *continuum* se incluye dentro del *continuum*.

En nuestro trabajo nos centraremos en el nivel semántico, dejando a un lado el aspecto sintáctico, el morfológico y el fonológico. Es decir, pretendemos adoptar una postura similar a la de Adelstein (2007a: 277), en la que se relaciona de manera estrecha la semántica léxica y la semántica discursiva.

5.4. Perspectivas de estudio del término

De acuerdo con Cabré (1999: 19 y 20), es posible atender a una triple concepción sobre los términos:

- a) Según la **concepción lingüística**, los términos forman parte del sistema léxico y comparten las características de la lengua común. La terminología sería, pues, «el conjunto de signos lingüísticos que constituyen un subconjunto dentro del componente léxico de la gramática del hablante» (ibíd.: 19). La diferencia con respecto al léxico no especializado se hallaría en la competencia del hablante oyente ideal: competencia general o especializada/restringida.
- b) Según la **concepción filosófica**, la terminología constituye las unidades cognitivas de conocimiento y aproximación al mundo, de representación de este conocimiento (de carácter especializado), puesto que «dan una idea de la organización del mundo especializado» (idem).
- c) Según las **diferentes disciplinas**, la terminología se encarga de expresar, comunicar y transferir el conocimiento especializado.

De esta forma, en las tres perspectivas coinciden la orientación finalista, el valor funcional y la concepción de conjunto de unidades, pero divergen en la concepción de la propia unidad: para la lingüística, son unidades de significación; para la filosofía, unidades de cognición; y para las disciplinas, unidades de denominación.

A esta triple concepción se suma su definición por la unión de tres componentes (Cabré, 2009: sin página):

- a) El **componente cognitivo**: los términos representan la percepción y categorización de la realidad.
- b) El **componente lingüístico**: el término se describe a través de las mismas propiedades que describen el resto de unidades lingüísticas.
- c) El **componente social**: los términos permiten la comunicación de los expertos entre sí.

A la hora de determinar su composición, seguimos a los autores de orientación lexicológica⁴⁵ (en concreto a Rondeau, 1984: 21), que, siguiendo el modelo saussureano del signo lingüístico como unión de significante y significado, atribuyen al término la denominación en el plano de la expresión y el concepto (*la notion*) en el plano del contenido. Desde nuestra perspectiva, el funcionamiento lingüístico, cognitivo y social del término queda sujeto a su inserción en el discurso.

5.5. Modelos de relación entre discurso común y discurso especializado

Como ya se mencionó en apartados anteriores, y basándonos en Cabré *et al.* (2007: 852), adoptaremos la denominación de *discurso especializado*⁴⁶ como contraposición al *discurso común*. Arntz y Picht (1995: 28) defienden la existencia de diversas *lenguas especializadas* (asimiladas según nuestro punto de vista a *discursos especializados*), cuyos límites se difuminan con respecto al área nuclear, al discurso común. Los modelos o teorías que tienen como objetivo explicar estos límites o relaciones son, principalmente, el de Baldinger (1952), el de Reindhardt (1966), el de Heller (1970) y el de Hoffmann (1985).

⁴⁵ De acuerdo con Adelstein (2004: 54), existen tres posiciones respecto al estatuto semiótico del término, que quedan justificados por los modelos semióticos empleados para la representación de la unidad terminológica. Estos modelos serían: (i) el modelo de cuatro campos de Wüster; (ii) los modelos triádicos de los seguidores de la TGT y (iii) el modelo de signo lingüístico sausseriano y los modelos triádicos de los autores de orientación lexicológica.

⁴⁶ Domènech (2007: 242-243) realiza una revisión sobre las denominaciones en torno al texto especializado, en las que incluye: lenguajes/lenguas de especialización o especializadas, lenguajes/lenguas especiales, lenguajes/lenguas con propósitos específicos, textos especializados, discursos especializados...

5.5.1. El modelo de Baldinger (1952)

El modelo usa tres círculos concéntricos. De esta forma, el círculo interior queda reservado para el léxico empleado en el discurso común; el segundo, para el léxico especializado cercano al discurso común; y el último, para el léxico especializado que se aleja del discurso común y se acerca al discurso especializado.

Este modelo emplea como criterio diferenciador la cercanía del léxico al tipo de discurso. En cierta forma, avanza ya la existencia de los grados, los límites difusos y la superposición de los círculos de un mismo individuo: «da ja alle 3 Kreise im gleichen Menschen zusammentreffen können» (Baldinger, 1952: 90).

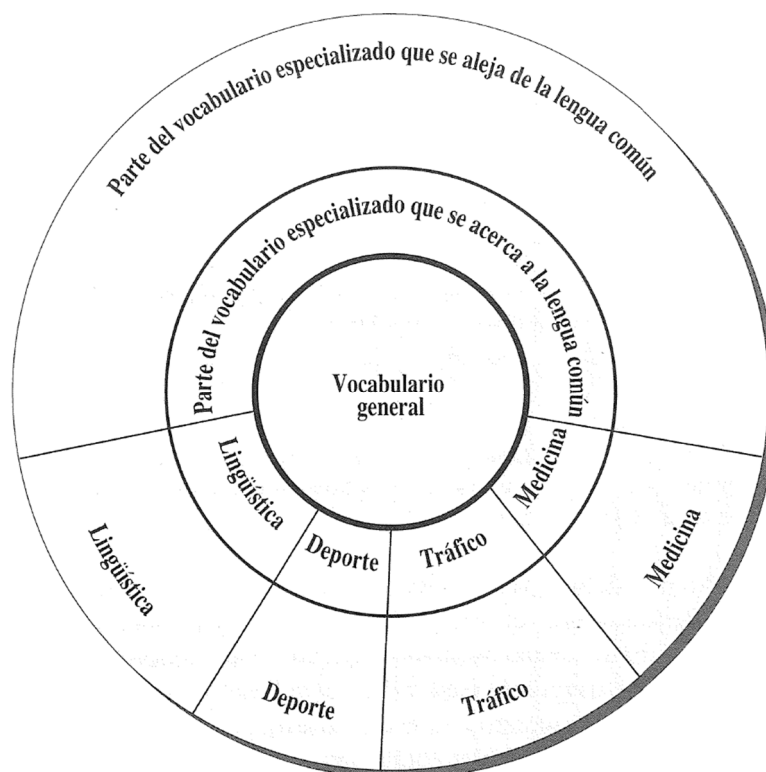


Ilustración 3: Modelo de Baldinger (1952)

5.5.2. El modelo de Reindhardt (1966)

Reindhardt (1966: 185) articula su modelo principalmente en torno al grado de especialización del vocabulario (de menor grado en el centro a mayor grado en la circunferencia), si bien atiende a cuatro criterios principales: a) el área; b) el grado de especialización; c) el tipo de medios léxicos utilizados; y d) el tratamiento de la materia (que es una perspectiva que recupera Cabré (1999: 154).

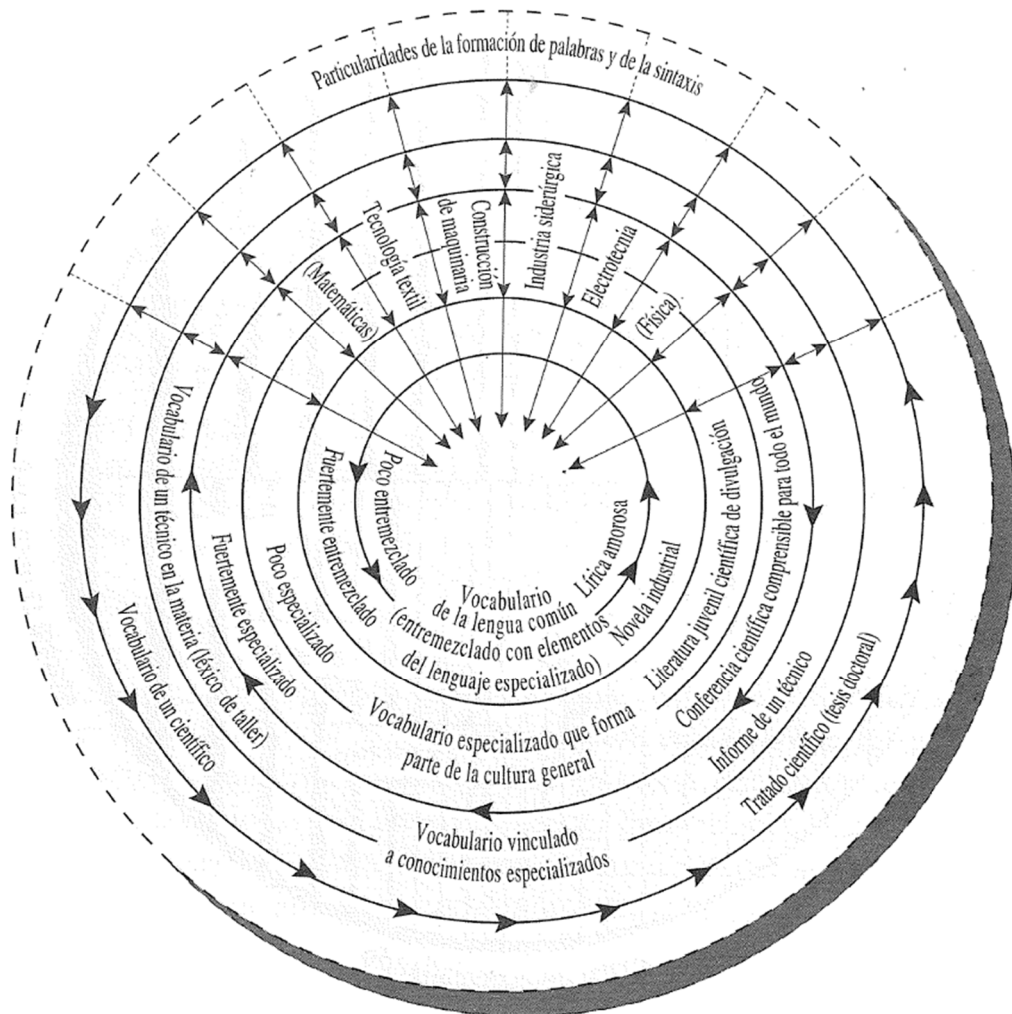


Ilustración 4: Modelo de Reindhardt (1966: 185), en Arntz y Picht (1995: 32)

5.5.3. El modelo de Heller (1970)

Heller (1970: 532) estructura su modelo en dos pares. El primer par, en la serie vertical, atiende a la especialización del léxico (especializado y no especializado). El segundo par, en la serie horizontal, incluye el criterio de comprensibilidad (comprensible o no comprensible para todo el mundo)⁴⁷. De esta forma, existirían grados tanto en la serie vertical como en la horizontal.

Heller sitúa los dialectalismos, jergalismos y extranjerismos cercanos al léxico no especializado en la serie vertical y cercano al léxico no comprensible para todo el mundo en la horizontal.

⁴⁷ Este criterio es asimilable al criterio cognitivo de la TCT.

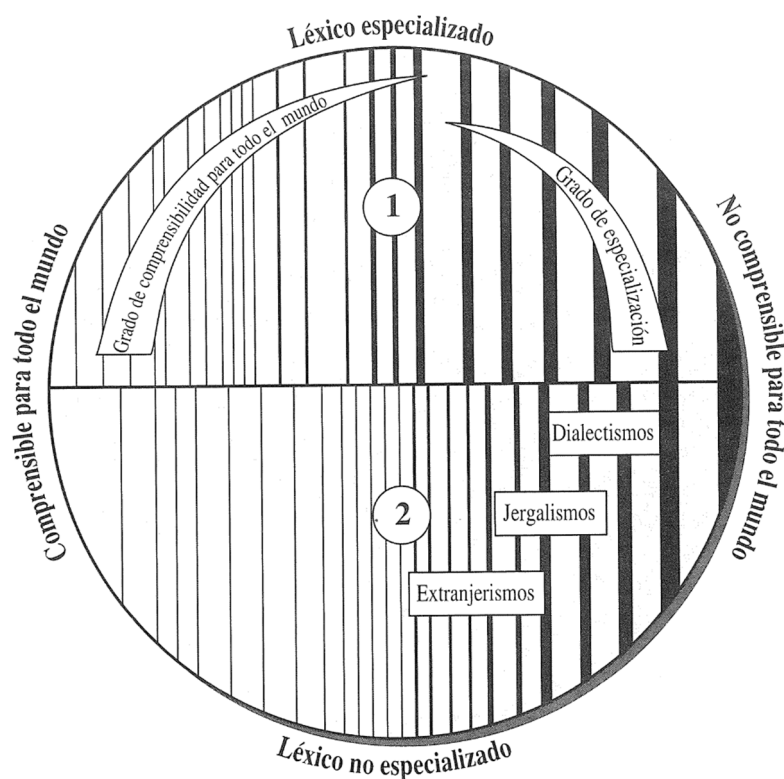


Ilustración 5: Modelo de Heller (1970: 532) en Arntz y Picht (1995: 31)

5.5.4. El modelo de Hoffmann (1985)

En su trabajo de 1985, Hoffmann diseña un modelo⁴⁸ que integra los aspectos relevantes para los diferentes discursos especializados y el discurso común, de forma que fuese posible identificar ciertos límites.

Nivel de abstracción	Forma lingüística	Ámbito	Participantes
Más elevado	Símbolos artificiales para elementos y relaciones	Ciencias fundamentales teóricas	Científico-científico
Muy elevado	Símbolos artificiales para elementos; lengua general para las relaciones (sintaxis)	Ciencias experimentales	Científico (técnico)-científico (técnico)
Elevado	Lengua natural con terminología y sintaxis controlada	Ciencias aplicadas y técnicas	Científico (técnico)-directores científico-técnicos de la producción material
Bajo	Lengua natural con terminología y sintaxis relativamente libre	Producción material	Directores científico-técnicos en la producción material-maestros, trabajadores, especialistas
Muy bajo	Lengua natural con algunos términos y sintaxis libre	Consumo	Representantes del comercio-consumidores

Tabla 6: Modelo de Hoffmann (1985)

⁴⁸ Hoffmann (1998: 31) afirma que el análisis del discurso especializado debe ocuparse sobre todo del discurso profesional y académico (partiendo de la clasificación de Gavranek, 1967: 365).

Desde nuestro punto de vista, los aciertos del modelo se hallan en la explicitación de la gradación de la serie vertical, la perspectiva de la abstracción y la relación que establece con otros parámetros como el ámbito y los participantes de la comunicación. No obstante, consideramos que dichas correspondencias son demasiado idealistas y no quedan avaladas por estudios empíricos que defiendan el modelo.

5.6. Definición y caracterización del término

En nuestra investigación, adoptamos la perspectiva⁴⁹ de la TCT sobre la definición de término, concretada en los trabajos de Cabré (1999: 132, 133 y 137; 2008: 13) y Cabré, Feliu y Tebé (2001), entre otros. Para esta perspectiva, el término es una unidad léxica poliédrica que se activa por condiciones pragmáticas en la que la forma y el contenido (es decir, la denominación y el concepto) son simultáneos e indisolubles.

- a) En cuanto a su **activación pragmática**, el término es una unidad léxica actualizable en el discurso, es decir, es una unidad discursiva; una unidad que activa un nuevo valor. En definitiva, una unidad de significación especializada. Además, Cabré *et al.* (2001: 176 y 177) consideran unas «condiciones de uso que representan una desviación⁵⁰ consciente respecto del uso referencial normal de los lexemas de la lengua general».
- b) En cuanto a su **forma**, el término está sujeto a variación y su denominación puede pertenecer al propio sistema lingüístico, a otro diferente o a sistemas simbólicos.
- c) En cuanto al **contenido**, esta parte de la «selección de rasgos adecuados a cada tipo de situación y determinados por el ámbito, el tema, la perspectiva de abordaje del tema, el tipo de texto, el emisor, el destinatario y la situación» (Cabré, 1999: 132). Es decir, el término está sujeto a variación conceptual ligada a cuestiones pragmáticas. Además, el concepto es percibido según la perspectiva del usuario especialista y es fruto de un proceso de conceptualización de la realidad que realiza una lengua. Este hecho explica las coincidencias en la percepción de la realidad (reflejada en la terminología) entre disciplinas, entre

⁴⁹ Para profundizar en las diversas definiciones que se han proporcionado sobre el término a lo largo de la investigación desde sus inicios sistemáticos en los años 70, véase el trabajo de Adelstein (2004: 72), que sintetiza en forma de tabla las propiedades del término y la palabra y sus perspectivas de abordaje en lingüística y aplicaciones terminológicas, en lingüística y cognición, y en lingüística y texto.

⁵⁰ Esta desviación también es patente en Condamines (1995: 226), que afirma que «these peculiarities do not concern many terms, but they offer the first clues to the deviant behaviour of terms as opposed to words (...)».

grupos científicos y entre discursos especializados (también entre lenguas). Los términos, pues, serían vistos como signos motivados del lenguaje, puesto que «además de denominar, significan en sí mismos⁵¹» (Cabré, 2008: 15).

A estos tres ejes se suman las funciones o componentes lingüístico, cognitivo y social/funcional (*vid. supra*), así como las funciones fundamentales de representación y comunicación.

5.7. Criterios para diferenciar *término* y *unidad léxica no especializada*

La diferenciación entre el término y la unidad léxica no especializada nace en la terminología tradicional como una necesidad epistemológica de contar con una unidad propia de análisis, basada en los postulados orientados al trabajo terminológico o a la planificación lingüística y «una perspectiva teórica más orientada a la representación del conocimiento» (Adelstein, 2004: 55).

Adelstein (*ibíd.*: 19-27) recoge en su investigación una síntesis de los criterios desde los que se ha definido el término, entre los que se incluyen el criterio definicional, el funcional, el comunicacional, el formal, el semántico, el semiótico, el cognitivo, el de enfoque disciplinar, el de origen y el de grado de especialización. Estos criterios han servido a lo largo del desarrollo teórico de la terminología como las claves de diferenciación entre *término* y *unidad léxica no especializada*.

Asimismo, defiende que la diferencia entre estos dos conceptos no es de estatuto semiótico, sino que se trata de «un uso, o bien uno de los significados posibles que adquiere una unidad del léxico en un contexto de comunicación especializada» (*ibíd.*: 72), por lo que la especificidad del término se halla en: a) la situación comunicativa de uso; b) la especialización de sus rasgos semánticos; c) el conocimiento que representa; y d) el modo de significar (*ibíd.*: 89)⁵².

Cabré (2008: 16) defiende que los términos y las unidades del léxico de la lengua participan de las mismas características y Cabré, Feliu y Tebé (2001: 304) indican que las diferencias se hallan en las condiciones cognitivas (su contenido) y sus condiciones pragmáticas (su uso) y no en las formales, semánticas o funcionales⁵³. Siguiendo estos

⁵¹ Condamines (1995: 226) sostiene al respecto: «terms also have a referential function, or to be more precise, they simultaneously have a signifying and referring function».

⁵² En Adelstein (2004: 73) se recogen los estudios que han analizado dicha especificidad.

⁵³ Formalmente, comportan las mismas estructuras y se explican por las mismas reglas gramaticales; semánticamente, ambas son unidades con significado representables mediante definiciones; funcionalmente siempre pertenecen a una de

niveles, Picht (1984: 59 y ss.) avanzó en *Métodos de terminografía y principios de investigación terminológica* que se podían excluir las diferencias morfológicas, morfológicas o sintácticas como criterios determinantes de diferenciación. Sí destaca la semántica como posible criterio diferenciador entre lo que nosotros denominamos *discurso común* y *discurso de especialización (lengua general y lengua profesional según Picht)*. Además, Picht (ibíd.: 212) determina que «las características distintivas entre *término* y *palabra* son precisión y pertinencia a un sistema de términos».

Alonso y DeCesaris (2006: 273 y 274) parten del punto de vista cognitivo como caracterizador de las ULE (unidades léxicas especializadas: representaciones lingüísticas de los conceptos especializados)⁵⁴ y proponen tres criterios para distinguir las UL y las ULE con el objetivo de determinar si son válidas y suficientes en esta diferenciación: criterio de frecuencia⁵⁵, criterio lexicográfico y criterio del especialista. Las autoras (ibíd.: 275) concluyen que tanto el criterio de frecuencia⁵⁶ como el lexicográfico (empleando diccionarios especializados) no son criterios totalmente discriminantes a la hora de distinguir las UL y las ULE⁵⁷. Tampoco consideran el criterio del especialista como un criterio suficiente, si bien las autoras creen que es el que posee la competencia cognitiva para discriminar si una UL representa o no un concepto específico. Por tanto, concluyen que:

Deben de [sic] haber otros rasgos más característicos más allá del plano léxico que permitan diferenciar el lenguaje del medio ambiente [un discurso especializado] de la lengua general [un discurso común] (Alonso y DeCesaris, 2006: 281) [los corchetes son nuestros].

Los criterios que usaremos en esta investigación se exponen en el apartado II. 8.1. (p. 126) y corresponden con un criterio subjetivo, uno textual, uno pragmático y uno cognitivo-funcional.

las categorías gramaticales existentes (es decir, no existen categorías gramaticales exclusivas para los términos); pragmáticamente, ambas participan en la interacción comunicativa (Cabré y Estopà, 2002: 145).

⁵⁴ También parten de la premisa de que no es posible establecer diferencias a partir de la forma.

⁵⁵ Consideran que una UL es una ULE si su frecuencia supera las cuatro ocurrencias en un texto.

⁵⁶ Estopà (1999: 146) rechaza la frecuencia como factor determinante del carácter especializado de una unidad «i encara és menys fiable en els textos de caràcter divulgatiu».

⁵⁷ De hecho, desde nuestro punto de vista, el criterio lexicográfico tiene el riesgo de convertirse en un criterio circular, ya que pretende defender la existencia de un término a través de su registro en el diccionario especializado pero los propios criterios de elaboración terminográfica parten del término en contexto como criterio de selección (por lo que en primer lugar hay que determinar si la unidad candidata es un término).

5.8. Variación especializada: diferencias entre situaciones de uso del léxico

Cabré (1999: 69-92) indica que la comunicación especializada y general se diferencia en la selección de unidades en las maneras de significar, en la frecuencia de funciones lingüísticas y en la organización del discurso, entre otros factores. La variada casuística y las diferentes perspectivas de análisis han favorecido la aparición y uso de diferentes denominaciones para hacer referencia a conceptos similares, como *lenguaje especial*, *lenguaje de especialización*, *lenguaje específico*, *lenguaje con fines específicos*, *lenguaje de grupo*, *lengua especializada*, *tecnolecto*, *jerga*, *argot*, etc.

Dado que las investigaciones sobre la terminología parecen coincidir en que los parámetros extralingüísticos son aquellos que determinan el discurso de especialización, cabe rescatar las palabras de Casas Gómez (2003: 570), que afirma que la variación especializada se halla como una dimensión independiente a caballo entre la diastratía y la diafasía⁵⁸, que abarca la variación jergal, la variación sectorial y la variación técnico-científica.

- a) La **variación técnico-científica** parte de una caracterización temática o perspectiva, en la que se intenta establecer una línea entre lo técnico (lo estrictamente especializado) y lo científico (a caballo entre lo común y lo especializado). Desde nuestro punto de vista, esta visión no es descriptiva ni explicativa de la realidad, puesto que no establece criterios concretos para la distinción de los dos polos, sino que queda sujeta a un criterio de percepción subjetiva.
- b) La **variación específica**⁵⁹, que afecta a los discursos sectoriales (gremios, oficios y grupos profesionales), se encuentra a caballo entre la variación jergal y la técnico-científica formando un *continuum*. De acuerdo con Rodríguez Díez (1981: 115; en García Antuña, 2009: 459), esta categoría queda caracterizada por: a) el protagonismo de los medios de masa para hacer menos críptico el argot; b) la interdisciplinariedad de determinadas áreas del discurso, que pasan a formar parte de la lengua común; y c) la individualización de los discursos por criterios sociológicos o temáticos.

⁵⁸ García Antuña (2009: 457) incluye aspectos que influyen en la diafasía, como la variación diamésica, las diferencias estilísticas y pragmáticas, los estilos de lengua, la variación diasexual y diageneracional y la distinción entre lenguaje común y técnico (lo que nosotros denominamos *discurso común* y *discurso especializado*).

⁵⁹ Esta perspectiva coincide con una de las acepciones de *jerga* del *DRAE* (2014, en línea): «[l]enguaje especial y no formal que usan entre sí los individuos de ciertas profesiones y oficios».

- c) La **variación jergal**⁶⁰, cuyas múltiples acepciones giran en torno a: a) la finalidad críptica de los discursos de grupos sociales marginados (es decir, la germanía histórica y moderna); b) la finalidad no críptica de la profesión o gremio (más cercana a la variación sectorial); y c) la caracterización de un grupo social según la actividad que realiza.

Por su parte, Picht (1984: 161) reconoce la falta de nitidez entre lo que él denomina *lenguas especiales* y *lenguas profesionales* (asimilables a *discursos sectoriales* y *discursos especializados*):

Otra interrelación queda visible al comparar el argot y la lengua profesional entre militares. Una de las diferentes esenciales entre argot y lengua profesional está en que aquél sólo da un nombre, una denominación, a una cosa ya denominada en la lengua general o profesional, mientras que ésta denomina conceptos de la profesión. Estos dos campos no siempre pueden ser separados con toda nitidez.

En efecto, Picht parte de la función comunicativa de la lengua dentro de una comunidad lingüística para hacer referencia al reflejo del aspecto sociolingüístico en las lenguas especiales (*Sondersprachen* o discursos sectoriales), que son empleadas por grupos sociales determinados por la edad, las aficiones o rasgos de actividad común (soldados, marineros, criminales, drogadictos, estudiantes, etc.). Para este autor (ídem), las características de las lenguas especiales son:

- El carácter monofuncional dentro de un marco social determinado.
- El número limitado de usuarios.
- La voluntad en el aprendizaje del tipo de discurso.
- El hecho de que no sea imprescindible para la existencia de la sociedad.

Es posible apreciar cómo esta caracterización se solapa en parte con la jerga y con el discurso especializado (o lengua profesional según Picht, 1984).

Desde nuestro punto de vista, y como conclusión de este apartado, es posible defender la variación especializada como un objeto condicionado por las variaciones de registro (diafasía) y por las variaciones del usuario (diastratía), relacionada con otros dos tipos de variación discursiva (a su vez, condicionadas por la diafasía y la diastratía). Por tanto, el eje de variación diafásico-diastrático abarcaría:

- a) El **discurso especializado**: caracterizado como medio de transmisión y comunicación científico-profesional entre los especialistas, profesionales e

⁶⁰ Esta perspectiva coincide con una de las acepciones de *jerga* del *DRAE* (2014, en línea): «[l]enguaje especial utilizado originalmente con propósitos crípticos por determinados grupos, que a veces se extiende al uso general; p. ej., la jerga de los maleantes».

investigadores de los campos de todas las ciencias (lingüistas, biólogos, ingenieros, abogados, etc.). La popularización y banalización a través de los medios, a veces en torno a un comportamiento prototípico idealizado y no necesariamente cercano a la realidad, empleado por grupos sociales concretos, pero no marginales, lleva a hablar de discursos sectoriales, pero a nuestro juicio se trata simplemente de una perspectiva de análisis diferente. Dentro de este tipo de discurso se incluyen todas las actualizaciones discursivas del grado de especialización textual, es decir, el texto especializado, semiespecializado y no especializado.

- b) El **discurso jergal**: caracterizado por la cripticidad denominativa de su discurso en torno a elementos de la realidad previamente denominados en el discurso común, vinculados a grupos sociales (a veces marginales) caracterizados por su pertenencia a una misma actividad o a un mismo sentimiento. Sería el discurso empleado por estudiantes, soldados, criminales, etc.

6. EL VALOR ESPECIALIZADO (I)

6.1. Nuevas aproximaciones: valor léxico y relevancia de la situación discursiva

La investigación en torno al concepto de *término* ha tratado de acotar su esencia de forma apriorística, intentando determinar su caracterización como término antes de conocer su propia naturaleza. Esto ha llevado a que las nuevas teorías descriptivas intenten explicar el término desde diferentes perspectivas al mismo tiempo, sin hacer referencia a una concepción estática dentro del sistema, reconociendo que «los términos no pertenecen a un ámbito, sino que son usados en un ámbito con un valor singularmente específico» (Cabré, 1999: 124)⁶¹ y que «el valor de un término se define por su relación con los demás términos que con él forman la estructura conceptual de un mismo ámbito temático» (Cabré, 2008: 14).

A nuestro juicio, la concepción de la situación discursiva y del valor activado de la unidad léxica son las dos principales perspectivas de superación de las ideas apriorísticas,

⁶¹ Cabré, Feliu y Tebé (2001: 307) reiteran esta idea cuando mencionan que «los términos no preexisten en las áreas temáticas, se utilizan en ellas» y Cabré (1999: 121) defiende que «los términos no pertenecen de manera natural a ningún ámbito» aunque reconoce «la transferencia terminológica que se produce entre ámbitos especializados o entre el registro general y el especializado [el discurso común y el especializado desde nuestro punto de vista]».

de forma que la relación entre los usuarios y la situación en la que se usan las unidades son las encargadas de activar el sentido especializado del término.

El valor constituye un elemento de la significación según Saussure (2007 [1916]: 138), que surge al comparar valores similares por oposición. Dicho de otra forma, el valor funciona dentro de un sistema de signos que establecen relaciones de forma opositiva. De esta forma, los valores de un signo no se definen positivamente por su contenido, sino de forma negativa por las relaciones que establecen con otros elementos del sistema (es decir, un signo es lo que otros no son; *ibíd.*: 141). Este valor no queda establecido por el individuo, sino que requiere el uso y consenso colectivo de los hablantes (*ibíd.*: 137).

La TCT explica los términos desde una teoría del léxico como valores de una unidad léxica (Cabré *et al.*, 2001: 305). Dado que no existen diferencias formales entre *término* y *unidad léxica no especializada* y cada uno de estos objetos es empleado en situaciones discursivas diferentes y aceptadas socialmente por sus usuarios, es posible tener en cuenta que una misma forma, un mismo significante, pueda actuar como *término* o como *no término* según su situación discursiva (Adelstein, 2004: 5). Es decir, el término sería un valor (no una unidad diferenciada en sí misma), un parámetro semántico-cognitivo (Adelstein, 2004: 7) condicionado pragmáticamente⁶² (condicionado por la situación discursiva) (Cabré y Estopà, 2002: 146). Cabré *et al.* (2001: 177) determinan que la especificidad del término está basada en criterios semántico-pragmáticos. Estos condicionantes sientan las bases del principio del valor terminológico (Cabré, 2008: 18).

Este principio se inspira en la noción de dinámica del sentido de Victorri (1996), que hace posible hablar del valor especializado como *sentido situado*, es decir, «al uso específico que se hace de estas unidades en un campo de especialización preciso» (Cabré, 2002b: 29).

En definitiva, el uso de una unidad léxica concreta en una situación discursiva específica puede dar como resultado que dicha unidad léxica adquiera o active un sentido particular preestablecido por la comunidad especialista, i. e. que adquiera o active un valor de contenido especializado (Cabré 2002a: 91; 2002b: 29). Además, cabe destacar que una misma unidad léxica puede activar, no solo uno, sino diferentes valores especializados

⁶² Alonso y DeCesaris (2006: 281) defienden la pragmática como criterio óptimo para indicar el valor especializado de una unidad apoyándose en las palabras de Cabré (1999: 26): «los aspectos pragmáticos son los que mejor permiten diferenciar los términos de las palabras. Pragmáticamente, términos y palabras se distinguen i) por sus usuarios; ii) por las situaciones en que se utilizan; iii) por la temática que vehiculan, y iv) por el tipo de discurso en que suelen aparecer».

según la situación temática (variación horizontal)⁶³ y diferentes grados de precisión, concisión y científicidad en dicho valor especializado (variación vertical) (Cabré y Estopà, 2002: 146). Estas características vienen dadas por el carácter discursivo del término, que permite activar sentidos diferentes en situaciones diferentes (Adelstein, 2004: 60 y 68). Según Cabré, Feliu y Tebé (2001: 306):

El valor terminológico o valor especializado de una unidad se activaría cuando las condiciones de un contexto de comunicación profesional así lo requieran. Esta activación consistiría en la activación de un sentido integrado por una selección de rasgos semánticos precisos y correspondientes al sentido especializado de una unidad en un ámbito especializado determinado. Al mismo tiempo, quedarían inhibidos otros sentidos conformados por una selección de estos mismos rasgos junto con otros no activados.

Una de las cuestiones que suscita debate es la consideración de comunicación profesional. Dicho de otra forma, el uso de ciertos términos fuera de sus situaciones prototípicas de uso, por ejemplo, en textos de divulgación de mayor o menor especialización, quedaría relegado a la consideración de *unidad léxica no especializada* y no de *término*.

Adelstein (2007a: 276 y 277), en su tesis doctoral, determina los tipos de factores que intervienen en la generación del significado especializado, entre los que se incluyen:

- a) La especificidad del significado especializado reside en la imbricación de información semántica y especificidad referencial.
- b) Dicha especificidad está condicionada por factores de la situación comunicativa, por factores oracionales y por factores sintáctico-semánticos locales.
- c) El campo temático y la perspectiva sobre el tema textual, y la situación comunicativa, en concreto los interlocutores, son los factores principales que determinan la generación del significado especializado.

Además, concluye que: (i) no puede sostenerse una relación determinista entre texto especializado y léxico de especialización, sino que esta relación debe concebirse en términos probabilísticos (lo que coincide con una de nuestras hipótesis y objetivos); y (ii) para determinar el significado especializado hace falta recurrir a las secuencias mínimas en las que ocurre la unidad léxica y no solo al nivel de especialización del texto⁶⁴ (Adelstein, 2007a: 280 y 281).

⁶³ Adelstein (2004: 82) ilustra estas diferencias temáticas: «piénsese en los “valores de palabra” que puede adquirir *cabeza* y en los “valores de término” que tiene en anatomía, geografía, informática o electricidad».

⁶⁴ Desde nuestro punto de vista, considerando el léxico como uno de los factores clave para determinar el grado de especialización, sería difícil argumentar que una unidad léxica activa un valor especializado recurriendo el grado de

Partiendo de la variación vertical del valor gradual de especialización, cabe considerar si este valor está ligado a la unidad de forma fija (aunque pueda ir cambiando en la diacronía según diversos condicionantes) o si, por el contrario, depende del usuario, es decir, de la percepción que tenga el usuario de la unidad. Piénsese por ejemplo en la unidad *zika* en los siguientes textos:

- (11) Una nueva enfermedad amenaza a las Américas, se trata de la infección por el **virus Zika** (VZIK), perteneciente a la familia Flaviviridae, genero flavivirus, transmitida por mosquitos del género Aedes. [...] Clínicamente es similar al dengue y la chikungunya, se caracteriza por mialgias, cefalea, artralgias principalmente de manos y pies, conjuntivitis no purulenta, las complicaciones hemorrágicas y neurológicas son raras. La mayoría de las infecciones son asintomáticas, y no se han reportado fallecimientos (Carvajal, Peña y Oletta, 2015)⁶⁵.
- (12) Los países afectados por el **zika** en Sudamérica están registrando una escalada inusitada en las peticiones de aborto, según un estudio publicado hoy. El **virus del Zika** solo causa síntomas leves en la mayoría de los casos, pero también está asociado a malformaciones cerebrales en el feto (Domínguez, 2016)⁶⁶.

Con total seguridad, consideramos que un lego en medicina (es decir, un usuario medio y común de la lengua) no estará familiarizado con la unidad léxica. Esta falta de familiaridad puede conllevar un esfuerzo cognitivo superior a otros discursos. Es decir, *zika* y sus variantes serían percibidos como una unidad léxica de alta densidad cognitiva⁶⁷. En ambos contextos consideramos que sería posible defender el valor especializado de la unidad, si bien su sentido y el sistema conceptual no son transparentes para el especialista y para el lego en la misma medida: para el especialista es un objeto específico analizable y para el lego es un objeto difuso, borroso cognitivamente. Desde nuestra perspectiva, el criterio de percepción del usuario o criterio subjetivo es el punto de partida para determinar el valor especializado. De hecho, Estopà (1999: 154) menciona «[h]em remarcat en negreta la part del segment que *creiem* que té valor terminològic [la cursiva es nuestra]»; es decir, que, a partir de la información situacional y atendiendo a los efectos cognitivos de las unidades léxicas, los usuarios son capaces de indicar, en un intervalo de valores graduales

especialización textual, ya que precisamente este grado se clasifica según diferentes valores y atributos (mayor o menor grado de especialización) dependiendo del grado de densidad léxica especializada del texto. En otras palabras: determinar que un texto es especializado implica tener en cuenta el nivel léxico, en concreto, la densidad léxica especializada; no obstante, determinar que una unidad léxica es especializada no implica que el texto en el que se inserta sea especializado, por lo que justificar el valor especializado de una unidad léxica exclusivamente a través de su situación textual podría conllevar ciertos problemas teórico-metodológicos.

⁶⁵ Fuente: CARVAJAL, Ana C.; PEÑA, Saúl O.; OLETTA L., José Félix (2015): «Infección por Virus Zika (VZIK): Arbovirosis emergente en las Américas», *Medicina Interna* 31(1), pp. 8-15. Disponible en: http://svmi.web.ve/wh/revista/V31_N1.pdf.

⁶⁶ Fuente: DOMÍNGUEZ, Nuño (22 de junio de 2016): «La alerta por el zika dispara las solicitudes de aborto en Sudamérica», *El País*. Disponible en: http://elpais.com/elpais/2016/06/22/ciencia/1466616529_299644.html.

⁶⁷ También recibe la denominación de *densidad conceptual* o *densidad informativa* según Cabré *et al.* (2007: 852-853).

y posibilidades, si consideran que una unidad léxica activa o no un valor especializado, es decir, si se considera o no *término*.

No obstante, existen dos casos problemáticos, a nuestro juicio, a la hora de emplear este criterio subjetivo:

a) Aquellas unidades léxicas que suelen usarse con mayor frecuencia en el discurso común y que con el paso del tiempo y el desarrollo científico acaban por terminologizarse: acaban activando un nuevo sentido (especializado) en una nueva situación de uso. Tan solo el usuario especialista sería capaz de vincular la denominación de estas unidades con un nuevo sentido especializado, ya que el usuario lego podría no identificar este nuevo sentido, este nuevo valor. Un ejemplo sería la unidad *bueno* en derecho (anteriormente expuesta), o la unidad *duro* o *espíritu* en lingüística.

(13) Concurren, en verdad, como asesores a fin de evitar avenencias o transacciones precipitadas o suicidas para los intereses de la parte, (v. el art. | | de la ley cit.) Sobre el acto de conciliación, tanto el demandante como el demandado deben concurrir acompañados cada uno por un hombre **bueno**, que han de ser españoles y gozar plenamente de sus derechos civiles (s. a., 2015)⁶⁸.

(14) **Duro**. En fonética utilizada en la descripción de las lenguas eslavas se usa el término ‘duro’, de carácter impresionista, como opuesto a blando o mojado, para aludir a las consonantes que no tiene palatalización, o que es muy escasa debido a que la lengua, debido a una intensión fuerte, apenas tiene contacto con el paladar duro. En ruso el rasgo duro se representa con [b] y el blando con [b̥] (Alcaraz Varó y Martínez Linares, 2004: 220)⁶⁹.

(15) A diferencia de la aspiración vocálica interna, la inicial, quizá por una cierta analogía con el **espíritu** áspero griego, que tenía el mismo valor y derivaba del mismo signo, no desapareció tan pronto ni de modo tan radical. Pero es fácil imaginar que la fonética sintáctica, es decir, la cadena fónica —no perdemos de vista que el latín fue una lengua hablada— daría lugar en numerosísimos casos a posiciones intervocálicas ocasionales o contextuales, que favorecerían su enmudecimiento (Fernández Martín, 1995: 278)⁷⁰.

b) Aquellas unidades léxicas con valor especializado que, debido a su divulgación a través de los medios, han acabado por integrarse en el discurso común, de forma permanente (como *ADN*, *virus*, *gripe* o *avión*) o efímera (como *prima de riesgo* o *aforamiento*) (*vid. infra*). El problema teórico que afecta a estas unidades es si deben considerarse siempre unidades léxicas especializadas (aunque con diferentes grados según el tipo de discurso) o si es posible que

⁶⁸ Fuente: [s. a.] (6 mayo 2015): «Hombre bueno», *UniversoJus.com*. Disponible en: <http://universojus.com/definicion/hombre-bueno>.

⁶⁹ Fuente: ALCARAZ VARÓ, Enrique y MARTÍNEZ LINARES, M.^a Antonia (2004): *Diccionario de lingüística moderna*. Barcelona: Ariel.

⁷⁰ Fuente: FERNÁNDEZ MARTÍN, Concepción (1995): «Límites precisos de la aspiración inicial en latín», *Habis*, 26, pp. 271-280. Disponible en: <https://dialnet.unirioja.es/descarga/articulo/58025.pdf>.

algunas unidades hayan perdido totalmente dicho valor especializado en el discurso común (*virus, avión, vacuna, gripe...*).

Estas unidades suponen un problema de carácter teórico y metodológico para la extracción terminológica, tanto manual como automática, puesto que la denominación no permite discriminar el valor especializado del texto si no se conoce de antemano la situación comunicativa en la que se inserta el término.

6.2. Dificultades para la identificación del término: denominación, comprensión y coste de procesamiento

El modelo de Adelstein (2007b) para analizar el valor de especialización, basado en una versión adaptada del modelo de estructura semántica de Pustejovsky (1995), resulta riguroso y exhaustivo para determinar el valor de especialización. Sus principales desventajas con respecto a nuestras necesidades metodológicas son el tiempo invertido en determinar el valor de especialización y su restricción al discurso especializado. En nuestro modelo integrador debe primar la rapidez en la selección de unidades candidatas a USE; debe servir de modelo o de ejemplo para los usuarios de los textos, para que puedan determinar rápidamente los términos.

Si continuamos atendiendo al mencionado criterio subjetivo en relación con la densidad cognitiva de la unidad léxica, es posible recurrir al argumento de comprensibilidad de una unidad léxica, bien por exclusivo desconocimiento del significado de la unidad, bien por desconocimiento del campo temático. Véanse los siguientes ejemplos, en los que podrían identificarse los tres niveles de variación propuestos anteriormente:

- (16) Desde hace algunos meses vengo oyendo un uso del verbo ‘**trolear**’ que ha trascendido el ámbito en el que nació. En internet, se utiliza el término inglés ‘troll’ para referirse a un usuario que se dedica dejar comentarios de dudoso gusto en blogs, foros, canales de Youtube y cualquier tipo de página que cuente con un sistema de comentarios con el único propósito de sembrar la discordia, soliviantar al personal e intentar amedrentar al propietario del espacio para dejar sus mensajes llenos de insultos, palabrotas, tomaduras de pelo, amenazas, etc. [...]. Así pues, ‘**trolear**’, ya con la grafía adaptada al español, pues no tiene la ‘ll’ del inglés, se ha estado utilizando durante los últimos años para referirse a la actuación, en internet, de estos usuarios (García Guerra, 2012)⁷¹.

⁷¹ Fuente: GARCÍA GUERRA, Miguel Ángel (19 de febrero de 2012): «Nuevo verbo: ‘trolear’». Disponible en: <http://www.magarciaguerra.com/2012/02/nuevo-verbo-trolea/>.

- (17) En la Universidad de Extremadura, en Cáceres, estamos organizando un congreso sobre literatura española que rendirá homenaje, en esta primera edición, al fallecido profesor Juan Manuel Rozas en el trigésimo aniversario de su **óbito**⁷².
- (18) Se asume que en los cortometrajes es muy posible que una persona física actúe como **productor**-empresario individual, y a la vez sea el guionista y el director de su película. En este supuesto, su remuneración como guionista o director se calcula de forma teórica, en aplicación de unas tablas publicadas por el ICAA anualmente junto a la convocatoria (MECD, 2010: 6)⁷³.
- (19) Contrato específico suscrito entre la empresa productora y el **productor ejecutivo**, así como la nómina referente a dicho contrato y documento acreditativo de la identidad de la persona a que se refiere. En el caso de existir contrato mercantil, se aportará el contrato y la factura correspondiente (MECD, 2010: 12)⁷³.

En el ejemplo (16) parece claro que el ejemplo de *troleo* corresponde con la jerga empleada en el entorno digital (y cada día más fuera de la red, en la conversación cotidiana entre los usuarios del mundo digital y la redes sociales, a modo de imitación del discurso *online*), puesto que se trata de una unidad léxica no sujeta a control de los especialistas y restringida a un uso social concreto. Por otra parte, el caso de *óbito*, en el ejemplo (17) sería difícil encuadrarlo dentro de la jerga o de un discurso sectorial, puesto que, a simple vista, podríamos hablar de un cultismo empleado en un nivel de variación fasolectal elevado. Aunque tampoco sería posible defender con el fragmento que aquí tenemos que se trate de un discurso especializado. Por último, en el caso de *productor* y *productor ejecutivo*, en los ejemplos (18) y (19), que representan conceptos diferentes, ¿se consideran términos si aparecen en los créditos de una película?, ¿y en el texto extraído de la legislación? En ambos se codifica el mismo sentido, pero parece que, aun existiendo un control conceptual similar, algunos usuarios considerarían que no son unidades léxicas especializadas en el contexto de la película. En definitiva, destacamos en esta parte que cuando hablamos de que se activa el valor especializado de una unidad léxica, nos referimos a que se activa uno de los sentidos del conjunto de acepciones posibles con los que puede contar una unidad en un contexto, condicionada esta acepción por la situación comunicativa.

Esta relación de los términos con sus situaciones de uso y con la densidad cognitiva lleva a cuestionar si dicha densidad siempre tiene las mismas características independientemente de la situación de uso de la unidad. En otras palabras, ¿la densidad cognitiva o coste cognitivo de *bueno*, *espíritu* o *duro* en el discurso común es similar a la

⁷² Fuente: Correo personal, 2016.

⁷³ Fuente: MECD – MINISTERIO DE EDUCACIÓN, CULTURA Y DEPORTE (2010): Criterios interpretativos en materia de reconocimiento del coste de una película, en aplicación de la orden CUL/2834/2009, de 19 de octubre, por la que se dictan normas de aplicación del real decreto 2062/2008, de 12 de diciembre, por el que se desarrolla la ley 55/2007, de 28 de diciembre, del cine, en las materias de reconocimiento del coste de una película e inversión del productor, establecimiento de las bases reguladoras de las ayudas estatales y estructura del registro administrativo de empresas cinematográficas y audiovisuales. Disponible en: <http://www.mecd.gob.es/cultura-mecd/dms/mecd/cultura-mecd/areas-cultura/cine/informacion-general/conceptos-cine-y-audiovisual/coste-de-peliculas/criterios-interpretativos-costes.pdf>.

que posee en el discurso especializado? ¿Suponen el mismo coste de procesamiento que *ADN*, *prima de riesgo* o *virus*? Parece que la percepción de la densidad cognitiva depende de:

- a) si la unidad parte del discurso común y tiende al discurso especializado (casos de *bueno*, *duro* o *espíritu*). En este caso, parece que la densidad cognitiva o densidad de procesamiento solo es mayor cuando se activa el valor especializado en una nueva situación.
- b) si la unidad parte del discurso especializado y tiende al discurso común (casos de *ADN*, *prima de riesgo* o *virus*). En este caso, parece que la densidad cognitiva solo es mayor cuando la unidad se emplea en el discurso común, lo que puede hacer que se use con cierta imprecisión conceptual⁷⁴.

Cabré (1999: 132), al respecto, menciona que «[la comunicación especializada] admite niveles de especialización diferentes, grados de opacidad cognitiva variados, índices diversos de densidad cognitiva y terminológica [...]». Cuando estas dificultades de inteligibilidad afectan a la vida del usuario (por ejemplo, en el campo jurídico, como mencionan Arntz y Picht (1995: 40)), puede interrumpirse la función comunicativa del texto:

El signo sin carga semántica, sin contenido, resulta sin valor al igual que un concepto sin signo. De ahí que el signo, para tener valor comunicativo, debe ser portador de un contenido y el contenido del signo debe ser conocido por las dos partes en el proceso de comunicación (Picht, 1984: 210).

Dichos problemas de comprensión están relacionados con la reducción del esfuerzo en la transmisión de la información (Sager, 1990: 105), lo que tiene conexión con el denominado Principio de Relevancia postulado por Sperber y Wilson (1995: 260), que defiende que:

Todo acto de comunicación ostensiva comunica la presunción de su relevancia óptima (principio comunicativo de relevancia). La cognición humana tiende a adaptarse hacia la optimización de la relevancia (principio cognitivo de relevancia) (Pons, 2004: 21).

⁷⁴ Arntz y Picht (1995: 40) afirman que a medida que un texto aumenta su grado de especialización, disminuye su comprensibilidad general, lo que también disminuye el número de usuarios que comprenden dicho texto. Por el contrario, cuando se incorporan al caudal léxico del discurso común algunas unidades léxicas normalmente usadas en el discurso especializado, «el oyente comprende con imprecisión o erróneamente muchos de los tecnicismos difundidos por los medios de comunicación y no puede emplearlos correctamente. Puede hablarse del fenómeno de *desterrminologización*».

El problema se encuentra en que, aunque un texto pretenda constituirse como acto de comunicación ostensiva, los factores cognitivos del destinatario (es decir, su esfuerzo de procesamiento) pueden interferir en el reconocimiento de un signo ostensivo, haciendo que dicho texto no se considere relevante, ni desde el punto de vista cognitivo ni desde el punto de vista comunicativo. Recordemos que en 1986 (p. 25), Sperber y Wilson afirmaban que en una situación concreta un supuesto se considera relevante cuando el esfuerzo requerido para procesarlo en esa situación es pequeño, por tanto, el Principio de Relevancia es un principio inconsciente ligado al conocimiento del destinatario.

6.3. Denominaciones de la terminología en torno al valor especializado

Dado que las investigaciones en torno a las situaciones de comunicación especializada han acabado por demostrar que existen otras unidades que codifican y transmiten el conocimiento especializado, la terminología como campo científico optó por acotar nuevos conceptos con nuevas denominaciones. Así, en primer lugar, surge, a nuestro juicio desde una perspectiva semántica, el concepto de *unidad de significación especializada* (USE) (Estopà, 1999: 25; Cabré *et al.*, 2001: 307; 2001b: 306) y, en segundo lugar, desde una perspectiva cognitiva, el concepto de *unidad de conocimiento especializado* (UCE) (Domènech, 1998; Cabré *et al.*, 2001b: 304; Cabré y Estopà, 2002: 150; Cabré, 2002a: 91): Las UCE se caracterizan por ser:

Unidades de distinto nivel descriptivo que constituyen los nudos de conocimiento de un texto o forman parte de ellos. La condición esencial para considerar que una unidad es una UCE es el tipo de conocimiento que transmite (condiciones cognitivas y semántica) y su uso en el discurso (condiciones pragmático-discursivas) (Cabrè y Estopà, 2005: 77).

Según Cabré *et al.* (2001b: 306 y 307) y Estopà (1999: 25), las USE se dividen en USE lingüísticas (ya sean códigos naturales o artificiales) (Cabrè y Estopà, 2005: 78) y no lingüísticas, entre las que se incluyen símbolos, nomenclaturas y otras fórmulas de notación icónica o gráfica⁷⁵.

Dentro de las formas lingüísticas, tanto las USE como las UCE, se concretan en diferentes unidades que giran en torno a un concepto clave: la unidad léxica especializada (ULE) (Alonso y DeCesaris, 2006: 271), también denominada unidad terminológica

⁷⁵ Como las USE no lingüísticas están alejadas del léxico, no las trataremos en este trabajo.

(UT)^{76,77} o término, caracterizada por su lexicalización, i.e. por estar gramaticalizadas como piezas léxicas (Cabré, 2002a: 100; Cabré y Estopà, 2005: 90) y por ser prototípicamente nominativas (Cabré, 2002a: 91). Cuando la USE/UCE es menor que la ULE, se habla de unidades morfológicas especializadas (UME), y cuando es mayor, se habla de unidades fraseológicas especializadas (UFE) (Adelstein, 2004: 77). A modo de síntesis podríamos concluir este apartado usando el siguiente esquema:

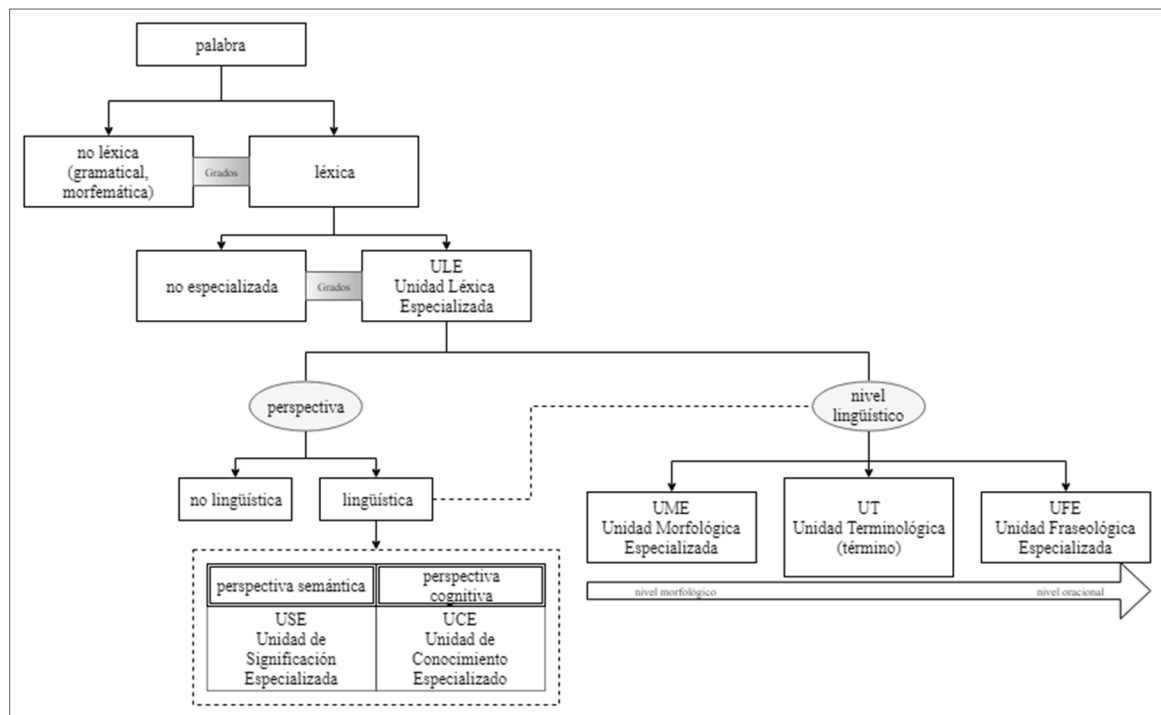


Ilustración 6: Esquema de denominaciones en torno al valor especializado [elaboración propia]

6.4. Fronteras del conocimiento especializado

Como viene comprobándose en este trabajo, resulta complejo establecer una línea divisoria entre lo que se considera especializado y lo que no, ya sea refiriéndonos al conocimiento o a la lengua (Cabré, 1999: 189). Por tanto, de nuevo, cabe recurrir a la noción de gradualidad entre lo general y lo especializado (Cabré *et al.*, 2007: 856), pues el conocimiento no puede tratarse de forma aislada, debido a su falta de uniformidad (Cabré y Estopà, 2002: 143 y 144), ya que «la ciencia expresa el conocimiento sobre la realidad natural y social de diferentes formas» (Estanny, 2001: 39).

⁷⁶ En Estopà (1999: 26, 29) las UT son las unidades de significación especializada lingüísticas nominales.

⁷⁷ Para Cabré y Estopà (2005: 151 y 152), la unidad terminológica es la «unidad léxica, cuya estructura corresponde a una unidad léxica de origen o producto de la lexicalización de un sintagma, que posee un significado específico en el ámbito al que se asocia y es necesaria en la estructura conceptual del dominio del que forma parte. Así pues, las condiciones que cumple una unidad para ser una UT son tres: a) estructura; b) especificidad semántica; y c) necesidad en la estructura conceptual».

Adelstein (2004: 27) esboza de forma acertada las cuestiones de las que partimos en este apartado:

¿Qué se entiende por especializado? ¿y por general? ¿A qué se aplican, realmente, estos adjetivos? ¿Al uso, al contenido transmitido o al conocimiento asociado? ¿A la situación o al contexto? ¿O, más bien, a todo ello conjuntamente? Creemos que, si bien el grado de especialización es para algunos autores el criterio de distinción por excelencia, se trata, en cambio, de un “macro criterio” que abarca la mayoría de los expuestos más arriba.

Cuando nos referimos a la dicotomía de *conocimiento común-conocimiento especializado* (siguiendo la misma terminología que para *discurso común y discurso especializado*), es posible partir de las relaciones entre denominación y sistema conceptual que establece la interdependencia de los dos conocimientos. Cabré (1999: 131) sintetiza, a partir de la adquisición de conocimiento, nuestro punto de partida en las siguientes palabras:

Ello no significa que esta adquisición parta siempre de cero, sino que admite distintos grados de solapamiento con el conocimiento adquirido previamente por el hablante: unas veces se aprende desde cero una denominación y un concepto no conocidos anteriormente; otras se aprende un nuevo concepto que el hablante no poseía y se asocia a una unidad formal ya interiorizada con otro sentido y/o aparejada a condiciones pragmáticas distintas; en otros casos se redistribuye la información conceptual que el hablante ya tenía asociada a una unidad denominativa; esta redistribución puede producirse en forma de aumento, pérdida o cambio de rasgos de contenido, o simultáneamente en cualquiera de sus combinaciones; otras veces se asocian únicamente nuevas condiciones pragmáticas a unidades conocidas previamente.

6.5. Aproximación a su definición

Temmerman (2001: 83) coincide con las palabras de Cabré (*idem*) cuando defiende que «terms are motor in the process of understanding as they link new understanding to previous understanding». La reflexión en torno a lo especializado parece coincidir en varios factores fundamentales: a) que se trata de un conocimiento reformulado, en tanto que parte del conocimiento previo que poseen los usuarios; b) que se caracteriza por su contemporaneidad y obsolescencia (Gutiérrez Rodilla, 1998: 2896; en Contreras, 2008: 21); y c) y en que están motivados por la evolución de la ciencia y la técnica (Gutiérrez Rodilla, 1998: 88; en Contreras, 2008: 102). En efecto, según Lara (1999: 51) «el conocimiento especializado surge ante requerimientos de precisión, de mayor

discriminación del conocimiento, ligados a necesidades económicas, jurídicas, políticas, etc. [...] por una mayor discriminación basada en el razonamiento y en operaciones más deductivas, basadas en teorías científicas» (Adelstein, 2004: 65).

Siguiendo a Temmerman (2001: 82), que se basa en el entendimiento de la realidad a través de marcos y modelos cognitivos (Fillmore, 1985; Fillmore y Atkins, 1992; Lakoff, 1987), puede decirse que existen tres factores que pueden contribuir a explicar el desarrollo de las unidades de entendimiento (y, por ende, las relaciones entre conocimiento común y conocimiento especializado):

- La necesidad de un mejor entendimiento.
- La interacción entre diferentes usuarios de la lengua.
- La estructura prototípica en la comprensión de categorías que puedan observarse de manera simultánea como resultado de y como una de las causas de la evolución semántica.

Precisamente los modelos cognitivos mencionados defienden la estructuración del conocimiento en relaciones y nodos (o conceptos), de forma que «tenir més coneixement d'un domini específic significa tenir una estructura d'aquell domini més elaborada, i aquesta elaboració consisteix a conèixer més conceptes (nodes) i més relacions (links) entre conceptes» (Domènech, 2006: 65).

6.6. Criterios para la distinción teórica del conocimiento especializado

En el trabajo de Rodríguez-Tapia (2016: 989) nos aproximamos a la posición de Cabré *et al.* (2001) y caracterizamos las situaciones especializadas donde tienen lugar los textos especializados por la representación y transmisión del denominado *conocimiento especializado*:

[...] partimos del hecho de que el conocimiento no es un objeto estático que pueda identificarse de forma unívoca con una disciplina, sino que se trata de una red de conexiones interrelacionadas entre sí, lo que explica, por ejemplo, el surgimiento de nuevas disciplinas, como la socioterminología, o de nuevos mecanismos de transmisión de información, como la banalización y la terminologización. En esta red de imbricaciones, Cabré *et al.* (2001a) distinguen dos tipos de conocimiento: *especializado* y *general* (al que nosotros denominaremos *conocimiento común* siguiendo la calificación que otorga Hoffmann (1999 [1985]) a las lenguas o discursos: *Gesamtsprache* (lengua general), *Fachsprache* (lengua especializada, o desde nuestro punto de vista, discurso especializado) y *Gemeinsprache* (lengua común).

A pesar de que no es posible hallar fronteras entre estos dos tipos de conocimiento, sí que consideramos viable acercarnos a las características esenciales que pueden atribuirse al conocimiento especializado y al conocimiento común. Si partimos de las características de la comunicación especializada y de las situaciones de comunicación, atendiendo a las condiciones de emisión, de recepción, la función lingüística básica y la temática (Cabré y Estopà, 2005: 69), podemos determinar que:

- a) En relación con las **condiciones de emisión**: el emisor debe controlar de forma óptima la estructura conceptual del ámbito (ídem; Cabré, 2002b).
- b) En cuanto a las **condiciones de recepción**, y, en especial, en cuanto al destinatario: Cabré y Estopà (ídem) defienden que el CE queda caracterizado por «el hecho de tener expectativa de recibir información, de ser informado, expectativa que no se da en otras situaciones comunicativas cuyo propósito general es de otra índole».
- c) Desde el punto de vista de la **función básica**: la comunicación especializada se inclina hacia la representación del conocimiento. Puede diferenciarse del discurso común, ya que este puede poner en funcionamiento diferentes funciones (como la emotiva, la estética, etc.).
- d) En cuanto a la **temática** y, en concreto, en relación con el conocimiento transmitido: solo es posible clasificar un conocimiento como especializado «si se representa y si se transmite respetando escrupulosamente la estructura conceptual del campo en cuestión» (Cabré y Estopà, 2005: 70). Estas autoras (ídem) asocian al texto especializado la «estructura informativa, mejor denominada cognitiva, que representa el conocimiento transmitido por el texto», así como la estructura formal, que se relaciona con el género textual de cada texto) y la gramatical. La aparición o diferente concentración de UCE en un texto dependerá, pues, de la estructura formal (es decir, de la tradición discursiva que establecen los géneros textuales, a nuestro juicio) y de la estructura cognitiva.

El trabajo de Cabré *et al.* (2001: 178) centra su atención en las diferencias entre CE y CC y en él se defiende que «el conocimiento general [para nosotros conocimiento común], en analogía con la competencia lingüística, es la información que el individuo medio comparte con los otros en virtud de su pertenencia a una comunidad, una cultura y una época determinadas», de forma opuesta al conocimiento especializado, que se caracteriza «por una estructuración más sistemática de sus contenidos» (ibíd.: 174) y una

perspectiva diferente desde la que se trata el tema (ibíd.: 178) (Rodríguez-Tapia, 2016a: 989; Rodríguez-Tapia y Moreno Paz, en prensa).

Asimismo, es posible defender que el CE no constituye un conocimiento completamente nuevo, puesto que nace de un conocimiento previo (del CC), es decir, se trata de una reformulación en diferentes grados del CC previo, (lo que permite explicar la complejidad para delimitar fronteras. Por este motivo, desde nuestro punto de vista (ídem), es posible percibir una relación de inclusión del CE y del CC dentro de la categoría *conocimiento general*.

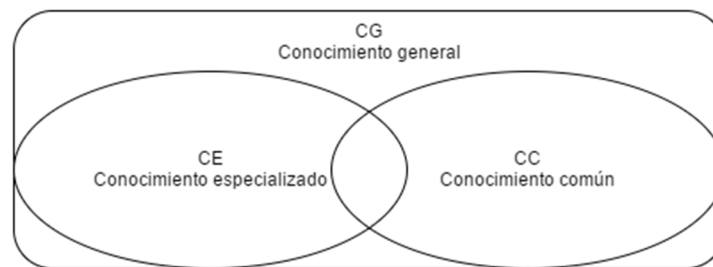


Ilustración 7: Esquema de relación entre conocimiento general, conocimiento común y conocimiento especializado

A modo de síntesis del mencionado trabajo de Cabré *et al.* (2001), y seleccionando los datos que más nos interesan en nuestro estudio, esbozamos la siguiente tabla (Rodríguez-Tapia y Moreno Paz, en prensa):

CONOCIMIENTO COMÚN	CONOCIMIENTO ESPECIALIZADO
Es el conocimiento compartido por una cultura, adquirido por el individuo desde su inserción y convivencia en la sociedad, y consolidado de manera no sistemática a partir de la experiencia con el mundo . Se estructura según el tiempo y los procesos históricos que permiten organizar la información según lo que aporta la experiencia al individuo .	No forma parte del bagaje común del individuo medio y representa una segmentación voluntaria y consciente de una parcela de la realidad . Asimismo, se inserta en contextos de instrucción explícitos y su adquisición parte de la reformulación del conocimiento general a partir de la experiencia del individuo con el mundo y se basa principalmente en la aceptación de hipótesis y enunciados que se basan en teorías sobre qué es y cómo funciona la realidad. Estas hipótesis se someten a prueba empírica, coherencia lógica o funcionalidad práctica .

Tabla 7: Características del conocimiento común y del conocimiento especializado (adaptado de Cabré *et al.*, 2001: 174 y ss.)

La diferencia principal, pues, se hallaría en la estructuración sistemática del CE con respecto del CC (Cabré *et al.*, 2001: 174). Al relacionar esta idea con la noción de valor

especializado, sería posible determinar que dicho valor quedaría activo si la unidad transmite conocimiento especializado. No obstante, ¿siempre que queda activo lo hace respetando escrupulosamente la estructura conceptual? ¿Qué sucede si el término se usa con fines divulgativos o didácticos o si se usa de forma incorrecta e imprecisa (destermnologización y pseudosinonimia)? (*vid. infra* apartado II. **¡Error! No se encuentra el origen de la referencia..**, p. **¡Error! Marcador no definido.**) Analicemos esta cuestión con los siguientes ejemplos:

- (20) La **prima de riesgo** española inició la jornada de hoy sin cambios, en 138 puntos básicos, pese a que el interés del **bono** nacional a diez años con el que se calcula volvió a caer hasta el 1,254 % desde el 1,257 % de ayer, que marcó su mínimo de cierre en más de catorce meses⁷⁸.
- (21) Según los datos de mercado recogidos por Efe, la rentabilidad del **bono** alemán a diez años, cuyo **diferencial** con la del español del mismo plazo marca el riesgo país, se elevó un poco, hasta el -0,111 % desde el -0,126 % previo⁷⁸.

Las unidades léxicas marcadas en negrita tienen uso en una situación de divulgación a pesar de que su uso frecuente o prototípico es el del discurso especializado. Teniendo en cuenta esto, ¿es posible hablar de que estas unidades codifican conocimiento especializado sistemático o, más bien, por el contrario, la activación del valor especializado está supeditada a las condiciones de recepción del usuario (del destinatario)? Cabe realizar estas cuestiones teniendo en mente que, aunque el medio de expresión es un medio de divulgación, el texto está orientado hacia un usuario destinatario con interés en o con conocimientos básicos sobre economía.

7. RELACIONES DENOMINACIÓN-CONCEPTO-SITUACIÓN DISCURSIVA

7.1. Fundamentación en el principio de variación

El principio de variación de la TCT permite explicar las relaciones lingüísticas que existen entre CC y CE, que, desde nuestro punto de vista, se inician con la necesidad de resolución de problemas de la sociedad. Estos problemas derivan en la reflexión de las disciplinas y la posterior generación de conocimiento especializado, que permite ser actualizado en diferentes aplicaciones intra o extradisciplinares. Cuando se consigue aproximarse a la resolución del problema inicial, se produce un trasvase de conocimiento de la realidad especializada de los campos y las disciplinas hacia la realidad cotidiana

⁷⁸ Fuente: EFE (30 de junio de 2016): «La prima de riesgo abre estable, en 138 puntos, y el bono sigue a la baja», *eldiario.es*. Disponible en: http://www.eldiario.es/economia/prima-riesgo-estable-puntos-sigue_0_532296830.html.

mediante la difusión y divulgación de la información a través de los medios de masas y medios educativos.

Esta relación no se presenta como un fenómeno nuevo, aunque sí mucho más nítido y perceptible a los ojos de los usuarios que participan en la comunicación, puesto que el acceso a la información y la eficiencia en su gestión es mucho mayor en la actualidad. Existen diferentes grados de relación entre el CC y el CE, dependientes de factores como la influencia del CE en la vida cotidiana de los usuarios o de si un fenómeno del CC es lo suficientemente descriptivo para ser asimilado por el CE. A favor de estos grados se encuentra Sager (1990: 60), que sostiene que «we have to accept [...] degrees of special reference and hence a different semantic force attached to a term according to the text in which a particular designation is encountered».

Por supuesto, el análisis lingüístico de los tipos de conocimiento y de sus interrelaciones es posible gracias a la actualización en el discurso (en concreto, en el nivel léxico), siendo posible estudiar la gran mayoría de condicionantes que afectan a denominación y concepto⁷⁹. Al respecto, Cabré (1999: 137) defiende la posibilidad de que los conceptos estén «sesgados por la época, cultura o características científicas del grupo del que los especialistas forman parte» y, además, se posiciona a favor del trasvase de conceptos entre diferentes campos del saber y entre el conocimiento común y el conocimiento especializado. Aunque es posible afirmar que la terminología se caracteriza por ser temáticamente específica, dicha especificidad «se actualiza permanentemente en el discurso, y por lo tanto no se basa en una adscripción previa a un dominio temático» (Cabré *et al.*, 2001: 308).

Las dos relaciones principales entre denominación, concepto y uso discursivo son la sinonimia y la polisemia, que se producen tanto en el seno de un mismo campo temático como entre campos temáticos diferentes y entre discurso común y discurso especializado:

Las denominaciones, por su parte, pueden coincidir totalmente (denominación y concepto) o parcialmente (sólo algunos aspectos del concepto) con unidades de otros campos. Si ello es así, los términos pueden presentar polisemia, en un doble sentido: a) una unidad puede ser reutilizada con el mismo significado en otro campo de conocimiento conservando los mismos rasgos conceptuales; b) de una sola unidad de base pueden desprenderse sentidos esencialmente coincidentes pero parcialmente específicos en función del campo en que se aplican. La variación

⁷⁹ Recordemos que, dado que las unidades del discurso especializado participan de las reglas de la gramática de la lengua y que, además, comportan variación de forma inherente, «los términos presentan elementos de variación formal que reflejan de forma natural sus condiciones de utilización en el discurso especializado, como las demás unidades de descripción de la lengua» (Cabré *et al.*, 2001: 308).

denominativa aceptada en la TCT como resultado de la observación de los datos en la realidad, se explica por la necesidad de adecuar la expresión a las características discursivas de cada tipo de situación comunicativa: ámbito, tema, perspectiva de abordaje del tema, tipo de texto, emisor, destinatario y situación (Cabré, 1999: 138).

A nuestro juicio, por tanto, las relaciones que es posible estudiar son las que siguen:

- a) La polisemia⁸⁰, en forma de terminologización y transterminologización, y la pseudopolisemia, en forma de banalización y variación conceptual.
- b) La sinonimia y cuasisinonimia, en forma de variación denominativa, y la pseudosinonimia, en forma de desteterminologización.

7.2. La terminologización

La terminologización constituye un tipo de polisemia que se produce cuando una denominación usada en el discurso común es adoptada por un campo temático del discurso especializado con un nuevo sentido, coincidente total o parcialmente con el sentido original. De esta forma, la unidad léxica original, usada en el discurso común, activa un nuevo sentido (un valor especializado) cuando es empleada en un discurso especializado en particular con unas condiciones pragmáticas concretas. El significado de la unidad léxica del discurso común, en principio, no se ve alterado ni modificado por la activación del nuevo valor en el discurso especializado. Picht (1984: 218-219) sintetiza de forma muy acertada el proceso de terminologización:

Terminologizar una palabra de la lengua general significa que la palabra adopta sobre su valor semántico ya existente un significado específico que corresponde a un concepto de un campo del saber. Así la palabra se convierte en un término con las características de éste: pertenencia⁸¹ a un campo del saber y ser parte del sistema de términos. Todo esto sucede sin que cambie el valor que la palabra tiene en la lengua general. El cambio semántico puede ser de distinto grado, en algunos casos apenas se percibe, en otros el significado se modifica casi totalmente. Sin embargo, en la superficie (= realización lingüística, el signo) no se observa alteración alguna.

La terminologización ha sido entendida como: a) un «trasvase importante de términos de la lengua común a los lenguajes especializados [...], que pasan a distintos ámbitos con significados precisos en cada caso, y por lo tanto, diferenciados entre sí» (Cabré, 1993: 168); b) «la asignación de un nuevo contenido conceptual a una forma léxica

⁸⁰ De forma relacionada, es posible incluir también la homonimia, cuyo funcionamiento se encuentra en las relaciones etimológicas.

⁸¹ Recordemos que según Cabré (1999: 121, 124) y Cabré, Feliu y Tebé (2001: 307), los términos no *pertenecen* a ningún campo del saber, sino que *son usados* en ellos.

conocida (a menudo perteneciente a la lengua común), muchas veces por medio de metáforas» (Arntz y Picht, 1995: 39); o c) movimiento de unidades de la lengua general a las especializaciones (López Rodríguez, 2001: 723). La terminologización es el proceso contrario a la desteminologización y guarda estrecha relación a nivel de frecuencia con la banalización.

En síntesis, una unidad léxica como *bueno* se terminologiza cuando en el campo temático del derecho se emplea con el sentido de ‘mediador en los actos de conciliación’ y es utilizada por un especialista en el campo. El mismo mecanismo puede defenderse cuando la unidad léxica *trabajo* se emplea como magnitud en mecánica o *ventana* como recuadro interactivo en informática. La especificidad de una ULE se halla en la especialización de rasgos semánticos de una unidad del discurso común (Meyer, Mackintosh y Varantola, 1997, 1998). La distinción, pues, de los usos especializados y no especializados en el proceso de terminologización está determinada por la especialización de dichos rasgos semánticos.

En el proceso de terminologización podrían observarse, por tanto, dos fases, en la que son analizables los parámetros de tipo de uso, significante o denominación, significado y activación de valor especializado. En los siguientes ejemplos, se pretende esquematizar el proceso de terminologización, en el cual las unidades léxicas (en estos casos, *duro* en lingüística y *débil* en ciencias de la computación) son empleadas en el discurso común con un sentido concreto, reconocible por el individuo medio como miembro de la sociedad, y, tras activar el valor especializado, condicionado por su uso pragmático, conviven junto con el valor no especializado. El símbolo «+» indica que sí se produce o tiene lugar el fenómeno y el símbolo «-» indica lo contrario.

- (22) La naturaleza fonética de la correlación “**duros**/blandos” se basa en dos rasgos: articulatoriamente, en la *palatalización/velarización*, según unos autores, o en la *palatalización/no palatalización*, según otros; acústicamente, sobre la distinción entre *sostenido/no sostenido* (Paños Roch, 1991: 6)⁸².

Duro	Antes del proceso		Después del proceso	
	Usado en discurso común	Usado en discurso esp. (LING)	Usado en discurso común	Usado en discurso esp. (LING)
Significante/ Denominación	+	-	+	+
Valor especializado	-	-	-	+

⁸² Fuente: PAÑOS ROCH, Natalia (1991): «Sobre el problema de la marca de correlación de los fonemas rusos palatalizados», *III Congreso de rusistas de España. Universitat de València. València, 16 y 17 de mayo de 1991*. Disponible en: http://www.uab.cat/Document/853/505/npanos_1.pdf.

Uso	Antes del proceso		Después del proceso	
	Usado en discurso común	Usado en discurso esp. (LING)	Usado en discurso común	Usado en discurso esp. (LING)
Significado	«Dicho de un cuerpo: Que se resiste a ser labrado, rayado, comprimido o desfigurado, que no se presta a recibir nueva forma o lo dificulta mucho» (DRAE, online).	-	«Dicho de un cuerpo: Que se resiste a ser labrado, rayado, comprimido o desfigurado, que no se presta a recibir nueva forma o lo dificulta mucho» (DRAE, online).	«Término [...] para aludir a las consonantes que no tienen palatalización, o que es muy escasa debido a que la lengua, debido a una intensión fuerte, apenas tiene contacto con el paladar duro». (Alcaraz Varó y Martínez Linares, 2004: 220)

Tabla 8: Terminologización de la unidad léxica duro en lingüística

- (23) Tipo de entidades: conjunto de entidades, que tienen los mismos atributos. Existen dos clases de tipo de entidad:

Débil: tipo de entidad que no tiene atributo llave propio.

Fuerte: tipo de entidad que posee un atributo llave propio (Malinowski, 1998: 145)⁸³.

Uso	Antes del proceso		Después del proceso	
	Usado en discurso común	Usado en discurso de esp. (INFOR)	Usado en discurso común	Usado en discurso de esp. (INFOR)
Significante/Denominación	+	-	+	+
Valor especializado	-	-	-	+
Significado	«De poco vigor o de poca fuerza o resistencia» (DRAE, online).	-	«De poco vigor o de poca fuerza o resistencia» (DRAE, online).	«Tipo de entidad que no tiene atributo llave propio» (Malinowski, 1998: 145).

Tabla 9: Terminologización de la unidad léxica débil en ciencias de la computación

7.3. La transterminologización

Constituye un tipo de polisemia interdominio (Meyer y Mackintosh, 2000) en la que se produce un «trasvase de unidades especializadas de un dominio a otro»⁸⁴ (Cabré, 1999: 100 y 121), normalmente, siguiendo una similitud metafórica (Meyer, Mackintosh y Varantola, 1997 y 1998). En este trasvase, las denominaciones pueden contener en cada caso «un significado parcialmente distinto en otros ámbitos de especialización» (Cabré,

⁸³ Fuente: MALINOWSKI, Elzbieta (1998): «El uso del modelo entidad-relación en el diseño de la representación de datos en las bases de datos orientadas a objetos», *Ingeniería: Revista de la Universidad de Costa Rica*, 8(1-2), págs. 141-153).

⁸⁴ Cabré se refiere en su publicación de 1999 (pp. 100 y 121) a la transterminologización como *pluriterminologización*.

1999: 140); «pueden aparecer en otros campos de especialización, con el mismo significado o con mínimas variaciones del mismo» (Cabré *et al.*, 2001: 308). Es lo que ocurre con unidades léxicas como a) *tronco* cuando se usa en el campo de la botánica y cuando se usa la misma unidad en anatomía; o b) *morfología*, referida al campo de la biología o la geología y, por otra parte, empleada en lingüística.

En la siguiente tabla, se esquematiza el proceso de transterminologización de la unidad *virus*, que, usada en el discurso especializado de la biología, con un sentido concreto y un valor especializado, se trasvasa posteriormente al campo de las ciencias de la computación, en el que su uso adquiere, de forma metafórica, un nuevo sentido (un valor especializado).

Virus	Antes del proceso		Después del proceso	
	Usado en discurso esp. (BIO)	Usado en discurso esp. (INFOR)	Usado en discurso esp. (BIO)	Usado en discurso esp. (INFOR)
Significante/ Denominación	+	-	+	+
Valor especializado	+	-	+	+
Significado	«Entities whose genome is an element of nucleic acid either DNA or RNA, which reproduces inside living cells and uses their synthetic machinery to direct the synthesis of specialised particles the virions, which contain the viral genome and transfer it to other cells» (Luria y Darnell, 1968) ⁸⁵ .	-	«Entities whose genome is an element of nucleic acid either DNA or RNA, which reproduces inside living cells and uses their synthetic machinery to direct the synthesis of specialised particles the virions, which contain the viral genome and transfer it to other cells» (Luria y Darnell, 1968) ⁸⁵ .	«Programa de ordenador que lleva a cabo acciones que resultan nocivas para el sistema informático» (Mur, Nieto y Molina, 1990: 15) ⁸⁶ .

Tabla 10: Transterminologización de la unidad léxica *virus* en biología y ciencias de la computación

7.4. La banalización

La banalización (Cabré, 1999: 100, 121 y 206; Adelstein, 1998) constituye un procedimiento de pseudopolisemia en el que una unidad léxica especializada comienza a aumentar su frecuencia de uso en el discurso común. La distinción, pues, de los usos especializados y no especializados en el proceso de banalización viene determinada por la atenuación o desaparición de la especificidad de los rasgos semánticos de la unidad léxica

⁸⁵ Fuente: LURIA, Salvador Edward y DARNELL, James E. (1968): *General virology*. New York [etc.]: Wiley.

⁸⁶ Fuente: MUR, Alfonso; NIETO, Pablo y MOLINA, Jesús (1990): *Virus informáticos*. Madrid: Anaya multimedia.

con respecto al uso discursivo de la misma. En este sentido, debe mencionarse que la banalización está ligada a distintos grados de desterminologización (*vid. infra* apartado II. 7.6., p. 120). La banalización es el proceso que tiene lugar cuando los medios de divulgación de masas y los medios educativos difunden el conocimiento especializado. *ADN, wifi, gripe A, dron* o *smartphone* son unidades banalizadas en mayor o menor medida. La banalización no es el proceso contrario a la terminologización a nivel semántico pero guarda relación con esta a nivel de frecuencia: mientras que la terminologización representa el fenómeno por el que una unidad léxica con uso en el discurso común aumenta su frecuencia en el discurso especializado asociada a un nuevo sentido o sentido metafórico (desde el punto de vista cognitivista), la banalización representa el fenómeno por el que una unidad léxica con uso en el discurso especializado aumenta su frecuencia en el discurso común asociada a un sentido similar, alterado a veces, fruto de una comprensión laxa del concepto por parte del hablante (es decir, asociada al fenómeno de desterminologización, en mayor o menor grado). La regulación de esta frecuencia (y consiguiente laxitud semántica) está motivada por eventos sociales e individuales.

Adelstein (1998: sin página) menciona que «ciertos términos presentan un grado tan alto de especificidad que tienen escasas posibilidades de banalizarse, por ejemplo, *cronolecto*». No obstante, diferimos en dicho argumento, puesto que, si pensamos en una unidad como *ADN*, es posible defender dicho «grado tan alto de especificidad» pero, como mencionábamos líneas más arriba, puede considerarse una unidad banalizada, dada su frecuencia en el discurso común. Por tanto, el criterio no se hallaría en la especificidad semántica de la unidad léxica, sino, más bien, en el grado de difusión y divulgación al que se ve sometido el conocimiento especializado que codifica la unidad (evidentemente, este CE aumenta su difusión a través de la recurrencia del término en el discurso común). En definitiva, entendemos la banalización en términos de frecuencia y la desterminologización en términos de laxitud semántica, si bien ambas suelen estar relacionadas.

En efecto, Adelstein menciona que algunos términos, pese a ser considerados «muy especializados, son incorporados al léxico común por razones culturales y de circulación de conocimiento, pero con restricciones de uso en cuanto al universo discursivo» (Adelstein, 1998: sin pág.), aunque también reconoce la posibilidad de «penetrar al léxico común sin restricciones» (*ídem*), como con el caso de *lente de contacto*.

Para Adelstein (*ídem*), la banalización es un proceso que conlleva creación neológica (en tanto que el discurso común incorpora el uso de nuevas formas denominativas y nuevos conceptos). No obstante, tan solo podría considerarse unidad neológica en el

discurso común, puesto que en el discurso especializado dependería de los criterios de neologicidad y aceptabilidad neológica (Cabré, 1993: 445; y Guerrero Ramos, 1995: 14-15).

Aunque la división entre DC y DE es un constructo teórico que pretende permitir reflexionar sobre el comportamiento de las unidades lingüísticas, lo cierto es que en el caso de la banalización, junto con otros casos, es complejo delimitar si un término se considera neologismo o no cuando este pasa a formar parte del DC. Entendiendo la lengua de forma holística, como un todo intercondicionado e interconectado, si una unidad léxica cumple con los criterios de neologicidad y de aceptabilidad neológica, será considerada una unidad neológica en todo el sistema. Sin embargo, si profundizamos en los tipos de discurso, en el par DC-DE, las diferencias no son tan claras, sobre todo cuando la unidad neológica en el DE ha sido fruto de una terminologización. Piénsese en el siguiente ejemplo:

Discurso común		Discurso especializado (INFOR)		Discurso común
ratón	→	ratón	→	ratón
	<i>Terminologización</i>		<i>Banalización</i>	
<u>Sentido original:</u> «Mamífero roedor de pequeño tamaño, de hocico puntiagudo y cola larga, de pelaje generalmente gris, muy fecundo y que habita en las casas» (DRAE, 2014, <i>online</i>).		<u>Nuevo sentido:</u> «Pequeño aparato manual conectado a una computadora, cuya función es mover el cursor en la pantalla para dar órdenes» (DRAE, 2014, <i>online</i>).		<u>Sentido original + nuevo sentido</u>

Tabla 11: Proceso de terminologización y banalización de la unidad *ratón*

Finalizado el proceso de terminologización, sería posible hablar de neonomia, en el sentido de que aparece una nueva denominación en el discurso especializado de la informática que proviene del uso discursivo común y adquiere un nuevo sentido (aunque un sentido metafórico basado en el significado de la unidad *ratón* en el DC). Cuando este neónimo acaba por banalizarse, es decir, aumenta su frecuencia de uso en el DC, solo podría hablarse de neología semántica (i. e. la unidad *ratón* incorpora un nuevo sentido, fruto de la activación del valor especializado en el DE). A nuestro juicio, el grado de desterminologización de *ratón* y de *ADN*, fruto de la consiguiente banalización, no es similar desde el punto de vista social, quizá debido al grado de familiaridad con cada objeto que suponen en la vida cotidiana de los hablantes. Dependiendo de la situación

comunicativa y del campo temático en el que se inscriba un discurso, la unidad léxica *ratón* activaría un sentido u otro.

El proceso de banalización nos lleva a reflexionar sobre si, como los neologismos en la diacronía, los términos pierden su carácter de término (es decir, su valor especializado) o si acaban por desteterminologizarse (su uso semántico se vuelve impropio con respecto al sistema conceptual o, en otras palabras, su uso no respeta en su totalidad el sentido especializado original). Según Adelstein (2007: 275), la relación entre los significados especializados y los no especializados de una misma unidad léxica guardan una relación de polisemia, dado que coinciden (en diferentes grados) en la activación de determinados elementos semánticos, postura que argumenta con el caso de la unidad *madre*.

Si tenemos en cuenta unidades como *tumor* o *cáncer*, que pueden considerarse unidades banalizadas, es posible percibir en ciertos contextos que se tratan como unidades sinónimas (aunque no lo son) o con cierta laxitud significativa cuando son empleadas por el individuo medio. Otros ejemplos pueden ser *tablet* o *wifi* en el campo de la informática; *virus* o *zika* en biología; o *prima de riesgo* o *déficit* en economía. Si bien, podemos estar de acuerdo en que estas unidades son unidades banalizadas, puede apreciarse un grado de banalización entre ellas que viene determinado, de alguna forma, «por el tipo de vinculación que en un momento dado se establece entre el dominio de origen y la vida cotidiana» (Adelstein, 1998). Por ello, proponemos la distinción de dos tipos de banalización, según el grado de frecuencia en el uso en el tipo de discurso (especializado o común): a) banalización efímera o aguda; y b) banalización perenne o crónica; y de dos criterios para evaluar los grados de banalización:

- a) Criterios de carácter **subjetivo** (sociocognitivos-discursivos principalmente), basados en la familiaridad que percibe el usuario de la unidad banalizada, que pueden recibir el nombre de *frecuencia subjetiva* o *naturalidad subjetiva*.
- b) Criterios de carácter **objetivo**, basados en la estadística a través de lingüística de corpus, que pueden recibir el nombre de *frecuencia objetiva* o *naturalidad objetiva*.

7.4.1. La banalización efímera

Este tipo de banalización se produce en un momento de la diacronía (normalmente en la sincronía) y, por razones socioculturales, la unidad léxica no acaba incorporándose al caudal léxico del discurso común, sino que su paso por el mismo es *efímero*. En otras palabras, se trata de aquellas unidades léxicas especializadas que, por causas que influyen en la vida cotidiana de los usuarios durante un periodo de tiempo, pasan a aumentar su frecuencia de uso en el discurso común pero que, una vez finalizado el evento o acontecimiento que inició el aumento de la recurrencia, acaban por limitar o finalizar su uso en el discurso común. Los siguientes ejemplos (*gripe A* y *prima de riesgo*) pretenden ilustrar la banalización efímera.

El caso de *gripe A* se remonta a la pandemia de H1N1 del año 2009. Su repercusión sociocultural a través de los medios de divulgación es comprobable en el corpus CORPES XXI, cuyo pico de frecuencia se encuentra en el período comprendido entre 2009 y 2010, sobre todo en España, en noticias y blogs. A partir de 2011, la unidad *gripe A* disminuye su frecuencia de uso en el discurso común, puesto que dicha pandemia se da por finalizada y, por tanto, deja de tener relevancia en los medios de divulgación.

Gripe A	Antes del proceso		Después del proceso	
Uso	Usado en discurso esp. (BIO)	Usado en discurso común	Usado en discurso esp. (BIO)	Usado en discurso común
Significante/ Denominación	+	-	+	+
Valor especializado	+	-	+	¿-?
Significado	Infección vírica causada por una variante del <i>Influenzavirus A</i> (H1N1) que causó una pandemia en 2009.	-	Infección vírica causada por una variante del <i>Influenzavirus A</i> (H1N1) que causó una pandemia en 2009.	

Tabla 12: Banalización de la unidad léxica *gripe A* en el discurso común

En la siguiente gráfica se comprueba el carácter efímero de la banalización de *gripe A*. Los datos han sido extraídos de CORPES XXI⁸⁷:

⁸⁷ A nuestro juicio, uno de los recursos más útiles para medir la banalización efímera son las publicaciones de la prensa o los blogs sobre actualidad, puesto que registran casi de manera inmediata el uso del léxico en el discurso.

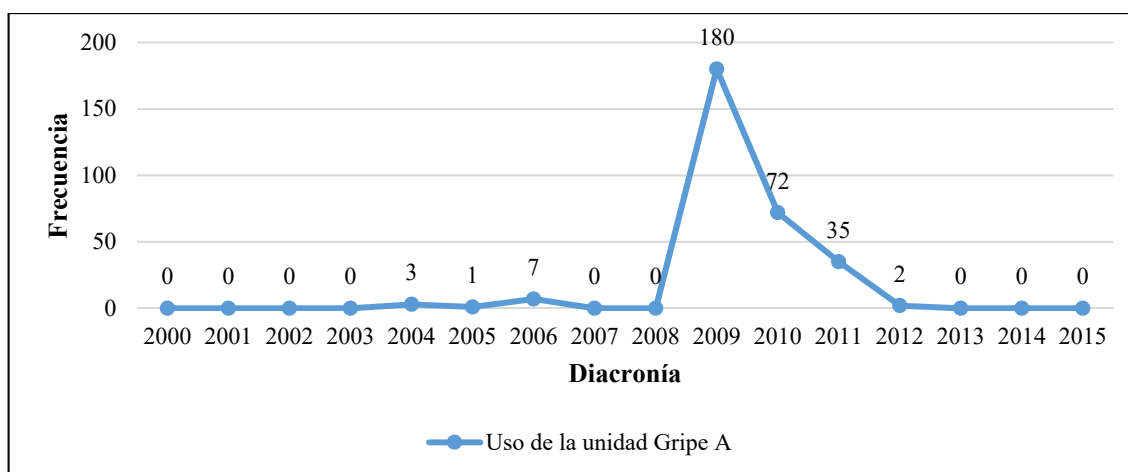


Gráfico 1: Frecuencia absoluta de gripe A entre 2000 y 2015 como banalización efímera según CORPES XXI

El caso de *prima de riesgo* es más cercano temporalmente y aún puede comprobarse en los medios de divulgación. Si usamos como fuente el corpus CORPES XXI, esta unidad léxica, aunque tiene registros desde 2001, tiene especial relevancia en el período 2011-2015, en España y en aquellos textos relacionados con la política, la economía y la justicia, sobre todo en forma de noticias y blogs. Tiene sentido hallar dichos datos, puesto que es a partir del crecimiento de los efectos de la crisis económica en España cuando la unidad *prima de riesgo* comienza a *esparcirse* por todos los medios de divulgación, si bien su significado especializado es percibido por el usuario medio con la laxitud significativa a la que hacíamos referencia anteriormente. Posiblemente, al finalizar el período de crisis económica, la unidad *prima de riesgo* acabe desapareciendo de los medios de difusión y finalice su etapa de banalización efímera.

Prima de riesgo	Antes del proceso		Después del proceso	
	Usado en discurso esp. (ECON)	Usado en discurso común	Usado en discurso esp. (ECON)	Usado en discurso común
Significante/Denominación	+	-	+	+
Valor especializado	+	-	+	¿-?
Significado	«Diferencia esperada a largo plazo entre la rentabilidad de la inversión en bolsa y el tipo de interés sin riesgo» (Gómez Montejo, 2010: 38) ⁸⁸ .	-	«Diferencia esperada a largo plazo entre la rentabilidad de la inversión en bolsa y el tipo de interés sin riesgo» (Gómez Montejo, 2010: 38) ⁸⁸ .	

Tabla 13: Banalización de la unidad léxica *prima de riesgo* en el discurso común

⁸⁸ Fuente: GÓMEZ MONTEJO, Ignacio (2010): «La prima de riesgo: ni tanta ni tan baja», *Bolsa: revista mensual de bolsas y mercados españoles*, 185, pp. 38-46. Disponible en: <https://dialnet.unirioja.es/servlet/articulo?codigo=3301888&orden=269037&info=link>.

7.4.2. La banalización perenne o crónica

La banalización perenne o crónica se produce cuando una unidad léxica del discurso especializado acaba incorporándose al caudal léxico del discurso común, tras reconocerse un evento u objeto de gran relevancia para la vida del usuario. Dicha unidad aparece en un momento de la diacronía y su uso se mantiene fijo con el paso del tiempo, hasta que el usuario entiende dicha unidad como una unidad más de su lexicón, con características semánticas especiales (como, por ejemplo, que su funcionamiento conceptual se escapa a su entendimiento) pero reconocible como una unidad cotidiana. Los siguientes ejemplos (*ADN* y *wifi*) pretenden ilustrar la banalización perenne.

Desde las publicaciones en torno al modelo de la doble hélice del ADN a mediados del siglo XX, pero sobre todo a partir del siglo XXI, debido a los avances científicos en genética, la unidad *ADN* ha aumentado su frecuencia en todos los medios de comunicación, llegando a incorporarse a los temarios de la educación secundaria obligatoria. Si bien su sistema conceptual a veces es desconocido por el usuario medio, este es capaz de determinar, en mayor o menor medida, a qué se refiere *ADN*. Debido a su relevancia científica y repercusión social, *ADN* ha acabado por incorporarse al caudal léxico del discurso común (prueba de ello es que haya sido incluido en el *Diccionario de la lengua española* de la RAE).

ADN	Antes del proceso		Después del proceso	
	Usado en discurso esp. (BIO)	Usado en discurso común	Usado en discurso esp. (BIO)	Usado en discurso común
Significante/ Denominación	+	¿+?	+	+
Valor especializado	+	-	+	¿-?
Significado	«Polímero de desoxirribonucleótidos de elevada masa molecular, constituido por bases púricas (adenina o guanina) y pirimidínicas (citosina o timina) unidas a moléculas de desoxirribosa, las cuales se relacionan entre sí por enlaces fosfato. Constituye el fundamento molecular de la herencia, con una estructura que se expresa en forma circular en las mitocondrias, como una sola cadena o, más frecuentemente, en los cromosomas del núcleo celular, como doble cadena antiparalela en doble hélice, en la que las bases púricas y pirimidínicas, portadoras de la información genética, están	-	«Polímero de desoxirribonucleótidos de elevada masa molecular, constituido por bases púricas (adenina o guanina) y pirimidínicas (citosina o timina) unidas a moléculas de desoxirribosa, las cuales se relacionan entre sí por enlaces fosfato. Constituye el fundamento molecular de la herencia, con una estructura que se expresa en forma circular en las mitocondrias, como una sola cadena o, más frecuentemente, en los cromosomas del núcleo celular, como doble cadena antiparalela en doble hélice, en la que las bases púricas y pirimidínicas, portadoras de la información genética, están unidas por enlaces de hidrógeno» (DTM, 2012) ⁸⁹ .	

ADN	Antes del proceso		Después del proceso	
Uso	Usado en discurso esp. (BIO)	Usado en discurso común	Usado en discurso esp. (BIO)	Usado en discurso común
	unidas por enlaces de hidrógeno» (DTM, 2012) ⁸⁹ .			

Tabla 14: Banalización de la unidad léxica ADN en el discurso común

En la siguiente gráfica se comprueba el carácter perenne de la banalización de ADN. Los datos han sido extraídos de CORPES XXI⁹⁰:

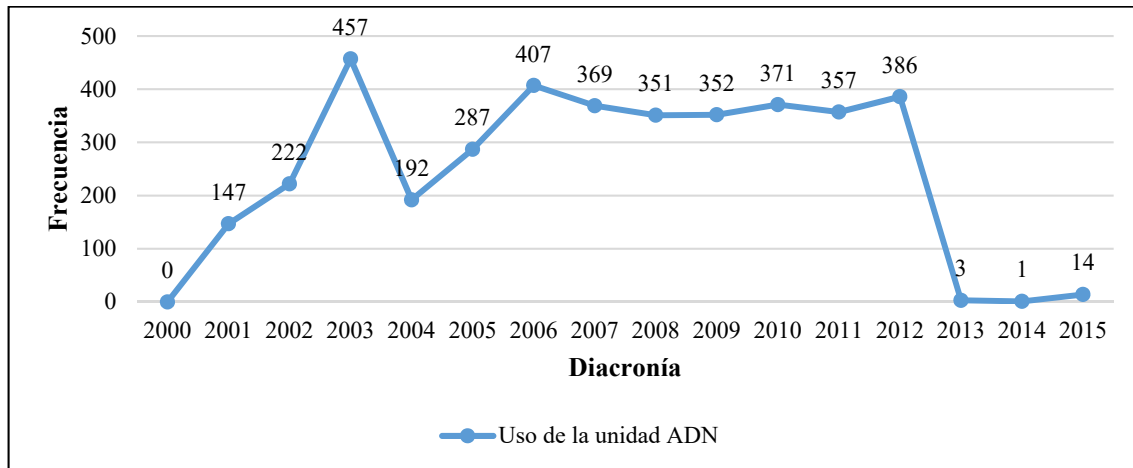


Gráfico 2: Frecuencia absoluta de ADN entre 2000 y 2015 como banalización perenne según CORPES XXI

Con *wifi* ocurre de igual forma, aunque su incorporación en el discurso común ha sido más reciente y nos atreveríamos a decir que incluso más rápida (también se incluye ya en el *DRAE*). El proceso de incorporación al discurso común es idéntico al de *ADN*, aunque sus causas son diferentes.

Wifi	Antes del proceso		Después del proceso	
Uso	Usado en discurso esp. (INFOR)	Usado en discurso común	Usado en discurso esp. (INFOR)	Usado en discurso común
Significante/Denominación	+	¿+?	+	+
Valor especializado	+	-	+	¿-?
Significado	«Conjunto de estándares para redes inalámbricas basado en las especificaciones IEEE 802.11 (especialmente la 802.11b), creado para redes locales inalámbricas, pero que	-	«Conjunto de estándares para redes inalámbricas basado en las especificaciones IEEE 802.11 (especialmente la 802.11b), creado para redes locales inalámbricas, pero que también se utiliza para acceso a internet» (Alegsa, 2010) ⁹¹ .	

⁸⁹ Fuente: DTM – Real Academia Nacional de Medicina (2012): *Diccionario de términos médicos*. Madrid: Editorial Médica Panamericana.

⁹⁰ Cabe mencionar que la caída de frecuencia en los años 2013, 2014 y 2015 puede deberse a que CORPES XXI todavía no contiene suficientes textos de los mencionados años.

Wifi	Antes del proceso		Después del proceso	
Uso	Usado en discurso esp. (INFOR)	Usado en discurso común	Usado en discurso esp. (INFOR)	Usado en discurso común
	también se utiliza para acceso a internet» (Alegsa, 2010) ⁹¹ .			

Tabla 15: Banalización de la unidad léxica wifi en el discurso común

La distribución de años y frecuencias absolutas a partir de los datos de CORPES XXI se relaciona en el Gráfico 3:

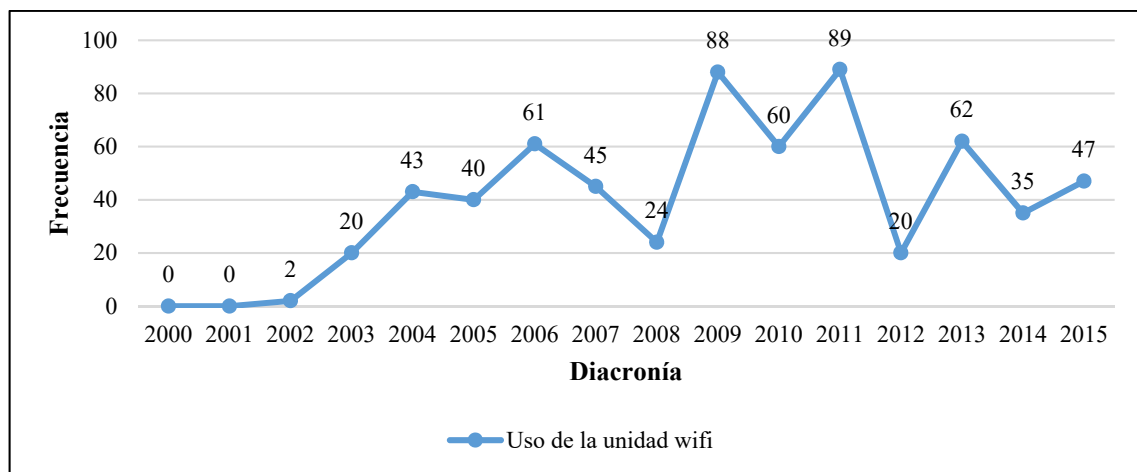


Gráfico 3: Frecuencia absoluta de wifi entre 2000 y 2015 como banalización perenne según CORPES XXI

Al respecto de los criterios para determinar los grados de banalización, sería útil investigar en futuros estudios sobre la asimilación de dichos criterios a los criterios de caracterización neológica (OBNEO, 2004; Vega Moreno, 2015), así como el papel de los procedimientos de reformulación en la banalización y su relación con la disminución del coste de procesamiento de los términos.

7.5. La variación denominativa y conceptual

La variación denominativa constituye un tipo de relación sinonímica que se produce entre dos o más unidades léxicas especializadas, que comparten un concepto común. En palabras de Freixa (2002b: 54): «el fenomen pel qual a una mateixa noció li corresponen diverses denominacions». La principal diferencia entre estas unidades se halla en el punto de vista o perspectiva que adopta la denominación. Cuando las similitudes se hallan en la denominación y las diferencias en el concepto hablamos de *variación conceptual*: «variació

⁹¹ Fuente: ALEGSA, Leandro (2010): «Definición de Wi-Fi», *Diccionario de informática y tecnología*. Santa Fe. Disponible en: <http://www.alegsa.com.ar/Dic/wi-fi.php>.

conceptual significa per a nosaltres variació en un mateix concepte: creiem que en el procés de denominació, un mateix concepte pot ser abordat de maneres diferents i que per aquesta raó poden aparèixer denominacions diferents» (Freixa, 2002b: 55).

La variación denominativa rompe con el principio de biunivocidad (relación 1:1 entre denominación y concepto) de la terminología wüsteriana, es decir, se aleja de la monosemia, considerada «la situación óptima de denominación» (Picht, 1984: 213).

Por ejemplo, cuando hacemos referencia a la característica principal del término, podemos mencionar su *valor especializado* o *valor terminológico*, su *uso especializado* o *uso terminológico* y su *carácter especializado* o *carácter terminológico*, pero siempre hacemos referencia al mismo concepto. Según Picht (1984: 215), la variación denominativa surge por creaciones *ad hoc* fruto de la falta de conocimiento de la terminología del campo del saber por parte de los propios profesionales o de los traductores.

A pesar de que estas razones pueden ser defendidas sin problema, consideramos que la propuesta de Freixa (2002a, 2002b, 2005, 2013) profundiza en mayor medida en la complejidad de la variación denominativa, que se articularía en torno a causas previas, dialectales, funcionales, discursivas, interlingüísticas y cognitivas.

De forma complementaria, es posible afirmar que la variación conceptual es fruto de polisemia parcial interdominio y la variación denominativa se produce a través de sinonimia intradominio⁹².

7.5.1. La variación conceptual a través de la polisemia parcial interdominio

Cuando un objeto de la realidad es analizado según diferentes disciplinas (es decir, cuando se trata de un análisis multidimensional), se generan diferentes conceptos, estrechamente relacionados, de un mismo objeto (solo una parte del concepto es común a todas las disciplinas; se diferenciarían, pues, en rasgos pertinentes). Freixa afirma al respecto:

Fixem-nos que Cabré no està tractant el fenomen pel qual es pot abordar de maneres diferents un únic concepte a l'interior d'una mateixa especialització, sinó que es refereix a percepcions diferents d'una mateixa realitat que van comportant diferents processos de conceptualització i, per tant, han conduït a diferents conceptes,

⁹² Noción basada en la polisemia intradominio de Meyer y Mackintosh (2000).

fenomen que es pot esdevenir entre diferents especialitats o entre una conceptualització especialitzada i una de no especialitzada (Freixa, 2002b: 55).

Por ejemplo, al pensar en la unidad *trabajo*, podemos referirnos al producto de su terminologización en diferentes campos o dominios como la economía, la sociología o la mecánica, en los que existe un núcleo semántico común, pero diferencias notables por cada disciplina.

Trabajo	
Acción y efecto de ocuparse de cualquier actividad física o intelectual.	DC
Medida del esfuerzo realizado por los seres humanos.	Economía
La ejecución de tareas que implican un esfuerzo físico o mental y que tienen como objetivo la producción de bienes y servicios para atender las necesidades humanas. El trabajo es por tanto la actividad a través de la cual el hombre obtiene sus medios de subsistencia por lo que tiene que trabajar para vivir o vive del trabajo de los demás (Sanchís, 2011: 17) ⁹³ .	DE
Magnitud física escalar producto de la fuerza ejercida sobre un cuerpo por su desplazamiento.	Mecánica

Tabla 16: Significados de trabajo en diferentes discursos especializados

7.5.2. La variación denominativa a través de la sinonimia intradominio

Se produce cuando un objeto de la realidad es analizado en el seno de un campo temático y la reflexión a lo largo de la diacronía y a través de diferentes escuelas provoca que surjan nuevas denominaciones para hacer referencia a un concepto. Es lo que ocurre con los casos de *término*, *unidad léxica especializada* o *unidad terminológica*, que son empleados dentro de la lingüística (en concreto, dentro de la terminología) con sentidos coincidentes, aunque cada unidad es fruto de una reflexión diferente.

Como avanzamos líneas antes, las causas de la variación denominativa son diversas, si bien pueden distinguirse dos niveles de causas. En el nivel superior se hallarían las causas previas y cognitivas (y constituyen las premisas de partida) y en el segundo nivel las causas dialectales, funcionales, discursivas e interlingüísticas. A continuación, adaptamos la clasificación de Freixa (2002b; en Freixa, 2013: 39) sobre los tipos y subtipos de causas variación, las cuales acompañamos de ejemplos.

TIPOS	SUBTIPOS	OBSERVACIONES	EJEMPLOS
NIVEL SUPERIOR			
Causas previas	—La redundancia lingüística	Las lenguas permiten intrínsecamente variar	

⁹³ Fuente: SANCHÍS, Enric (2011): *Trabajo y paro en la sociedad postindustrial*. Valencia: Tirant Lo Blanch, p. 17.

TIPOS	SUBTIPOS	OBSERVACIONES	EJEMPLOS
	—La arbitrariedad del signo lingüístico —Las posibilidades de variación de la lengua	la manera de hacer referencia a la realidad y de expresarla (dispone de recursos morfológicos diferentes).	
Causas cognitivas	—Imprecisión conceptual —Distanciación ideológica —Diferencias en la conceptualización		
SEGUNDO NIVEL			
Causas dialectales	—Variación geográfica —Variación cronológica —Variación social	Es necesariamente externa a la lengua	<i>underground - subway</i>
Causas funcionales	—Adecuación al nivel de lengua —Adecuación al nivel de especialización	Es necesariamente interna a la lengua	<i>cáncer de cérvix uterino - cáncer de útero</i>
Causas discursivas	—Evitar la repetición —Economía lingüística —Creatividad, énfasis y expresividad	Hacen referencia a causas estilísticas	<i>síndrome de la inmunodeficiencia adquirida - sida</i>
Causas interlingüísticas	—Convivencia del término “local” con el préstamo —Diversidad de propuestas alternativas	Las variantes concurren por razones sociolingüísticas	<i>ratón - mouse</i>

Tabla 17: Tipos y subtipos de causas de variación denominativa según Freixa (2013)

7.6. La desteterminologización

El proceso de desteterminologización hace referencia al uso impropio o disolución conceptual de un término, ya se emplee en el discurso común o discurso especializado. Meyer y Mackintosh (2000: 113) explican la desteterminologización de la siguiente forma:

When a term migrates into general language, it may undergo two types of semantic changes. On the one hand, the essence of the terminological sense (i.e. its domain sense) may be retained. On the other hand, semantic changes can be more substantial, involving a significant dilution of the terminological sense.

Meyer y Mackintosh (ídem), desde la perspectiva de la divulgación médica, consideran que cuando «el receptor del texto no percibe su significado con la misma exactitud y profundidad que el experto», se produce la distorsión del conocimiento científico en el traslado de este al gran público (Campos, 2013: 49). Esta también es la

perspectiva de Arntz y Picht (1995: 40), que afirman que, a medida que un texto aumenta su grado de especialización, disminuye su comprensibilidad general, lo que también disminuye el número de usuarios que comprenden dicho texto. Por el contrario, cuando los términos se incorporan al caudal léxico del discurso común «el oyente comprende con imprecisión o erróneamente muchos de los tecnicismos difundidos por los medios de comunicación y no puede emplearlos correctamente. Puede hablarse del fenómeno de *desterminologización*».

Por su parte, Montalt y Shuttleworth (2012: 16), también partiendo del ámbito sanitario, apuestan por un enfoque dialógico entre médicos y pacientes (enfoque que puede trasladarse a cualquier relación discursiva) y redefinen el concepto de *desterminologización*:

It is a process of recontextualisation and reformulation of specialised terms aiming at making the concepts they designate relevant to and understandable by a lay audience. This process is motivated by specific cognitive, social and communicative needs, and takes place as part of a broader process of recontextualisation and reformulation of discourse.

Con todo, diferimos de esta definición, ya que lo que aquí se propone es cambiar el sentido de *desterminologización* para referirse, precisamente, al mecanismo de *recontextualización discursiva*, un concepto diferente, en el que tienen gran protagonismo los procedimientos de reformulación formal y conceptual (Rodríguez-Tapia, 2015). Por ello, nos posicionamos a favor del sentido que proponen Meyer y Mackintosh (2000: 113) y Arntz y Picht (1995: 40).

Como defendíamos en trabajos anteriores (Rodríguez-Tapia, 2015: 9), consideramos que la *desterminologización* puede constituir:

- a) Un **procedimiento consciente**, que tendría lugar entre especialistas cuando utilizan las formas denominativas indistintamente sin atender a la precisión conceptual. Sería el caso del empleo de *unidad de significación especializada*, *unidad de conocimiento especializado* o *término* de forma indiferente sin distinguir matices conceptuales. Sería, pues un procedimiento de cuasisinonimia o sinonimia parcial, que tiene lugar cuando el grado de semejanza conceptual es tal que, hasta los especialistas y profesionales, aun conociendo dicha diferencia, los usan como sinónimos (Picht, 1984: 216).
- b) Un **procedimiento inconsciente**, que sucede cuando el lego emplea las formas denominativas indistintamente por desconocimiento de la precisión conceptual.

Un ejemplo sería la indistinción entre *acento* y *tilde*, *traducción* e *interpretación* o *tumor* y *cáncer* del que se hace uso en la lengua común. Se trataría, pues, de un caso de pseudosinonimia, en el que se produce un uso equívoco de términos que el usuario cree que son sinónimos (Picht, 1984: 216). Es el caso del uso de *tumor* y *cáncer* como sinónimos (véase apartado dedicado a la banalización), lo que, desde nuestro punto de vista, permite conectar dichas nociones: la desterminologización va ligada a la banalización, puesto que el aumento de frecuencia del uso producido en el segundo posibilita la disolución del significado especializado de la unidad léxica en cuestión.

7.7. Esquema de relaciones

En síntesis, la Ilustración 8 pretende sintetizar los diferentes procesos analizados en este apartado usando como marco el discurso especializado (y sus diferentes tipos) y el discurso común, así como la variación denominativa y conceptual:

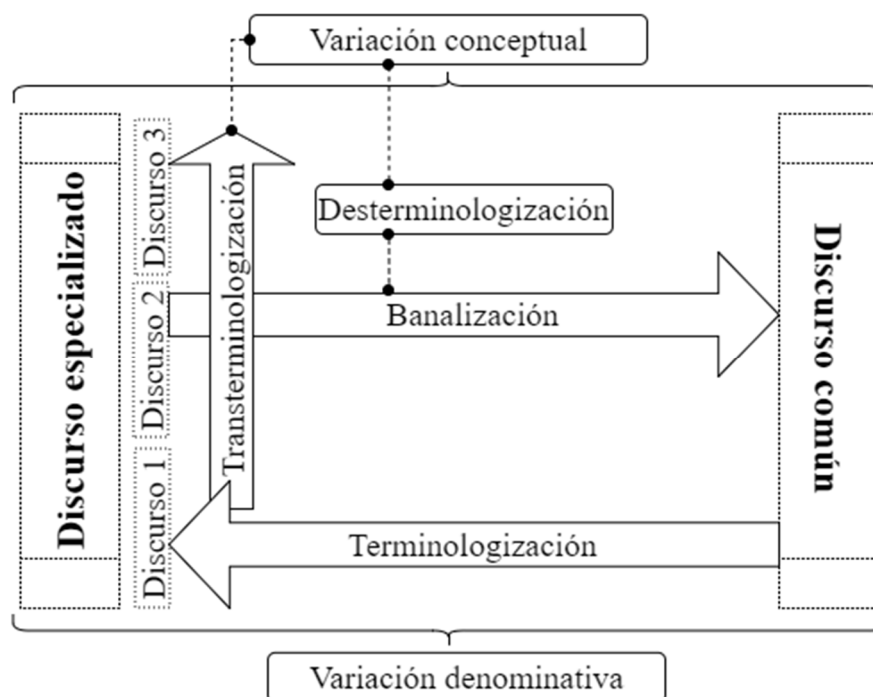


Ilustración 8: Esquema de relaciones concepto-denominación-discurso [elaboración propia]

- a) La **terminologización** se produce cuando una unidad léxica del discurso común pasa a emplearse en el discurso especializado con un nuevo sentido (un valor especializado), como en el caso de *duro* en lingüística y de *débil* en ciencias de la computación.

- b) La **transterminologización** se produce cuando la unidad léxica se trasvasa entre discursos disciplinares, como en el caso de *virus* entre biología y ciencias de la computación. La transterminologización va ligada a la variación conceptual.
- c) La **banalización**, de forma contraria a la terminologización, se produce cuando una unidad léxica empleada en un discurso especializado pasa a ser empleada en el discurso común. Es el caso de *ratón*, *ADN* o *gripe A*. Cabe distinguir entre banalización efímera y banalización perenne o crónica. Suele estar vinculada con grados de desteterminologización.
- d) La **variación denominativa** es el proceso por el que coexisten en el discurso en general dos o más denominaciones para un mismo concepto, como en el caso de *término*, *unidad léxica especializada* o *unidad terminológica*.
- e) La **variación conceptual** es el proceso que tiene lugar a causa de la multiperspectiva desde la que se observan los objetos de la realidad, como en el caso de la unidad *trabajo*. En la variación conceptual sientan sus raíces la transterminologización o la desteterminologización.
- f) La **desteterminologización**, de forma consciente o inconsciente, se produce cuando la unidad léxica se usa o comprende con imprecisión conceptual, como en el caso de *tilde* y *acento* o *tumor* y *cáncer*. Suele ir ligada a la banalización.

7.8. Los términos crípticos y délficos

Cuando nos aproximamos a la relación entre concepto, denominación y discurso adoptando una posición basada en el léxico, es posible apreciar que existen términos marcados de diferente forma:

[S]e observa una selección del léxico dependiente de dominio, de léxico ya sea formal y semánticamente o solo semánticamente. Formal y semánticamente, porque existen vocablos que solo aparecen en el discurso de una determinada especialización. Solo semánticamente porque, aunque se trate de vocablos utilizados también en otras materias o en la comunicación general, estos adquieren un sentido específico dentro de cada dominio especializado, sentido que se ubica en el conjunto de sentidos que conforman el significado global de una unidad léxica (Cabré y Estopà, 2005: 75).

Así, podemos distinguir, de acuerdo con la terminología de Riggs (1988), seguida también por Campos Plaza y Campos Martín (2009: 686 y 687), dos tipos de unidades léxicas especializadas⁹⁴:

- a) Los **términos crípticos** (Riggs, *ibíd.*: 117): se trata de unidades léxicas especializadas marcadas semántica y formalmente. No son transparentes semánticamente para el usuario medio y pueden ser identificadas rápidamente como términos, dada su peculiaridad morfológica. Algunos ejemplos pueden ser *ácido desoxirribonucleico*, *alolexo*, *heterocomposición*, *hipotaxis* u *osteomalasia*⁹⁵.
- b) Los **términos délficos** (Riggs, *ibíd.*: 28, 109, 116, 154): se trata de unidades léxicas especializadas marcadas semánticamente (es decir, que activan un valor especializado) pero no marcadas formalmente (el lego podría tener problemas para identificarla como término). Su denominación deriva del oráculo de Delfos y hace referencia a *lo ambiguo*. Los términos délficos pueden ser producto de una terminologización. Algunos ejemplos pueden ser:
 - i) en ciencias de la computación: *fuerte* y *débil* referidos a una variable, así como *invocar*, *converger* o *ejecutar*;
 - ii) en lingüística, en concreto en gramática, *espíritu* o *duro*;
 - iii) en traducción, los términos *localización*, *problema* o *dificultad*;
 - v) en el ámbito jurídico, la unidad *fallo*;
 - vi) en mecánica, las unidades *ciega* y *mariposa* referidas a una tuerca;
 - vii) en documentación, los términos *ruido* y *silencio*;
 - viii) en organización civil, *parroquia*; y
 - xix) en finanzas, la unidad *apalancamiento*.

Los términos délficos suponen un problema para la labor terminológica, ya sea del propio terminólogo sistemático o para el mediador interlingüístico. A veces, debido al desconocimiento del campo temático (como ya avanzaba Picht, 1984: 215), algunos términos no son identificados como tales, corriendo el riesgo de ser tratados de forma diferente. Montero y Faber (2008: 190) explican al respecto las razones que pueden llevar a esta confusión terminológica:

- a) por desconocimiento del significado o valor pragmático de las unidades terminológicas y fraseológicas usadas en el texto porque no están familiarizados

⁹⁴ Tanto los términos crípticos como los délficos pueden ser objeto de banalización terminológica.

⁹⁵ En nuestros ejemplos, es común la construcción morfológica mediante formantes grecolatinos.

con el ámbito de uso (temporal, socio-profesional y geográfico), el grado de especialización del texto, el nivel de normalización, la frecuencia de uso, etc.;

b) por la duda sobre el valor terminológico de una unidad monoléxica de uso frecuente en el discurso general, cuestión que tendrán que resolver para poder tratarla de forma adecuada (si es un término tendrá que encontrar una unidad terminológica en la lengua de llegada y si no lo es, podrá utilizar otros recursos);

c) por la duda sobre el valor terminológico de las unidades poliléxicas que no representen una estructura frecuente o prototípica o sean muy extensas.

8. EL VALOR ESPECIALIZADO (II): CONCLUSIÓN

En síntesis, es posible defender que la categoría de término está ligada a la condición de activación de *valor especializado*, es decir, cuando la unidad léxica codifica conocimiento especializado. La activación de este valor y la codificación de dicho conocimiento están condicionadas por la situación discursiva, es decir, los interlocutores, el tipo de texto y la función que se pretenda cumplir con dicho texto.

Este condicionamiento resulta de la imposición de condiciones por parte de dichos parámetros. Por ejemplo, la relación discursiva entre emisor y destinatario (y, a nuestro juicio, sobre todo el perfil del destinatario) afectan a la función textual y esto conlleva la selección de ciertas estrategias recontextualizadoras que permitan la comprensibilidad del texto, que, por su parte, también impone reglas de construcción, fruto del tipo textual. Para defender la existencia del término, pues, deben evaluarse dichos condicionantes.

8.1. Criterios para determinar el valor especializado

A pesar de que la terminología empleada en algunas disciplinas puede estar marcada formalmente (como puede observarse en medicina —*cardiopatía*— o en lingüística —*alolexo*—), pueden diseñarse estrategias que permitan la búsqueda automática de dicha terminología. En cuanto a los términos délficos, el problema existe en su denominación, que, a no ser que haya sido incluida previamente en una base de datos, no permite la extracción automática de la terminología.

Evidentemente, el criterio semántico es criterio *sine qua non* para catalogar una unidad léxica como término. Todas las unidades léxicas especializadas están marcadas semánticamente (activan el valor especializado según la situación). Pero esta marca no tiene por qué ser formalizable: *hombre bueno* (derecho), *duro* (lingüística), *variable* (ciencias de la computación). Pero, dado que no todos los términos cuentan con marcas formales (criterio morfológico) y dado que existen condiciones de variación en el discurso que derivan en procesos semántico-pragmáticos diversos (terminologización, banalización, etc.) es necesario subrayar la existencia de todo un aparataje teórico que influye en la identificación de la terminología, como pueden ser los criterios subjetivo, pragmático, textual, cognitivo o morfológico.

Si aplicamos la estrategia seguida por el modelo de Vega Moreno (2015: 252-295) para la selección de neologismos, basado en los criterios del OBNEO, a la determinación

del valor de especialización, podrían proponerse los siguientes criterios, a nuestro juicio, por orden de prioridad:

- a) **Criterio cognitivo:** hay sospechas de que la unidad sea un término. En el texto, el usuario considera que el sentido de la unidad léxica se desvía del sentido que tiene en el uso del discurso común (en caso de que exista este uso). El receptor es, en primera instancia, el que se inclina hacia determinar el valor especializado. Con todo, esta primera sospecha debe ser verificada a través de diferentes filtros (que son los criterios siguientes).
- b) **Criterio textual:** ¿dónde aparece? ¿cuál es su marco? El tipo de texto o situación en la que se usa el término es un criterio necesario, pero no suficiente para su identificación terminológica. La superestructura de los textos impone ciertas restricciones y condiciones al uso del léxico, el registro, la sintaxis, etc. que permiten contar con más argumentos que contribuyan a la identificación terminológica.
- c) **Criterio pragmático:** ¿quién lo usa? ¿a quién se dirige? El perfil del emisor y el destinatario, dependiendo de su grado de especialista, es, como en el criterio textual, un criterio necesario, pero no suficiente para la identificación terminológica de la unidad léxica. La relación discursiva entre los interlocutores impone condicionantes discursivos que se actualizan en estrategias recontextualizadoras.
- d) **Criterio cognitivo-funcional:** ¿representa o comunica conocimiento especializado? Teniendo en cuenta los criterios textual y pragmático (es decir, en qué situación se usa el término, quién lo usa y a quién se dirige) puede analizarse qué función pretende cumplir el texto en el que se inscribe la terminología y, junto con esto, si el candidato a término codifica conocimiento especializado o no. En caso de que pueda defenderse la existencia de este conocimiento especializado y de una función primordialmente representativa, podría defenderse la catalogación de la unidad candidata como término.

Podría sumarse el **criterio morfológico** a este listado de criterios. La existencia de marcas formales que permitan identificar el término puede ser en realidad el punto de partida (porque puede coincidir con el subjetivo), o, de forma contraria, el argumento final. A nuestro juicio, solo es relevante cuando la denominación del término ya se ha vinculado con anterioridad a un sentido especializado. De hecho, existen estudios que consideran que

el criterio morfológico no es relevante, como el de Picht (1984: 59 y ss.), Adelstein (2004: 5) o Cabré, Feliu y Tebé (2001b: 35).

De forma sintética, los criterios podrían organizarse de la siguiente forma, aunque hay que tener en cuenta que es posible superponer criterios y que, a efectos prácticos, suelen ser interdependientes en mayor o menor medida.

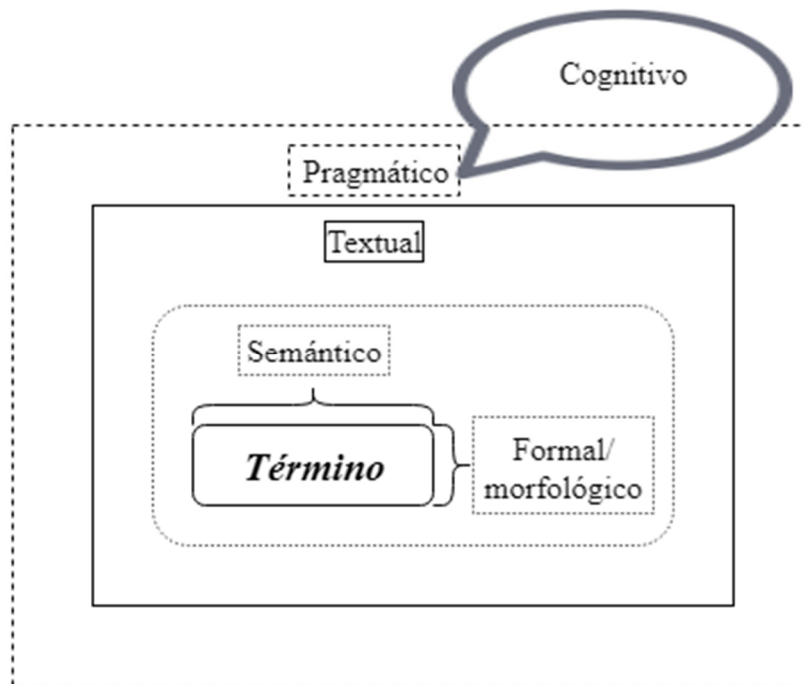


Ilustración 9: Esquema de criterios de caracterización terminológica y criterios de terminologicidad [elaboración propia]

A pesar de que adoptar un criterio cognitivo de partida (apoyado o no en el criterio morfológico) pueda ser considerado un criterio no científico, la existencia de filtros posteriores que recurran a argumentos lingüísticos y no a suposiciones permite trabajar con un método acorde con el objeto y el objetivo del análisis, de forma que se evita la completa subjetividad del usuario. De hecho, Chomsky (2006 [1968]: 112) usa el criterio de la intuición del hablante-oyente ideal como criterio para evaluar la gramaticalidad de las oraciones.

8.2. El principio de activación del valor especializado en el discurso común

Ahora bien, las condiciones antes mencionadas se aplican a la situación discursiva óptima (o prototipo de comunicación especializada) para la activación del grado de especialización. En otras palabras: el tipo de texto en el que aparece la unidad candidata a término, los interlocutores alrededor de los cuales gira dicha unidad y la función que

cumple la misma constituyen las condiciones óptimas para que la unidad active su valor especializado. No obstante, como ocurre con el fenómeno de banalización, puede darse el caso de que la situación discursiva no sea la óptima para que la unidad candidata a término active el valor especializado y, aun así, lo active. Dicho de otra forma, ¿es posible que el valor especializado quede activo en las unidades que usan los no especialistas? Si dicho valor se considera una categoría semántico-pragmática (y, por tanto, depende de la situación comunicativa, i.e. interlocutores, tipo de texto, etc.), cuando dicha situación comunicativa cambia, es lógico considerar que el valor especializado no se active.

Dos argumentos que usan algunos investigadores para discriminar el valor especializado en una situación comunicativa no especializada, atendiendo, sobre todo a su perspectiva cognitiva, son:

- a) Los términos pertenecen a un **sistema conceptual** analizado y delimitado por los especialistas. Cuando un lego usa una unidad léxica que codifica un concepto de dicho sistema conceptual, es posible defender que nos encontremos ante un término. Por ejemplo: un lego en una conversación coloquial, en un mensaje en una red social, en un foro o en un blog, puede hablar de *cáncer*, *tumor*, *tiroiditis de Hashimoto*, *pantalla táctil*, *wifi*, *hipoteca*, *sentencia* o *resolución*. Todas, por ello, en tanto que pertenecen a sus respectivos sistemas conceptuales, se considerarían términos.
- b) Ahora bien, el segundo argumento que podría esbozarse es que, a no ser que el **emisor** controle el sistema conceptual, no será posible que produzca textos especializados. Como el elemento característico del texto especializado suele ser la terminología, sería posible considerar que, bajo este argumento, no puede quedar activo el valor especializado de las unidades léxicas, puesto que dicho emisor no es capaz de ubicar el término en el sistema conceptual. De esta forma, el término queda usado como mera etiqueta, desprovista de su significado especializado, y, por tanto, empleando un significado aproximado. El término está constituido por denominación o forma (que es más o menos fija) y contenido (que es variable según la situación comunicativa). Al cambiar la situación, hablamos de que cambia el tipo de texto, el emisor, posiblemente el destinatario o la perspectiva. Todos estos cambios afectan, evidentemente, al contenido, que pasa a ser otro contenido. Esto hace considerar que en los textos producidos por legos no se usan términos.

Cabe, no obstante, hallar varios contraargumentos para cada una de estas cuestiones:

- a) En el caso del **sistema conceptual**, es evidente que, en la realidad, todo es ubicable en un sistema conceptual abstracto, por lo que sería posible afirmar, por tanto, que todo es término. Así, cuando el lego usa *comprar* o *vender* (conceptos básicos del campo de la economía), *hipoteca*, *fiebre*, *gato* o *perro*, que forman parte del sistema conceptual de sus respectivas materias, ¿sería posible considerarlas términos? A nuestro juicio no, puesto que su grado de incorporación en el caudal léxico de la lengua común a través de su frecuencia de aparición en el discurso común permite considerar que dichas unidades se usan en su forma de unidad léxica de la lengua común (o forma de unidad léxica no especializada o de no término).
- b) En el caso de la necesidad de que el **emisor** sea especialista y controle el sistema conceptual para producir textos especializados, creemos que esto no entra en conflicto con la posibilidad de que use términos, es decir, de que el valor especializado de ciertas unidades léxicas quede activo según ciertas condiciones de imitación, que se derivan en un uso impreciso de la terminología. La conclusión inmediata sería que los legos usan términos, pero con laxitud semántica e imprecisión, es decir, posiblemente desteterminologizados.

Por tanto, podemos negar los dos extremos:

- a) Se niega que la unidad usada por un lego sea término por el hecho de pertenecer al sistema conceptual de la materia.
- b) Se niega que la unidad usada por un lego no pueda emplearse como término por no tener perfil de especialista en ese ámbito.

Es necesario, por tanto, encontrar un criterio intermedio que permita explicar cómo se activa el valor especializado y por qué se activa en situaciones variadas y complejas de comunicación (no) especializada. Por tanto, sería conveniente tener en cuenta las palabras de Cabré y Estopà (2002: 146):

[...] la condición de valor especializado no es única ni estática, sino que, por un lado, una misma unidad léxica puede activar diferentes valores especializados en función de la temática en la que se use (variación horizontal); y, por otro lado, una misma unidad léxica dentro de un mismo ámbito especializado puede ser más o menos especializada pues el rasgo de valor especializado lleva asociado valores graduales de precisión, concisión y científicidad (variación vertical).

Hablaríamos, pues, de grado de valor especializado y dependería del grado de banalización o de frecuencia de uso/grado de incorporación en la lengua común. Para medir este grado de frecuencia o incorporación en la lengua común, podría recurrirse a tres estrategias, que también conllevan problemas teóricos y metodológicos:

- a) El primero consiste en realizar consultas sobre frecuencia en **corpus de referencia**. A continuación, exponemos diferentes casos de candidatos a término usados en el discurso común (conversación coloquial) entre legos con su correspondiente valor de frecuencia absoluta y número de casos por millón en el subcorpus de textos de español de España en el corpus CORPES XXI:

Ejemplo	Frecuencia absoluta	Casos por millón
He ido al banco a solicitar una <i>hipoteca</i> .	1.116	4,51 casos por millón
La <i>sentencia</i> dice que el piso es mío.	4.586	18,54 casos por millón
Le han dicho que tiene un <i>tumor</i> .	2.313	9,35 casos por millón
Le han diagnosticado un <i>cáncer</i> .	5.880	23,77 casos por millón
Han hecho recortes por el <i>déficit</i> .	2.922	11,81 casos por millón
Mi hijo tiene <i>fiebre</i> .	1.778	7,18 casos por millón
Dale a “ <i>ejecutar</i> ” en esa ventana.	429	1,73 casos por millón
Han pedido el <i>divorcio</i> .	871	3,52 casos por millón
El móvil no se conecta al <i>wifi</i> .	499	2,01 casos por millón
Vaya <i>depresión</i> tiene ⁹⁶	¿2.666?	¿10,78 casos por millón?
Le han diagnosticado <i>depresión</i> ⁹⁶	¿2.666?	¿10,78 casos por millón?

Tabla 18: Frecuencias absolutas de términos en CORPES XXI

Como puede comprobarse, los resultados son muy dispares y, al contrastar los diferentes casos, se observa que, en el corpus, lo que de verdad puede evaluarse, es la relevancia sociocultural de un fenómeno a través del léxico en un punto de la diacronía. Esto es, puede comprobarse la repercusión del *cáncer* a través de su banalización, en contraste con la de *divorcio* o *fiebre*, a nuestro juicio, más cercanas a la forma de *no término*. Por lo tanto, consideramos que la consulta de frecuencias en corpus no nos proporciona datos concluyentes sobre el carácter terminológico de una unidad léxica.

- b) La segunda estrategia consistiría en aplicar los **criterios** para distinguir **neologismos**, esta vez, aplicados a la distinción de la activación del valor especializado. En primer lugar, por tanto, se identifica un candidato a término determinado por la percepción psicológica/subjetiva del destinatario. En segundo lugar, se comprueba en una obra lexicográfica: si aparece en el diccionario, existe

⁹⁶ Este caso muestra la frecuencia absoluta de la forma *depresión*, sin distinguir situaciones discursivas, por lo que no es posible discriminar los sentidos en los casos por millón.

sospecha de que se haya incorporado al caudal léxico de la lengua común. No obstante, existen términos que se incluyen en el diccionario y que no se usan con frecuencia en la lengua común, como pueden ser *semasiológico* u *onomasiológico*. Además, hallamos el problema que suponen palabras de la lengua común como *divorcio*, *depresión*, *sentencia*, etc. Se trata de términos banalizados que ya ha asimilado la lengua común. Por tanto, tampoco podemos defender el criterio lexicográfico para nuestro objetivo.

- a) Para la tercera estrategia, cabría recurrir a los **tipos de emisores** según su **competencia cognitivo-comunicativa**, de forma que puedan graduarse de mayor a menor competencia. La siguiente tabla muestra los tipos de emisores vinculados a textos que pueden producir y la posibilidad de activar el valor especializado según dichas condiciones.

Grado de especialización	Emisor	Tipo de texto	Activación del valor especializado
+	Especialista	Artículo científico	✓
		Tesis	✓
		Artículo científico de introducción	✓
	Instruido	Prensa divulgativa para instruidos	¿?
		Prensa especializada de divulgación general	¿?
		Prensa general	¿?
	Lego	Consulta en revista de divulgación general	¿?
		Consulta en foro a especialistas	¿?
		Conversación coloquial	¿?
-			

Tabla 19: Posibles casos de activación del valor especializado según los tipos de texto y los emisores

Así, puede comprobarse que el perfil del instruido y del lego suponen el mayor problema, puesto que no es sencillo evaluar su nivel de competencia cognitivo-comunicativa, lo que nos hace sospechar que la activación del valor especializado conlleva más problemas con estos perfiles. Es más, podemos defender que el criterio cognitivo es el más útil pero complejo en su aplicación para resolver nuestro problema.

Para explicar, por tanto, el fenómeno de activación de los diferentes valores de especialización, proponemos el siguiente principio de activación del valor especializado en el discurso común.

En primer lugar, debe plantearse la comunicación lingüística como una sucesión de planos Y graduales, condicionados por características discursivas (cognitivas y pragmáticas) X y Z, de forma que cada plano Y constituye al fin y al cabo un plano de

actualización discursiva: un *tipo de discurso*. Gráficamente, el planteamiento inicial puede esbozarse como sigue:

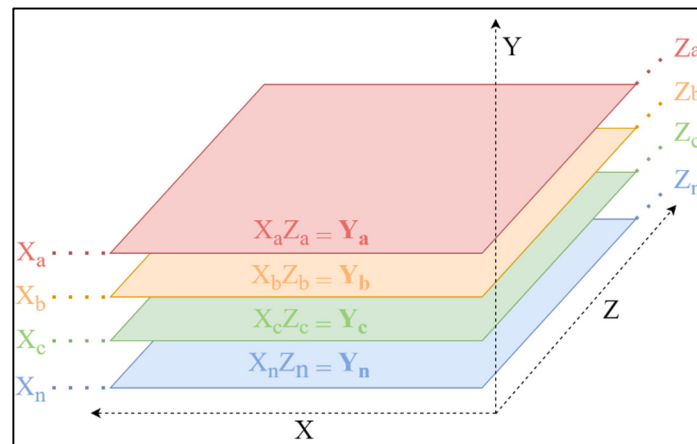


Ilustración 10: Ejes X, Z e Y en el principio de activación del valor especializado en el discurso común [elaboración propia]

La leyenda para este gráfico y para el principio que esbozamos en este apartado (que se desarrolla a continuación) es la siguiente:

X = coordenadas de variación pragmática 1: interlocutores y función textual.

Z = coordenadas de variación pragmática 2: tipo textual, registro, etc.

XZ = la situación discursiva, variables de actualización del discurso lingüístico (pragmáticas): interlocutores, función textual, tipo textual, registro, etc. Son graduales.

Y (= XZ) = plano de actualización lingüística (tipo de discurso). Es gradual. De esta forma, Y_a , Y_b , Y_c e Y_n son diferentes tipos de discursos (son grados de actualización discursiva; en nuestro caso, grados de especialización) condicionados por las variables que aparecen en los ejes X y Z. En este trabajo se usa Y_a para hacer referencia al discurso especializado e Y_n para hacer referencia al discurso común.

UL = unidad candidata a término.

T_1 = activación del valor especializado.

T_0 = no activación del valor especializado.

Tabla 20: Leyenda para el principio de activación del valor especializado en el discurso común

Así, el plano Y_a , por ejemplo, representaría el plano del discurso especializado, condicionado por la conjunción de características $X_a Z_a$, que podrían ser la participación de interlocutores especialista-especialista, una función textual eminentemente representativa del conocimiento y un tipo de texto artículo científico. Los planos Y_b , Y_c , etc.

representarían otros planos de actualización discursiva diferentes, dado que las características XZ también lo son. No obstante, debe recordarse que tanto X como Z son ejes graduales, por lo que pueden adquirir valores diferentes para un mismo plano.

El valor especializado de una UL solo se activa si en su discurso de aparición (en el plano Y_a) se cumplen ciertas condiciones favorecidas por la *situación discursiva especializada* (que, recordemos, vienen dadas por la unión de X_a y Z_a). De forma esquemática:

UL según X_aZ_a en $Y_a = T_1$
Es decir, la unidad léxica candidata a término empleada en el discurso Y_a (discurso especializado) según las características de la situación discursiva X_aZ_a permiten la activación del valor especializado de la mencionada unidad léxica, lo que concuerda con el concepto de sentido situado.

Tabla 21: Principio de activación del valor especializado en el discurso especializado

Para explicar que en Y_n (el plano del discurso común) una palabra se *actualiza* en término (es decir, activa su valor especializado) es necesario que ese término *porte* de forma implícita (o virtual) las condiciones de actualización X_aZ_a del plano Y_a . De esta forma, podría esquematizarse como sigue:

UL(X_aZ_a) en $Y_n = T_1$
UL en el plano Y_n (el del discurso común), que tiene coordenadas de actualización X_nZ_n , se activa si UL lleva consigo virtualmente las variables de activación X_aZ_a del plano Y_a , es decir: el valor terminológico se activa si en Y_n se usa UL con el sentido que tendría usado entre especialistas en situación de comunicación especializada (i. e. si UL en su situación prototípica se activa como valor semántico-pragmático).

Tabla 22: Principio de activación del valor especializado en el discurso común

Así, de forma gráfica, podría representarse de la siguiente forma:

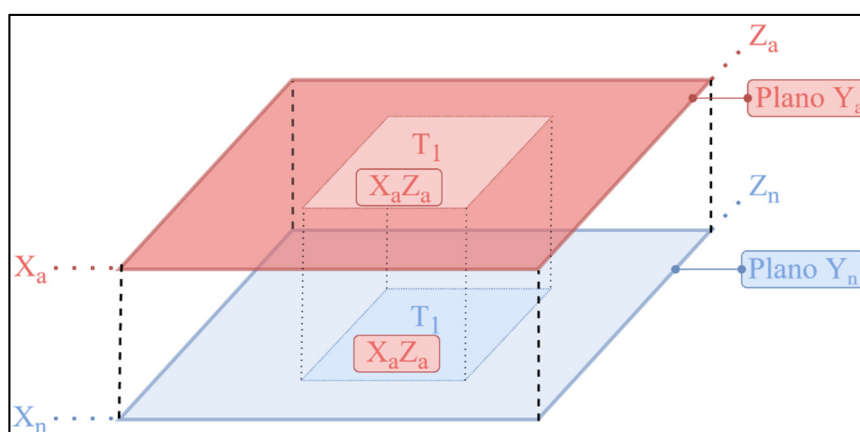


Ilustración 11: Principio de activación del valor especializado en el discurso común [elaboración propia]

Y si usamos como ejemplo la unidad léxica *tumor*, podría explicarse dicha fórmula y dicho gráfico de la siguiente manera:

<i>Tumor</i> usado en el discurso Y_a (discurso especializado)	$Z_a X_a \rightarrow T_1$	En el discurso especializado, la situación discursiva (especialista-especialista, función textual representativa y tipo de texto artículo científico) hacen que <i>tumor</i> se considere término.
<i>Tumor</i> usado en el discurso Y_n (discurso común)	$Z_n X_n(Z_a X_a) \rightarrow T_1$	En el discurso común, la situación discursiva (lego-lego, función textual comunicativa y tipo de texto conversación coloquial) hacen que <i>tumor</i> solo se active si se emplea simulando o reflejando las características de la situación prototípica de uso del término (es decir $X_a Z_a$). Dicha excepción en la situación comunicativa lleva consigo obligatoriamente un uso destemologizado con grados de laxitud semántica con respecto al significado del término original.

Tabla 23: Contraste del principio de activación del valor especializado en el discurso especializado y en el discurso común

Así, si un lego o instruido se dirige a un especialista y le habla de *depresión*, *ansiedad*, *divorcio*, *fiebre*, *déficit*, *ejecutar*, etc., el emisor usa el término como forma imitada, cognitivamente no tiene carga conceptual del plano Y_a , pragmáticamente solo imita las condiciones de $X_a Z_a$. Aunque es posible comprobar que existen ciertas diferencias entre cada caso.

De esta forma, el principio de activación del valor especializado en el discurso común rezaría que: una unidad léxica activará su valor especializado y, por tanto, será considerada término siempre que se use bajo sus condiciones pragmáticas de activación óptima o, en su defecto, *cuando* dichas condiciones se encuentren ligadas a la mencionada unidad léxica de forma virtual a causa de la *intención* del emisor de activar el valor especializado en una situación discursiva que no es óptima para la unidad léxica.

El problema de este principio se halla en cómo determinar *cuándo* se produce esa virtualización de las condiciones y cómo se identifica la *intención* del emisor.

Podríamos concluir, por tanto, que el instruido es el único capaz de, no siendo especialista, simular las condiciones de comunicación especializada (debido a su formación). El lego, sin embargo, no es capaz de simular dicha condición de comunicación especializada. El problema vuelve a replantearse cuando el lego emplea unidades léxicas monorreferenciales, exclusivas de la comunicación especializada, con escasa frecuencia en el discurso común y que plantean un problema desde la cognición, puesto que el emisor, con total seguridad, tiene escasa competencia en el ámbito del que trata:

- (24) Hola buenas tardes el día 2 de agosto sufrí una caída de escaleras golpeándome con el filo la región lumbar muy cerca de la columna, inmediatamente se me inflamó el área y desde entonces tengo dolor al tacto y sigo teniendo inflamado esa misma área como un bulto que es doloroso al toque, fui al medico general y me dijo que es un esguince de cadera me mandó 3 inyecciones de **B12** con **lidocaína** y **dexametazona**, **meloxicam** con **metocarbamol** y aplicar una crema de **diclofenáco** pero ya llevo mas de una semana y

aunque eh visto algo de mejoría la inflamación no cede, sigo teniendo la bolita. No me molesta al caminar ni tengo debilidad en manos ni pies solo la espalda alta siento como cansancio y hormigueo como ya dije solo el área me molesta al tacto, al recargarme o acostarme boca arriba. Sera algo serio? tengo 31 años soy varón no eh tenido temperatura alta ni nada de eso⁹⁷.

(25) Nuevo Docu-programa de tv llamado ADN te ayuda a solucionar cualquier tipo de dudas sobre parentescos a través de pruebas de ADN. Si te sientes identificado con algunos de los siguientes casos, o conoces a alguien, llama e informate. 626 105 351⁹⁸.

9. EL TEXTO ESPECIALIZADO Y EL GRADO DE ESPECIALIZACIÓN

En este capítulo tratamos de justificar el grado de especialización textual tomando el texto especializado como punto de partida, puesto que es este objeto el que, dependiendo de la situación comunicativa, se ve modificado y acaba configurándose como diferentes realizaciones lingüísticas. Por lo tanto, en el marco de la comunicación especializada (esbozada en el apartado II. 2.2., p. 41) analizamos a continuación la definición del texto especializado como objeto idealizado y en su máxima condición de especialización, así como sus características distintivas para, a partir de dicha caracterización, situar el objeto texto especializado en el eje vertical de gradación y describir las teorías y métodos que se han empleado para estudiar y clasificar los textos del grado de especialización.

9.1. Primera definición y aproximaciones al texto especializado

En primer lugar, cabe subrayar que la caracterización epistemológica del texto especializado surge de la necesidad comunicativa que tienen los especialistas de hacer referencia especial para denominar sus objetos propios (Cabré *et al.*, 2001: 176). El texto especializado es producto de la variación diastrático-diafásica y puede describirse de igual forma que los textos divulgativos, es decir, a partir de sus condiciones discursivas y sus características textuales (Cabré y Estopà, 2005: 36), que, según Da Cunha *et al.* (2011: 266) se identifican analizando «the text author, the potential reader, the structural organization and the lexical units' selection».

⁹⁷ Fuente: Emmanuelortega (14 de agosto de 2016): «Caída de escalera bulto en zona lumbar», *Portalesmedicos.com*. Disponible en: http://www.portalesmedicos.com/foros_medicina_salud_enfermeria/ubbthreads.php/posts/195227/Caida_de_escalera_bulto_en_zon.

⁹⁸ Fuente: Anónimo (2 de febrero de 2009): «Re: Prueba de paternidad gratis en Docu-programa», *Portalesmédicos.com*. Disponible en: http://www.portalesmedicos.com/foros_medicina_salud_enfermeria/ubbthreads.php/topics/8532/Prueba_o_test_de_paternidad.

Así, el texto especializado puede definirse usando las siguientes citas, que se complementan:

[...] la producción lingüística que sirve para expresar y transmitir conocimiento especializado, que tiene una serie de características lingüísticas que le confieren especificidad en el conjunto de textos producidos en una lengua, y que presenta una serie de características pragmáticas que determinan los elementos específicos del proceso de comunicación (tema, usuarios y situación comunicativa) (Cabré *et al.*, 2001: 181).

[...] productos predominantemente verbales de registros comunicativos específicos, que se refieren a temáticas propias de un dominio de especialización, y que responden a convenciones y tradiciones retóricas específicas; por lo tanto, en dependencia del tipo de disciplina pueden ser más o menos dependientes de la cultura y la época dada [...]. Los textos especializados se realizan en clases textuales específicas del discurso de especialización (artículo de investigación, ponencia, artículo de divulgación científica, comunicados científicos a la prensa, etc.) (Ciapuscio y Kuguel, 2002: 43).

Los textos especializados son las producciones lingüísticas, orales o escritas, que se producen en escenarios de comunicación profesional y sirven exclusivamente a una finalidad profesional. Se reconocen los escenarios profesionales por los interlocutores que actúan en la situación, por el tratamiento de una temática relativa al dominio o dominios concernidos por la profesión y por la finalidad esencial de buscar la información del receptor, aunque para ello se utilicen estrategias discursivas distintas (Cabré, 2007a: 90).

Con ello, analizamos los elementos pragmáticos, lingüísticos, cognitivos y estructurales que caracterizan el texto especializado.

9.2. Características del texto especializado

9.2.1. Características pragmáticas

El concepto de *entorno* (Coseriu, 1962: 309) ha sido determinante en los estudios posteriores sobre la comunicación humana, puesto que pretende explicar cómo la situación comunicativa (concepto posterior al de *entorno*) afecta a la estructuración de los discursos. Dentro de la lingüística comunicativa, la teoría del registro de Halliday (1982 [1976]) permite explicar la situación comunicativa atendiendo principalmente a tres dimensiones, que coinciden, a nuestro juicio, con los factores extralingüísticos (o pragmáticos) que

defiende Hoffmann (Hoffmann, 1998 [1987]: 85): las variables sociales, la intención, la situación y el objeto comunicativos. Las variables que propone Halliday son:

- a) El **campo**: hace referencia al tema tratado. En el caso del texto especializado, hace referencia a un tema especializado⁹⁹ o, de forma más precisa, a la perspectiva desde la cual se trata un tema concreto, normalmente, una perspectiva que no comparten los interlocutores del discurso común, es decir, el público en general.
- b) El **modo**: hace referencia al canal de comunicación, que podría asimilarse al tipo de texto mediante el cual se estructura el contenido y a la variación diamésica, siendo posible distinguir textos escritos y orales. En el caso del texto especializado, que puede adoptar forma escrita u oral, suele hablarse de ciertos tipos textuales de estructura más o menos firme o que aceptan escasa modificación, como los artículos científicos, las patentes o las tesis doctorales, entre un largo etcétera.
- c) El **tenor**. Puede ser **tenor funcional** (denominado simplemente tenor), que hace referencia a las funciones lingüísticas que activa el texto (*vid. supra* apartado II. 2.5., p. 50) o **tenor interpersonal** (o tono), que hace referencia a la relación existente entre los interlocutores (*vid. supra* apartado II. 2.4., p. 46), normalmente distinguible en registros más o menos formales. En el texto especializado, el tenor funcional suele corresponder con la función referencial de Jakobson (en concreto con la función representativa anteriormente expuesta) y el tenor interpersonal suele ser formal entre especialistas.

9.2.2. Características cognitivas

El texto especializado representa y transmite conocimiento especializado (*vid. supra* apartado II. 6.4., p. 99). Toma como referencia un dominio cognitivo idealizado propio de los especialistas de un ámbito (Domènech, 2006: 49), de forma que el emisor (especialista) tiene el control conceptual del sistema: es decir conoce los nodos y las relaciones entre nodos.

⁹⁹ A nuestro parecer, no es posible hablar de temáticas especializadas, puesto que, *stricto sensu*, no existen. Lo que existe son maneras o perspectivas especializadas (con mayor abstracción) de tratar temas concretos pero epistemológicamente y ontológicamente consideramos que no es posible defender la existencia de *temáticas especializadas*.

9.2.3. Características lingüísticas

Más allá de las características pragmáticas que condicionan el texto y de las cognitivas que lo vehiculan, Cabré *et al.* (2011: 302) plantean la pregunta sobre si es posible hallar características lingüísticas en el texto especializado que lo diferencien del no especializado.

Además de la precisión, concisión, impersonalidad y sistematicidad vinculadas con el texto especializado (Freixa, 2002b: 179), se describen y desarrollan algunas características del texto especializado, tanto léxicas como gramaticales, en los trabajos de Kocourek (1991) y Sager, Dungworth y McDonald (1980). Son estas dos últimas referencias las empleadas como marco teórico de contraste en el proyecto ESPETEX, del proyecto *TEXTERM-2. Fundamentos, estrategias y herramientas para el procesamiento (2003-2005)*, que tenía por objetivos

[c]omprobar a través de un corpus suficientemente representativo si se confirman las características gramaticales que los manuales de lenguajes especializados atribuyen a los textos de especialización.

En caso de que no se confirmen en parte o totalmente, intentar encontrar y establecer algunos de los factores gramaticales específicos que diferencian los textos especializados (Cabré, 2007a: 97-98).

Este primer acercamiento derivó en diversos trabajos que pretendían diferenciar la especialización de textos a nivel lingüístico usando un corpus de un mismo tema pero diferenciado en cuanto a los emisores (especialista y periodista) (Cabré *et al.*, 2010: 453), de forma que pudiesen emplearse las «diferencias más representativas como criterios para la elaboración de una herramienta semiautomática de detección de textos especializados» (Cabré *et al.*, 2010: 454), que, de hecho, se pone en marcha en el trabajo de Da Cunha *et al.* (2011). Este argumento ya lo defendieron Cabré *et al.* (2001: 181) al afirmar que los textos especializados «presentan algunas características idiosincráticas susceptibles de ser detectadas en soporte informático por medio de procesos automáticos o semiautomáticos».

9.2.3.1. Características gramaticales

A juicio de Cabré *et al.* (2007: 853), los elementos gramaticales que predominan en el discurso especializado deben comprobarse de forma empírica sobre «corpus reales representativos de la diversidad de las producciones especializadas». De acuerdo con Cabré

(2007a: 105 y ss.), el texto especializado puede caracterizarse gramaticalmente según los siguientes elementos:

- Predominio de los sustantivos.
- Predominio de la nominalización de formas verbales.
- Expansión adjetival de los sustantivos.
- Ausencia de la segunda persona del singular y del plural.
- Uso considerable de la tercera persona del singular, reforzada con el sujeto impersonal.
- Predominio del presente de indicativo respecto al tiempo pasado.
- Predominio de la voz pasiva.

Publicaciones posteriores con el mismo fin (Cabré *et al.*, 2010: 455) constatan la predominancia de las secuencias «sustantivo + adjetivo calificativo» y los sustantivos acabados en «-ción» en el texto especializado, así como otros elementos como la conjunción *o*, la voz pasiva o marcadores metadiscursivos¹⁰⁰. A los elementos descritos se añaden nuevos hallazgos, como son la predominancia de pronombres tónicos y átonos y los tiempos en subjuntivo (ibíd.: 459).

Para constatar estos elementos, Cabré *et al.* (2010) trabajan sobre un corpus etiquetado gramaticalmente (*POS tagging*) con textos especializados y no especializados de prensa, de los cuales extraen de forma automática los patrones gramaticales para cada subcorpus. El rasgo lingüístico se considera relevante si su frecuencia de aparición en los dos subcorpus difiere en más de un 20 % (ibíd.: 456).

9.2.3.2. La sintaxis

La sintaxis ha sido uno de los elementos menos estudiados empíricamente en la investigación en torno al grado de especialización textual. Numerosos son los autores que defienden el componente sintáctico del discurso especializado como uno de los elementos diferenciables entre el discurso especializado y no especializado. Así, Ciapuscio (2003: 30) defiende que «los factores funcionales, situacionales y temáticos tienen su correlato en el nivel de la forma lingüística, tanto en la sintaxis como en el léxico».

¹⁰⁰ En contraste, el discurso común estaría caracterizado gramaticalmente por «la predominancia de otros rasgos en el discurso general, como son los nombres propios, los tiempos en pasado y los verbos en 1.ª persona del singular» (Cabré *et al.*, 2010: 458).

Desde una perspectiva muy general, Hoffmann (1987) determina el grado de libertad de la sintaxis en el paradigma de la forma lingüística, de manera que cuanto más elevado es el nivel de abstracción de un texto, más controlada es la sintaxis (*vid. supra* Tabla 6, p. 84). En su trabajo de 1985 (105 y ss. en Arntz y Picht, 1995: 43), Hoffmann sintetiza alguna de las características morfosintácticas que atribuye a los textos especializados:

- a) Uso de las formas verbales en presente, especialmente en tercera persona del singular. A esta idea Martínez Linares (2007) también suma la tendencia del empleo de formas verbales no personales, como el infinitivo, los gerundios o los participios, en construcción absoluta o concertada.
- b) Uso de la pasiva o la pasiva refleja.
- c) Tendencia a la nominalización en detrimento de la verbalización (Cabré y Estopa (2005).
- d) Tendencia de la aparición del adjetivo.
- e) Tendencia al uso del singular con más frecuencia que el plural.

En cuanto a las características sintácticas sistemáticas del texto especializado, Cabré y Estopa (2005) defienden la existencia de estructuras poco complejas y el uso de oraciones de escasa longitud, si bien apuntan excepciones como los textos argumentativos o jurídicos. Vivanco (2006) también apuesta por la coordinación y la yuxtaposición como uniones más representativas aunque pueden expandirse mediante la subordinación. A esto cabe puntualizar los elementos que defiende Vázquez y del Árbol (2006: 159-163), ya que afirma que en la oración simple predominan las oraciones declarativas con verbo copulativo y en la oración compuesta abundan los nexos completivos y relativos con frecuente uso de la subordinación sustantiva y adjetiva. En el lado opuesto, no suelen aparecer oraciones exclamativas ni interrogativas.

Basándose en la propuesta de Mitrofánova (1985), Álvarez (2008) determina en un estudio sobre un corpus español-ruso (tan solo 9 textos médicos) la inesperada frecuencia de oraciones simples no complejas, y la escasa presencia de oraciones coordinadas. También destaca la presencia marginal de los giros de gerundio (lo que contradice lo postulado por algunos investigadores) y la frecuencia de los giros de participio.

A nuestro juicio, en la literatura científica existen algunas contradicciones sobre el tipo y longitud de las oraciones que caracterizan al discurso especializado. Desde nuestro punto de vista, debería atenderse, en primer lugar, al grado de especialización y, en segundo lugar, al espacio que ocupan dichas oraciones, tanto en el tipo de texto como en el espacio

discursivo (es decir, determinar si la oración aparece en un blog, un artículo científico o una revista de divulgación, entre otros, y, además, si forma parte de un resumen, de una exposición introductoria, de un párrafo sobre metodología o incluso una conclusión; concepto de *secuencia textual* de Adam (1990, 2001 [1992])). Consideramos que se trata de factores que pueden incidir en la descripción de la actualización sintáctica en el discurso especializado (y, por extensión, en la descripción del grado de especialización textual). Por estas razones, la sintaxis no queda incluida dentro de nuestro estudio, ya que, desde nuestro punto de vista, todavía son necesarios estudios empíricos que demuestren la vinculación que existe entre la sintaxis y el tipo, la secuencia o la clase textual. A nuestro juicio, se trata de una de las líneas principales de investigación que esta tesis debe proyectar.

9.2.3.3. Características léxicas

Como exponíamos en el apartado II. 5.3. (p. 77), y teniendo en mente las citas que lo avalan, los investigadores coinciden en considerar el léxico especializado (la terminología¹⁰¹) como el elemento característico del texto especializado, siempre caracterizado por factores de índole funcional y situacional (Ciapuscio, 2003: 90). No obstante, las investigaciones en torno a dicho elemento y en torno a su frecuencia parecen indicar que no se trata del elemento discriminante del grado de especialización textual aun cuando sí que es posible hablar de que existe una correlación clara entre el grado de especialización y el número y tipos de unidades léxicas especializadas que contiene el texto (Domènech, 2006: 73).

Ciapuscio halla ciertas diferencias en cuanto a la presencia y tratamiento de la terminología en los textos del grado de especialización: en los textos más especializados se produce una alta densidad terminológica y escasa reformulación; en el nivel intermedio escasea la terminología y se perciben formas parafrásticas, definiciones, ejemplos, etc., y en el nivel más divulgativo se evita la terminología o se define y explica (Ciapuscio, 2003: 90) Esto concuerda con algunas de las conclusiones de la tesis de Freixa (2002b: 265), que sostiene que en los textos más especializados la variación denominativa es un fenómeno menos frecuente que en los textos no especializados, tanto porque existe un número menor de nociones representadas por más de una denominación, como porque los grupos polidenominativos presentan un número inferior de denominaciones diferentes.

¹⁰¹ Las características del término como objeto lingüístico se exponen en el apartado II. 5.6.

A pesar de que la hipótesis que relaciona densidad conceptual o cognitiva (y, en definitiva, densidad terminológica) y grado de especialización parece ser la hipótesis más planteada, lo cierto es que los estudios empíricos (Domènech, 2006: 207, 212-213) no demuestran dicha creencia:

- a) La **densidad terminológica**, entendida como el número de términos que contiene un texto en relación con el número total de palabras, parece ser una medida insuficiente para obtener un parámetro cuantitativo válido para discriminar un texto en función de su nivel de especialización. De hecho, la tesis doctoral de Domènech (2006: 208, 212) no solo muestra que las diferencias de densidad terminológica entre niveles son poco significativas, sino que tiene lugar, precisamente, de forma inversamente proporcional al grado de especialización.
- b) La **diversidad terminológica**, entendida como el número de términos diferentes que contiene un texto en relación con el número global de ocurrencias terminológicas, también parece una medida insuficiente para obtener un parámetro cuantitativo válido para discriminar un texto en función de su nivel de especialización. Sin embargo, esta diversidad, medida a partir de un cálculo matemático que tenga en cuenta el número de ocurrencias terminológicas de un texto, el número de términos diferentes, la frecuencia de aparición de cada término y la proporción global dentro del texto de la distribución de esta frecuencia, parece ser una medida bastante refinada que puede permitir obtener un parámetro cuantitativo estadísticamente válido para discriminar un texto en función de su nivel de especialización (ibíd.: 216).

9.2.4. Características superestructurales

Recordamos que Dijk (1992 [1983]: 143) defiende que la superestructura determina el orden global del texto. Según Ciapuscio y Kuguel (2002: 43), los textos especializados seleccionan «clases textuales específicas» (nosotros lo denominamos *tipos textuales*)¹⁰² del discurso especializado: artículo de investigación, ponencia, artículo de divulgación

¹⁰² En esta tesis doctoral, no deben confundirse dos términos estrechamente relacionados, que la bibliografía científica ha debatido desde hace décadas: (i) empleamos el término *clase* desde el punto de vista de las ciencias de la computación para hacer referencia a la variable dependiente del estudio, es decir a los tres estadios del grado de especialización; no empleamos el término *clase* en el sentido de Isenberg (1987: 101) (*vid. supra* apartado II. 3.3., p. 57); (ii) por otro lado, empleamos el término tipo textual para hacer referencia a una superestructura, a una forma específica de texto definida dentro de una tipología textual (Isenberg, ídem).

científica, comunicados científicos a la prensa, etc. Sobre esto precisan Cabré y Estopà (2005: 36) que «desde el punto de vista formal, los textos especializados no se distinguen de los generales más que por el hecho de seleccionar géneros textuales específicos».

9.3. Principio de variación y eje vertical

Teniendo como punto de partida el principio de variación en la comunicación, cabe traer a la mente de nuevo la importancia de la estructuración discursiva en torno a los ejes horizontal y vertical, empleados en multitud de investigaciones (Reindhardt, 1966; Hoffmann, 1998 [1987]: 62-68; Cabré, 1993: 141; 1999: 156; 2002b: 30; Cabré, 2002: 93; Cabré y Estopà, 2002: 146; Domènech, 2006: 72; entre un largo etcétera) como constructo teórico para explicar la gradación. En nuestro caso, emplearemos el eje vertical como eje estructurador de la gradación de especialización (es decir, de la perspectiva y abstracción con la que el texto trata el tema).

Cabré (2002: 19), por ejemplo, hace hincapié en la gradación como consecuencia de la multitud de situaciones en los que se puede actualizar la situación comunicativa¹⁰³:

De acuerdo con este criterio, los textos pueden ser más o menos precisos, concisos, sistemáticos, literales, objetivos, impersonales o polifónicos en función de diferentes variables: el tema, la función dominante, los destinatarios, el ámbito de comunicación o la estrategia comunicativa.

Podemos decir, en consecuencia, que se trata de criterios graduales que pueden representarse en un eje en el que cada texto adquiriría un determinado valor para cada uno de los criterios.

Sin embargo, esta gradación, evidentemente, se articula en torno a dos realizaciones totalmente opuestas que toma como referencia el máximo y mínimo valor de especialización: el texto especializado y el texto no especializado (Cabré, 2002: 22).

9.4. Clasificación binaria y *continuum*

Así pues, es posible distinguir dos posturas a la hora de realizar una tipología del texto especializado. La primera propone una división binaria en texto especializado y no especializado (Sager, Dungworth y McDonald, 1980), de acuerdo, principalmente, con el hecho de que el emisor sea especialista o no respectivamente. Para Cabré, esta perspectiva

¹⁰³ No obstante, en esta publicación, Cabré (2002: 31) se acerca al *continuum* según el usuario del texto.

no es posible, puesto que «una tipología textual no puede reducirse a una jerarquía binaria¹⁰⁴» (2002b: 22) sino que debe responder a multitud de dimensiones y criterios, ya que entre dos objetos completamente opuestos existe una «amplia gama de realizaciones que se ubica en un eje gradual» (idem) (Rodríguez Tapia, 2016a: 991).

La segunda postura, más aceptada y teorizada en la terminología descriptiva, se sitúa a favor del *continuum* o grado de especialización. Esta postura queda defendida en diversas publicaciones de tendencia funcionalista como las de Balboni (1986; cit. en Ciapuscio y Kuguel, 2002: 41), Gläser (1982 y 1993), Jacobi (1987) y Lauffler-Laurian (1983 y 1984), a los que se suman Cabré (1999, 2002b y 2008) y las publicaciones de Ciapuscio (1998), Ciapuscio y Kuguel (2002) y Ciapuscio, Kuguel y Otañi (2005).

La posición a favor del *continuum* toma como referencia inmediata los dos objetos textuales opuestos, que permiten contrastar sus características como una idealización del texto especializado y una idealización del texto no especializado. Entre ellos, existe una escala gradual de realizaciones «de tal manera que hay palabras o conceptos más o menos especializados o generales según el contexto y el uso que se les otorgue. Hay, pues, niveles de especialización y generalidad en todos los textos» (Cabré *et al.*, 2001: 178). Así, gráficamente podría representarse dicha posición teórica de la siguiente forma:

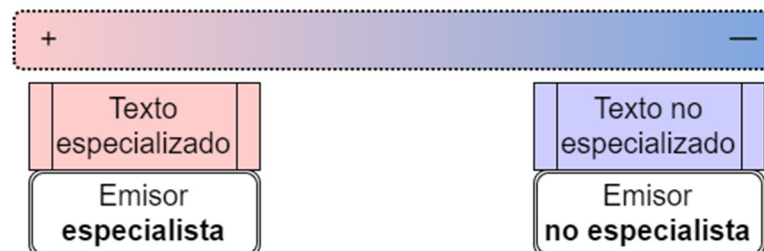


Ilustración 12: Gradación entre polos opuestos según la TCT [elaboración propia]

Este fue el modelo que seguimos en anteriores trabajos (Rodríguez-Tapia, 2015, 2016a, 2016b). No obstante, y como consecuencia de las conclusiones proyectadas en dichas investigaciones, a pesar de la utilidad teórica de esta división, consideramos más eficiente dividir cada polo en diferentes grados, de forma similar a la tipología con la que trabaja la TCT. De esta forma, de acuerdo con este modelo teórico, un texto será especializado si está emitido por un especialista y contará con grados de mayor o menor

¹⁰⁴ En los trabajos de Cabré sobre identificación del texto especializado (Cabré, 2007; Cabré *et al.*, 2010; Cabré *et al.*, 2014) se parte del emisor como criterio para determinar si un texto es especializado o es general, por lo que en principio, se trabaja con un criterio binario. Esto, no obstante, puede implicar el riesgo de incluir textos que, aun siendo emitidos por un especialista, se dirijan a un receptor no especialista (y por tanto, repercuta en la actualización lingüística del mismo). De esta forma, sería posible que un texto se clasifique como texto especializado pero en el grado de especialización sea próximo realmente al texto no especializado.

experticia según la situación comunicativa. A esto, desde nuestro punto de vista, habría que añadir que el mismo fenómeno se aplica al texto no especializado, emitido (de forma idealista) por un no especialista, y graduado de menor a mayor abstracción en el tratamiento de la temática.

Así, divergimos en la propuesta de la TCT en la división de las posibilidades del grado de especialización, tal y como sintetizamos en la siguiente tabla:

PROPUESTA DE LA TCT			NUEVA PROPUESTA		
Emisor		Clase texto	Clase texto		Emisor
Especialista	Especializado	(Gradación)	Gradación	Especializado	¿Especialista?
Controla el conocimiento especializado y es capaz de crear nuevo CE <i>Un médico en un artículo científico</i>				Controla el conocimiento especializado y es capaz de crear nuevo CE <i>Un médico en un artículo científico</i>	
<i>Un periodista especializado en medicina en un artículo para una revista de divulgación especializada</i>			Gradación	Semiespecializado	¿Especialista?
No especialista		No especializado/ divulgativo		¿Especialista?	
No especializado		No controla el CE y no es capaz de crear nuevo CE <i>Un estudiante de secundaria en un trabajo de clase</i>	Gradación	No controla el CE y no es capaz de crear nuevo CE <i>Un estudiante de secundaria en un trabajo</i>	

Tabla 24: Contraste de clasificación del grado de especialización entre la TCT y nuestra propuesta

Las diferencias principales se hallan en que, para la TCT, el emisor determina en un primer nivel la posibilidad de clasificar un texto según el grado de especialización, siendo necesario que el emisor sea un especialista para distinguir grados de especialización dentro del texto especializado. Así, sería posible distinguir textos muy especializados, textos semiespecializados y textos divulgativos. Nuestro modelo difiere en este sentido en que el emisor debe conjugarse con otros criterios, como el destinatario y las características lingüísticas del texto (en definitiva, la situación comunicativa en su conjunto), para clasificar el texto dentro de alguno de los diferentes grados en los que podemos dividir los diferentes niveles del *continuum*.

En relación con los interrogantes en el emisor de nuestra propuesta, en principio se trata de una aproximación siguiendo un hablante-oyente ideal que queda por demostrar empíricamente. Nuestra hipótesis considera que un especialista es capaz de usar sus conocimientos para diseñar textos especializados, semiespecializados o divulgativos,

independientemente del control conceptual que ejerza sobre el texto. De hecho, consideramos que este control conceptual solo afecta a la ausencia de confusiones nocionales en los textos, pero esto no repercute en la configuración discursiva del texto, que es, al fin y al cabo, lo que clasificamos según el grado de especialización.

A pesar de que esta tabla tiene por objetivo solamente ilustrar aquellas diferencias más importantes con respecto al modelo de clasificación de la TCT, cabe decir que en esta representación del modelo no incluimos el destinatario, que consideramos uno de los criterios más importantes para determinar el grado de especialización, como tratamos en el capítulo dedicado al análisis de resultados.

De hecho, si precisamos la representación gráfica del modelo, el texto semiespecializado debería hallarse a caballo entre los polos opuestos, de forma que los límites compartirían características y sería posible hablar del texto semiespecializado. En efecto, a no ser que se demuestre lo contrario y que se constituya como un objeto diferente, según esta representación, un texto especializado que se aproxime al texto no especializado podría ser considerado semiespecializado.

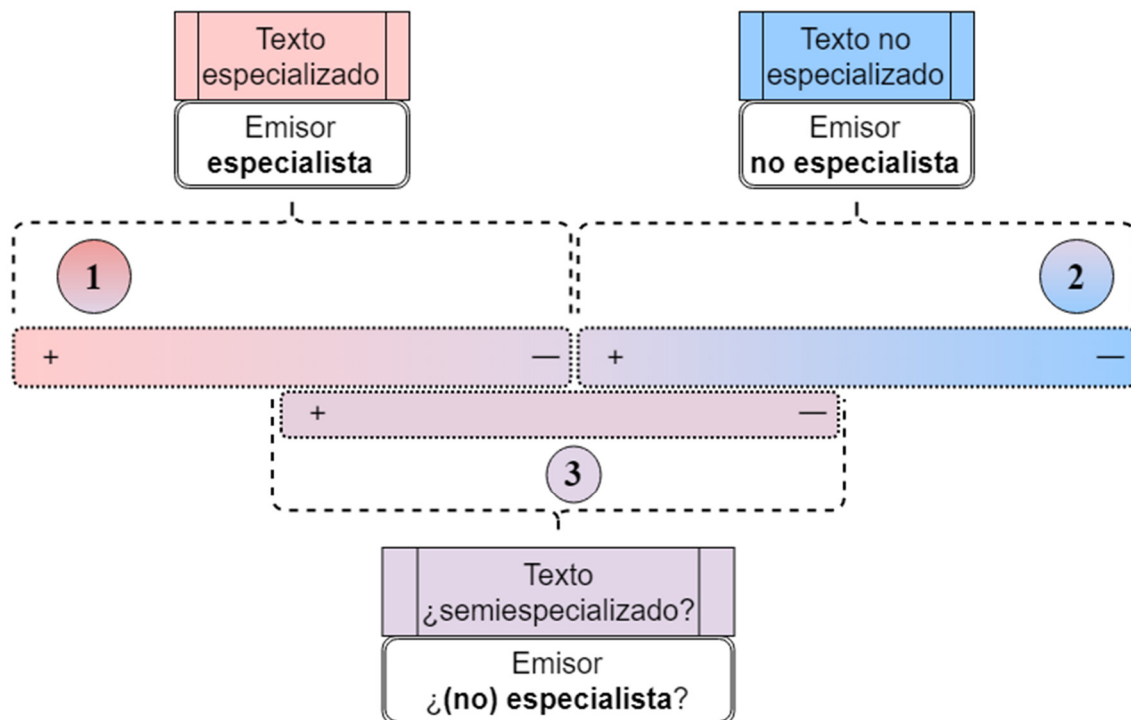


Ilustración 13: Gradación en y entre polos opuestos según nuestro modelo [elaboración propia]

Esta división nos hace comprobar que, entre el grado 1 y el grado 2, puede existir toda una gama de realizaciones que, bajo nuestro punto de vista, pueden coincidir en ciertas características. En ese caso, y dado que contamos con dos objetos textuales, suponemos que podemos hallar un grado 3, fruto de dichas coincidencias pragmático-lingüísticas y, en

ese caso, un tercer objeto textual, al que nos referimos como *texto semiespecializado*, a cuya definición nos aproximamos en Rodríguez-Tapia (2016b: 246). El problema subyace en los límites de clasificación entre cada objeto, para los cuales hay que determinar los criterios que permiten discriminar el grado. También debe tenerse en cuenta que el propio objeto de texto semiespecializado constituye un objeto gradual, que puede o no coincidir o superponerse con los grados correspondientes al texto especializado o al texto no especializado.

Para Cabré (2002a: 90) el grado de especialización en el texto depende «del valor que adquieren las variables que intervienen en su producción-transmisión-recepción», por lo que también es posible defender que la activación de esta propiedad no es intrínseca a cada texto, ni tampoco homogénea en su composición: es posible que en un mismo texto existan fragmentos de diferente grado de especialización (Cabré *et al.*, 2001: 180) (Rodríguez-Tapia, 2016a: 992).

Las mencionadas variables se ajustan a los parámetros de caracterización del texto especializado, es decir, condiciones cognitivas, semántico-pragmáticas y gramaticales (Cabré *et al.*, 2001: 181) y permiten, por tanto, adoptar diversos criterios para construir la tipología textual. Por ejemplo, atendiendo a la función transmisora de conocimiento según los interlocutores, Cabré (2002b: 31) divide los textos en:

[...] textos que transmiten el conocimiento de especialista a especialista, textos que lo transmiten de especialista a aprendiz de especialista, y textos de amplia difusión destinados al público interesado, pero sin competencia específica en la materia. A este último grupo de textos se les denomina también textos de divulgación especializada.

No obstante, este es tan solo uno de los criterios a los que se podría atender para clasificar los textos según su especialización; otros criterios podrían partir de la densidad terminológica en un texto, de la fijación estructural del mismo, de las funciones lingüísticas, del nivel de concisión, etc.

9.5. Modelos y métodos de clasificación del grado de especialización textual

Esta variedad en la selección de criterios ha derivado en el desarrollo de diferentes modelos y métodos de análisis para clasificación textual, dado el interés que supone la capacidad de predicción que permite la construcción de tipologías textuales (Cabré,

2002b; 21), que nos permite trabajar con tendencias o generalizaciones. Para Ciapuscio, Kuguel y Otañi (2005: 108):

En efecto, una tipología del texto de especialización que pueda dar cuenta de los distintos grados de especialización en los que el conocimiento especializado es instanciado en texto permite, a su vez, profundizar el estudio de cómo la comunicación de este conocimiento se adecúa no solo a la especificidad temática sino también a la experticia de los participantes y, de este modo, indagar acerca del modo en que esta instanciación impacta en el nivel de la formulación, es decir, en la terminología.

Una revisión de las tipologías textuales según diferentes modelos lingüísticos puede encontrarse en la publicación de Ciapuscio y Kuguel (2002: 43 y 44), que tratan las tipologías basadas en rasgos exclusivamente lingüísticos (Harweg, 1968), otras propuestas por parte de los modelos funcional-comunicativos basadas en rasgos internos y externos a los textos (Sandig, 1972) y según la jerarquía de criterios funcionales, situacionales y estructurales (Brinker, 1988), así como otras tipologías multidimensionales de modelos cognitivos basados en representaciones prototípicas en distintos niveles (Heinemann y Viehweger, 1991: 147).

Cabré (2002b: 19) también propone un modelo de análisis para la tipologización, basado en los estudios de Beaugrande y Dressler (1997) y Dijk (1980), entre otros. Así, distingue tres bloques estructurales:

- La estructura formal: se relaciona con el género y tipo textuales.
- La estructura informativa: se relaciona con la organización de la información.
- La estructura lingüística: se relaciona con las estructuras fonológica, ortográfica y morfológica, además de la sintáctica.

Por otra parte, el modelo de Monterde Rey (2002b: 124 y 125) engloba aspectos pragmáticos, estilísticos, funcionales y estructurales; y el de Ciapuscio y Kuguel (2002: 37-47), constituyéndose como modelo de corte cognitivo-comunicativo multinivel y multidimensional, abarca multitud de puntos de vista, cuyos aspectos más novedosos o relevantes son sintetizables en la siguiente tabla:

-
- Perspectiva de función como efecto del texto en contexto de interacción social.
 - Jerarquización y secuencialización funcional que posibilita clasificar los textos como monofuncionales o plurifuncionales.
 - Marco interaccional que contempla: papel social de los interlocutores según el grado de competencia sobre un área de conocimiento, tiempo y espacio en el que se inscribe el texto, categorización del texto según represente una fuente primaria o secundaria, y perspectiva sobre el tema.
-

-
- Normas estilísticas que contemplan elementos extralingüísticos y no verbales que condicionan aspectos léxicos y sintácticos.
 - Relevancia del nivel léxico y la densidad terminológica.
 - Nivel microestructural y macroestructural.
-

Tabla 25: Características principales del modelo de análisis del texto de Ciapuscio y Kuguel (2002)

Por nuestra parte, en Rodríguez-Tapia (2016a: 995-1001) diseñamos un modelo de análisis textual cualitativo y clasificación cuantitativa para el grado de especialización. Este modelo adopta una posición sincrética entre los modelos anteriormente descritos, e introduce elementos de las teorías de Cabré (1993, 1999) y Arntz y Picht (1995).

En cuanto al análisis y la caracterización del tercer objeto textual, el texto semiespecializado, son muy escasas las publicaciones que establecen criterios para distinguirlo. Normalmente, los estudios tan solo reconocen los tres niveles de especialización (normalmente atendiendo a abstracción o al tipo de emisor) pero no se proponen criterios como se hace con el texto especializado y el texto no especializado.

Así, por ejemplo, hallamos estudios que simplemente se refieren al texto semiespecializado como objeto textual, pero no se mencionan los criterios empleados para discriminarlo de las otras posibilidades textuales: «el corpus textual se extrajo de dos revistas que corresponden al discurso económico científico semiespecializado» (Diéguez, 2006: 515). En otros estudios, la aproximación al texto semiespecializado se realiza como objeto de transición, haciendo referencia al registro y al carácter divulgativo:

En nuestro trabajo consideraremos textos semiespecializados aquellos que, siendo especializados, tiendan a rebajar el registro científico-técnico como lo hace la divulgación, así como aquellos textos aparentemente divulgativos que traten el tema desde una perspectiva prácticamente teórica con un lenguaje y características correspondientes a un registro más elevado (Prieto Velasco, 2008: 71).

En definitiva, puede defenderse que entre el texto especializado y el texto no especializado existe un eje gradual de realizaciones lingüísticas. Estas, se acercan más a un polo o a otro dependiendo de la situación comunicativa. De acuerdo con nuestro punto de vista, es posible identificar un tercer objeto textual en la transición de dicho grado: el texto semiespecializado. Para ello, es necesario proponer un modelo de análisis multicriterio que integre todas las variables que pueden afectar al grado de especialización y que permita identificar los puntos de similitud y divergencia entre cada uno de los tipos de texto, especialmente aquellos relacionados con el perfil cognitivo-comunicativo de los interlocutores, la densidad terminológica y la densidad reformulativa.

10. LOS PROCEDIMIENTOS DE REFORMULACIÓN FORMAL EN TERMINOLOGÍA

La reformulación desempeña un papel importante a favor de la comprensión lingüística y, en el caso del texto especializado, se constituye como un procedimiento esencial para regular la comprensibilidad según el receptor. Por tanto, consideramos útil prestar atención al nivel microestructural para comprobar la relación entre el uso de la terminología y la activación de procedimientos de reformulación formal. Con este término nos referimos a las estrategias que existen para la interpretación de un texto o unidad y su posterior restitución y explicitación para el destinatario de la información, de forma que la modificación afecta tan solo a la estructura formal o también a la semántica y, por tanto, al sentido o al conocimiento decodificado (Rodríguez-Tapia, 2015: 4).

En anteriores trabajos (Rodríguez-Tapia, 2015 y 2016b) tratamos el problema de la reformulación en terminología, y concluimos que la incidencia de procedimientos de reformulación era similar en los diferentes grados de especialización, si bien los procedimientos empleados eran diferentes (Rodríguez-Tapia, 2015: 16-26, 28-29). Con todo, se proyectaban diferentes líneas de investigación futuras que afectaban tanto a la forma de emplear la herramienta de trabajo como a su propia configuración teórica:

a) Implementar, perfilar y acotar los instrumentos de análisis relacionados con la caracterización discursiva y con la estructura textual así como la tabla de análisis de PRFC [...]. En esta última herramienta, debe intentarse establecer equivalencias cuantitativas o porcentuales entre el número de procedimientos que se registran dentro del bloque conceptual y dentro del bloque formal, ya que existe una gran diferencia entre el número de los procedimientos observados en los dos bloques. Por ello, proponemos seguir profundizando en el estudio de los procedimientos de reformulación conceptual y prestar más atención a los procedimientos con mayor incidencia y relacionar sus causas.

[...]

a) Estudiar el comportamiento e incidencia de los PRFC en otras disciplinas para comprobar si ponen en marcha procedimientos diferentes o adicionales.

b) Determinar si la tradición discursiva es similar en las disciplinas humanísticas y en las ciencias naturales o tecnológicas, así como señalar qué implicaciones tiene la tradición discursiva en los PRFC.

c) Determinar si la clasificación según la especialización se realiza de la misma en campos disciplinares diferentes y si los tipos textuales obtenidos (TD, TSE y TE) son similares en cantidad o configuración textual (Rodríguez-Tapia, 2015: 29-30).

Así, este capítulo, sobre la base de los trabajos mencionados (Rodríguez-Tapia, 2015 y 2016b), pretende: a) reflexionar sobre el concepto de reformulación en general y sobre la reformulación en terminología en particular, y b) implementar la herramienta de

trabajo diseñada (ídem), simplificándola de forma que sea más manejable para nuestros propósitos.

10.1. La reformulación lingüística

La reformulación constituye un procedimiento de organización y progresión del discurso mediante el cual el hablante vuelve sobre un segmento de información anterior (implícito o explícito), lo reinterpreta y lo presenta de nuevo de forma diferente (Garcés Gómez, 2006: 654). Queda caracterizado por un proceso retroactivo que posibilita «explicar, rectificar, reconsiderar, recapitular o separarse de la formulación anterior» (Garcés, 2008: 69)¹⁰⁵, lo que permite a los hablantes poner en funcionamiento su conciencia metapragmática (Verschueren, 1999, 2000; Reyes, 2002, Portolés, 2004) y «elegir las formulaciones lingüísticas que consideran más adecuadas en cada momento para configurar el texto» (Garcés, 2008: 69). Además, permite avanzar en la comprensión del contenido y la intencionalidad del texto, ya que se trata de un procedimiento que garantiza la cohesión textual (Bach *et al.*, 2003: 1). Desde el punto de vista de Gülich y Kotschi (1987 y 1995), la reformulación tiene un marcado carácter interactivo, ya que su función principal es evitar los problemas de intercomprensión, tal y como avanzábamos en la descripción de las fases que se ponen en marcha (Rodríguez-Tapia, 2015: 6).

El proceso de reformulación puede constituirse como un proceso de autorreformulación, en el que el propio hablante es el que inicia la reformulación (por ejemplo, cuando un usuario explica o define un término o expresión con otras palabras), o como un proceso de heterorreformulación, en el que un hablante diferente reformula el enunciado original por causas relacionadas con la comprensibilidad del mismo, como en textos dialogados o polifónicos entre varios usuarios, en los que uno comenta, matiza o explica (entre otros) las palabras del otro.

La reformulación puede explicarse desde una perspectiva cognitiva y pragmática empleando la teoría de la relevancia (Sperber y Wilson, 1995: 260), pues el interlocutor identifica la posible falta de comprensión de un enunciado (y su consiguiente falta de relevancia comunicativa) y recurre al proceso de reformulación para garantizar que se alcanza la relevancia óptima. Cortés y Camacho (2005: 199-205) también se posicionan a favor de esta perspectiva, puesto que consideran la reformulación como un proceso que

¹⁰⁵ Garcés (2008) trabaja en su estudio con textos periodísticos principalmente.

tiene como objetivo facilitar la comprensión al receptor. De acuerdo con Garcés (2008: 75), las razones que motivan la reformulación son las que siguen:

- a) se pretende explicar lo dicho en una formulación previa;
- b) se quiere recapitular lo anterior mediante una nueva formulación;
- c) se hace una consideración y se expresa desde un punto de vista;
- d) se quiere establecer una separación de distinto grado;
- e) se rectifica lo dicho en un segmento anterior.

En el caso del español, las investigaciones se han centrado, sobre todo, en la clasificación de los marcadores que ponen en funcionamiento dicha operación, aunque encontramos multitud de publicaciones que tratan la diversidad de mecanismos de reformulación, como sintetiza Carmona (2015: 288), atendiendo a las publicaciones de Fuentes Rodríguez (1993), Martín Zorraquino y Portolés (1999), Cortés y Camacho (2005), Garcés (2008) o López Serena y Borreguero (2010):

Fuentes Rodríguez (1993)

- | | |
|--------------------|---|
| a) Parafrástica | f) Por condensación (etiqueta, conclusión y recapitulación) |
| b) Explicación | g) Por extensión (enumeración o definición y ejemplificación o particularización) |
| c) Denominación | |
| d) Corrección | |
| e) No parafrástica | |

Martín Zorraquino y Portolés (1999)

- a) Reformuladores explicativos
- b) Reformuladores rectificativos
- c) Reformuladores de distanciamiento
- d) Reformuladores recapitulativos

Cortés y Camacho (2005)

- a) No parafrásticos
 - Concreción propia
 - Resumen
 - Generalización
 - Relieve
 - Restricción
 - Rectificación
- b) Parafrásticos

Garcés (2008)

- a) Explicación
- b) Rectificación
- c) Recapitulación
- d) Reconsideración
- e) Separación

López Serena y Borreguero (2010)

- a) Parafrástica o explicativa
 - b) No parafrástica
 - Recapitulación
 - Reconsideración
 - Distanciamiento
-

Tabla 26: Algunos trabajos en español sobre los mecanismos de reformulación (Carmona, 2015: 288)

Como es posible observar en dicha tabla, las investigaciones hasta ahora (no solo en español) han distinguido dos tipos de reformulación¹⁰⁶:

¹⁰⁶ En Rodríguez-Tapia (2015) asimilábamos los conceptos de reformulación parafrástica y reformulación no parafrástica a los de reformulación explicativa y reformulación imitativa (Fuchs, 1994) respectivamente. En esta investigación,

- a) la **reformulación parafrástica**, que establece una relación de equivalencia semántica entre el segmento de referencia y el segmento reformulado; y
- b) la **reformulación no parafrástica**, que instaura un cambio de perspectiva enunciativa (Gülich y Kotschi, 1987; Bach *et al.*, 2003; García Negroni, 2009: 48).

10.1.1. La reformulación parafrástica

Es un procedimiento que establece una «relación de equivalencia discursiva basada en una equivalencia semántica» (Garcés, 2008: 71) de forma gradual, pero al mismo nivel jerárquico. De esta forma, la reformulación parafrástica no es una propiedad de los enunciados, sino un procedimiento, «una estrategia cognitivo-lingüística» que anula las diferencias discursivas a favor de las semejanzas (Garcés, 2008: 72).

Atendiendo al funcionamiento semántico-pragmático del posible marcador reformulador, Rossari (1990, 1997) afirma que, si se puede identificar una relación de identidad semántica entre los dos miembros enunciados, puede hablarse de reformulación parafrástica. La reformulación parafrástica puede ponerse en funcionamiento a través de marcadores discursivos reformulativos lingüísticos, a través de marcas tipográficas o a través de procedimientos metalingüísticos, por lo que es pertinente resaltar que el uso del marcador discursivo no es prescriptivo ni obligatorio en la reformulación parafrástica. Las siguientes tablas reflejan los procedimientos que permiten identificar la reformulación parafrástica atendiendo a dos criterios: a la función de la operación reformulativa (Tabla 27) y al procedimiento que introduce la operación reformulativa (Tabla 28).

REFORMULACIÓN PARAFRÁSTICA	MARCADORES
Expresiones metalingüísticas y marcadores de reformulación	Ser, denominar, llamar, significar, recibir el nombre de, referirse a, aludir, decir, querer decir, conocer como, o, es decir...
Marcas tipográficas	Coma, dos puntos, guion [raya], paréntesis...

Tabla 27: Marcadores de reformulación parafrástica según el procedimiento que introduce la operación reformulativa (Bach *et al.*, 2003; en Rodríguez-Tapia, 2015: 7)

REFORMULACIÓN PARAFRÁSTICA	MARCADORES
Operaciones de reformulación explicativa	es decir, esto es, lo que es lo mismo, dicho de otro modo/de otra manera, (dicho) en otras palabras/en otros términos, vale decir, o sea, o, i.e.

optamos por emplear de forma exclusiva las denominaciones de Gülich y Kotschi (i. e. reformulación parafrástica y no parafrástica) y considerar que la terminología de Fuchs hace referencia a conceptos hiponímicos de aquellos.

**Operaciones de reformulación
rectificativa** (o) mejor dicho, (o) más bien, digo

Tabla 28: Marcadores de reformulación parafrástica según la función reformulativa (García Negroni, 2009: 51)

Siguiendo dichos criterios de carácter semántico, se distinguen tres tipos de reformulación parafrástica (Gülich y Kotschi, 1987; Garcés, 2008: 73)¹⁰⁷:

a) **Expansión**: el segmento reformulador es más amplio semánticamente que el enunciado de referencia. Se subdivide en **explicación** y **precisión**¹⁰⁸.

(26) En estos casos el miocardio inicialmente es *normal*, esto es, **no existe un problema intrínseco de contractilidad**¹⁰⁹.

b) **Reducción**: el segmento reformulador condensa el contenido semántico del enunciado de referencia. Se subdivide en **denominación** y **resumen**¹¹⁰.

(27) *Los medicamentos más utilizados en la actualidad para controlar el colesterol son las denominadas estatinas*¹¹¹.

c) **Variación**¹¹²: ambos segmentos presentan rasgos semánticos equiparables.

(28) *Digital* (o **digoxina**): medicamento que aumenta las contracciones del corazón para aliviar los síntomas de insuficiencia cardíaca¹¹³.

(29) *La insuficiencia cardíaca diastólica (también denominada «insuficiencia cardíaca con función sistólica conservada»)* [...] ¹¹⁴.

10.1.2. La reformulación no parafrástica

En la reformulación no parafrástica (Gülich y Kotschi, 1995: 44; Bach *et al.*, 2003: 2) se produce un cambio de perspectiva enunciativa en la nueva formulación, que queda sujeta a una relación de subordinación jerárquica con respecto del enunciado original de referencia (Garcés, 2008: 72). Para Rossari (1990, 1997), si el marcador reformulativo

¹⁰⁷ En los ejemplos entre paréntesis, el segmento de referencia aparece marcado con *cursiva*. El segmento reformulador aparece marcado con **negrita**. El marcador, en caso de que exista, aparece marcado con subrayado.

¹⁰⁸ La explicación aporta información adicional sobre el enunciado reformulado. La precisión matiza un aspecto concreto del enunciado reformulado.

¹⁰⁹ Fuente: Cardiofamilia (s. a.). «Concepto e importancia de la insuficiencia cardíaca», Cardiofamilia. Disponible en: <http://www.cardiofamilia.org/etiologia-y-fisiopatologia/concepto-e-importancia.html>.

¹¹⁰ La denominación da nombre a la información contenida en un enunciado (normalmente consiste en reducir un texto a un segmento léxico o fraseológico), mientras que el resumen sintetiza la información contenida en un enunciado (normalmente consiste en elaborar un texto menor que el texto origen).

¹¹¹ Fuente: Junta de Castilla y León (s. a.). «Insuficiencia Cardíaca», Aula de Pacientes de Castilla y León. Disponible en: <http://www.saludcastillayleon.es/AulaPacientes/es/guias-aula/insuficiencia-cardiaca/guia-informativa-insuficiencia-cardiaca.ficheros/162150-Gu%C3%ADa%20Insuficiencia%20Cardiaca.pdf>.

¹¹² Bach *et al.* (2003) denominan a este procedimiento *redenominación*, que, desde nuestro punto de vista, es análogo a la variación denominativa (Freixa, 2005).

¹¹³ Fuente: <http://www.insuficienciacardiaca.com/insuficiencia-cardiaca/index.htm>.

¹¹⁴ Fuente: Cardiofamilia (s. a.). «Concepto e importancia de la insuficiencia cardíaca», Cardiofamilia. Disponible en: <http://www.cardiofamilia.org/etiologia-y-fisiopatologia/concepto-e-importancia.html>.

hace que se produzca una reinterpretación del segmento reformulado según una perspectiva enunciativa nueva, se trata de un caso de reformulación no parafrástica. En la reformulación no parafrástica, el empleo del marcador de reformulación (como *en suma*, *en todo caso*, *en conclusión*, etc.). De hecho, su supresión acaba con la operación de reformulación (García Negroni, 2009: 49).

De acuerdo con el criterio pragmático (es decir, según el grado de cambio de perspectiva entre miembro de referencia y miembro reformulado), se distinguen tres grados de reformulación no parafrástica (de menor a mayor grado): la recapitulación, la reconsideración y la separación o distanciamiento, así como una categoría de invalidación.

García Negroni (2009: 51) distingue los siguientes marcadores de reformulación no parafrástica para las siguientes operaciones:

REFORMULACIÓN NO PARAFRÁSTICA	MARCADORES
Operaciones de recapitulación o condensación	a modo de conclusión, en síntesis, en resumen, en conclusión, en suma, recapitulando, en definitiva...
Operaciones de reconsideración	después de todo, al fin de cuentas, al fin y al cabo, de hecho, en efecto...
Operaciones de distanciamiento	de todas formas/maneras/modos, de cualquier forma/modo/manera, en todo caso, al menos, en realidad, en el fondo...
Operaciones de invalidación	en fin, (bueno) no...

Tabla 29: Marcadores de reformulación no parafrástica según la función reformulativa (García Negroni, 2009: 51)

10.1.3. Los constituyentes de la reformulación

El funcionamiento de la reformulación está condicionado por ciertos aspectos que afectan a su estructura básica, a la (no) obligatoriedad del empleo del marcador de reformulación, al segmento sobre el que incide la reformulación o al carácter explícito o implícito del segmento de referencia (Garcés, 2008: 76-79). A continuación, exponemos los dos primeros aspectos, que serán los que se tendrán en cuenta en el análisis de nuestra investigación:

- a) La **estructura básica** de los procedimientos de reformulación cuenta con un segmento de referencia, un posible conector y un segmento reformulador.

(30) Digital (o digoxina): medicamento que aumenta las contracciones del corazón para aliviar los síntomas de insuficiencia cardíaca¹¹³.

- b) En cuanto a la **presencia del marcador**: cuando existe grado de equivalencia semántica más o menos explícita, existe la posibilidad de introducir o no un marcador reformulador.

- (31) Cuando llega poca cantidad de sangre a los tejidos, los nutrientes y el oxígeno disminuyen por lo que se produce fatiga o debilidad muscular¹¹⁵.
- (32) La Hipertensión arterial (HTA), el tabaquismo y la hipercolesterolemia son los factores de riesgos más importantes para sufrir un evento cardiovascular mayor (Furones, 2003)¹¹⁶.
- (33) Ortopnea: se llama así a la sensación de falta de aire cuando se está tumbado en la cama (Saceda, s.a.)¹¹⁷.

10.2. La reformulación en el texto especializado¹¹⁸

De acuerdo con Bach (2001: 217)¹¹⁹, la reformulación de los textos científicos puede tratarse de dos formas distintas:

- a) La producción de un nuevo texto divulgador a partir de otro más especializado. Por ejemplo, cuando se adapta la información de un artículo científico sobre el funcionamiento del cáncer a una revista de divulgación científica.
- b) La recuperación dentro de un mismo texto especializado de algún elemento que el productor formula de nuevo. Por ejemplo, cuando se proporciona la sigla entre paréntesis para una denominación.

Ambos sentidos se relacionan con la noción de traducción interlingüística o *rewording* de Jakobson (1975) y prestan atención al aspecto del grado de especialización textual. En otro trabajo, Bach *et al.* (2003: 2) profundizan en la segunda forma y añaden una tercera, relacionada con la situación en que un marcador de reformulación parafrásico vincula una unidad terminológica con una unidad más transparente o con otra unidad terminológica y concluyen que esta siempre es posterior a la formulación (concepción unidireccional de la reformulación) (Rodríguez-Tapia, 2015: 6).

¹¹⁵ Fuente: <http://www.elmundo.es/elmundosalud/especiales/corazon/b.html>

¹¹⁶ Fuente: FURONES MOURELLE. Juan Antonio (2003): «¿Los diuréticos tiziados y los beta bloqueadores siguen como primera elección en el tratamiento de la hipertensión arterial?», *Biomed. Revista electrónica de Biomedicina*, 1(1), pp. 43-48.

¹¹⁷ Fuente: SACEDA, David (s.a.). «Insuficiencia cardíaca», Webconsultas. Disponible en: <http://www.webconsultas.com/salud-al-dia/insuficiencia-cardiaca/insuficiencia-cardiaca-11552>.

¹¹⁸ Las dos siguientes secciones han sido extraídas y adaptadas de Rodríguez-Tapia (2015: 4-6).

¹¹⁹ El trabajo de Bach, fruto de propuestas teóricas como la de la Teoría de la Argumentación y del modelo integrador de Ginebra, se centra en los conectores reformulativos variacionales y no variacionales en catalán, que nos ayudará a prever algunos de los comportamientos de estos conectores reformuladores en los textos semiespecializados en español.

10.3. Los procedimientos de reformulación formal en el texto especializado

Los procedimientos mediante los cuales se produce variación conceptual o variación en la representación formal de los términos reciben el nombre de procedimientos de reformulación formal y conceptual (PRFC).

La finalidad de estos procedimientos es aportar información relevante al proceso comunicativo para conseguir que el destinatario comprenda el contenido del mensaje que se transmite (Rodríguez-Tapia, 20015: 4-5).

Los PRFC pueden identificarse a través del análisis textual, resumido de la siguiente forma: en primer lugar, se usa un término; en segundo lugar, se identifica un problema de comprensión por parte del destinatario o una tradición discursiva que afecta al término; y en última instancia se emplean las estrategias de reformulación. El proceso anteriormente enumerado pretende determinar si, al redactar un texto, existe un potencial problema de comprensión por parte del destinatario.

Tradicionalmente, el criterio que condiciona la activación de la reformulación es semántico y cognitivo¹²⁰, puesto que el fin del PRFC es garantizar la comprensión del término mediante el uso de unidades o estructuras equivalentes conceptualmente¹²¹. No obstante, en este proceso también puede producirse variación conceptual, que depende de la funcionalidad de la clase textual (Ciapuscio, 1999: 103).

Este problema de comprensión depende por completo de las condiciones discursivas y pragmáticas, que muestran, en rasgos generales, si el destinatario tiene acceso al conocimiento especializado que codifican los términos. Es decir, el emisor del texto determina si, por ejemplo, la unidad *insuficiencia cardíaca* es un problema cognitivo para el destinatario, dependiendo de si este es un especialista, un instruido o un lego en la materia y ejecuta los procedimientos pertinentes para que el destinatario pueda acceder al conocimiento especializado codificado, por ejemplo, mediante el uso de ejemplos, de definiciones o de metáforas (Rodríguez-Tapia, 2015: 5).

¹²⁰ Este criterio cognitivo se manifiesta en el marco de constante evolución que defiende la Teoría Sociocognitiva de la Terminología (Temmerman, 2000 y 2002).

¹²¹ Según Arntz y Picht (1995: 191), dos términos se consideran equivalentes si existe entre ellos identidad conceptual, lo que Felber (1984: 153), desde el punto de vista interlingüístico, percibe en términos de gradación, y diferencia la equivalencia conceptual plena, la intersección, la inclusión o la inexistencia de equivalencia conceptual. Si no existe equivalencia conceptual (inter o intralingüística), Goffin (1985) propone diversas soluciones, entre las que se hallan el préstamo o calco léxico, la creación denominativa o la creación de un equivalente explicativo. Las soluciones que propone Goffin forman parte del catálogo de PRFC que analizaremos más adelante (Rodríguez-Tapia, 20015: 5).

Ahora bien, a nuestro juicio, el criterio cognitivo tradicional no es el único condicionante para la activación de los PRFC, ya que su uso también puede tener una motivación discursiva o estilística relacionada, bien con el tipo de texto y su tradición discursiva, bien con las intenciones retóricas del usuario. Por ejemplo, la variación denominativa, que incorporamos a nuestro modelo de estudio, es un tipo de PRFC y Freixa demostró en su trabajo de 2005 que dicha variación estaba motivada por causas estilísticas principalmente.

10.4. Nuestra postura teórica y metodológica

En Rodríguez-Tapia (2015: 14; 2016b: 233) exponíamos una herramienta de análisis de procedimientos de reformulación formal y conceptual según el siguiente esquema¹²²:

Procedimientos de reformulación formal y conceptual			
Procedimientos de reformulación conceptual	Procedimientos de reformulación formal		
	Formas lingüísticas		Formas no lingüísticas
	Textuales	VD	
Determinologización			Gráficos
Banalización			Tablas
Metáfora	Definición	UT diferente	Ilustraciones
	Reformulación parafrástica	→ Eponimia	
	→ Referencias intratextuales y extratextuales	Formas fraseológicas	Fórmulas alfanuméricas por convención
	→ Marcadores de reformulación parafrástica	Símbolos	
	Reformulación imitativa	Siglas, acrónimos y abreviaturas	
		Neología y préstamos	
		Hiperonimia	
		→ Ejemplificación	

Tabla 30: Procedimientos de reformulación formal y conceptual según Rodríguez-Tapia (2015: 14; 2016b: 233)

En esta tesis doctoral, optaremos por modificar dicho esquema para ajustarnos a nuevos objetivos, de forma que nos centraremos en exclusiva en los procedimientos de reformulación formal de formas lingüísticas y, en concreto, mediante procedimientos de reformulación parafrástica (PRF). Las razones que justifiquen esta modificación están

¹²² La justificación teórica de cada sección y procedimiento se desarrolla en Rodríguez-Tapia (2015: 12-16).

relacionadas con el perfil lexicalista-discursivista de nuestra tesis doctoral, así como con dos de los objetivos de la misma (página 19):

- a) Identificar la relevancia de cada uno de los parámetros que contribuyen a la distinción del grado de especialización.
- b) Evaluar los parámetros que intervienen en la identificación del texto semiespecializado y que resultan de posible aplicación tanto en la configuración del texto especializado como en la del no especializado.

A nuestro juicio, la reformulación desempeña un papel fundamental en la configuración semántico-cognitiva del texto y es dependiente de las condiciones pragmáticas, lo que puede repercutir en la clasificación de los textos según su grado de especialización.

Así, la herramienta que emplearemos en este trabajo sigue la siguiente estructura según los criterios que mencionamos a continuación:

- a) Una estructuración según la **función** que cumple el procedimiento de reformulación (expansión, reducción o variación), según Gülich y Kotschi (1987).
- b) Un conjunto de **procedimientos de reformulación** por cada función, a saber:
 - (i) Para la **expansión**: el procedimiento de **definición**, según Arntz y Picht (1995), Bessé (1997) y Monterde Rey (2004), al que se asimilan explicación y precisión según Gülich y Kotschi (1987).
 - (ii) Para la **reducción**: el procedimiento de **denominación**, explicitado a través del uso de a) **siglas y acrónimos** según Monterde Rey (2004); b) **unidades terminológicas diferentes** según Freixa (2002a, 2002b y 2005); c) **neologismos y préstamos** según Goffin (1985), Cabré (1993) y Guerrero Ramos (1995); y d) la **ejemplificación** según Campos (2013).
 - (iii) Para la **variación**: el procedimiento de variación denominativa, explicitado a través de los mismos procedimientos que la reducción, a excepción del procedimiento de ejemplificación.
- c) Un conjunto de **marcadores de reformulación**, con carácter metalingüístico y tipográfico, para cada función, según Bach *et al.* (2003) y García Negroni (2009: 51).

A continuación, sintetizamos dicha herramienta en forma de tabla acompañada de algunos de los marcadores de reformulación a través de los cuales se puede identificar dicha estrategia reformulativa.

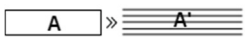


PROCEDIMIENTOS DE REFORMULACIÓN PARAFRÁSTICA			
Función	Expansión 	Reducción 	Variación 
Procedimientos	Definición Explicación Precisión Ejemplificación	Denominación Siglas y acrónimos UT diferente (incluidos epónimos y unidades fraseológicas) Neologismos (incluidos préstamos) Ejemplificación	Variación denominativa ¹²³ Siglas y acrónimos UT diferente (incluidos epónimos y unidades fraseológicas) Neologismos (incluidos préstamos)
Marcadores metaling.	ser, denominar(se), llamar(se), significar, recibir el nombre de, referirse a, aludir, decir, querer decir, conocer como, es decir, esto es, también, lo que es lo mismo, dicho de otro modo/de otra manera, (dicho) en otras palabras/en otros términos, o sea, o, i.e., etc.		conocer como, es decir, esto es, también, lo que es lo mismo, dicho de otro modo/de otra manera, (dicho) en otras palabras/en otros términos, o sea, o, i.e., etc.
Mecanismos tipográfic.	Coma, dos puntos, raya, paréntesis...		

Tabla 31: Herramienta de análisis de procedimientos de reformulación parafrástica [elaboración propia]

¹²³ Recordemos que la variación denominativa es un fenómeno de sinonimia.

III. FUNDAMENTOS METODOLÓGICOS

1. LA LINGÜÍSTICA DE CORPUS COMO INSTRUMENTO DE TRABAJO

El interés investigador por demostrar y relacionar hechos lingüísticos a partir de pruebas empíricas derivó en el surgimiento de la lingüística de corpus. Sus ventajas investigadoras rompieron con la hegemonía generativista que había surgido a finales de los años cincuenta y los estudios de la lingüística pasaron de estudiar las posibilidades teóricas de un hablante-oyente ideal a estudiar el uso natural de la lengua en textos reales (Biber, Conrad y Reppen, 1998: 1; en Torruella, 2017: 31 y 32).

Como señala Torruella (2017: 23), es posible distinguir dos acepciones para el término *lingüística de corpus*:

por una parte, la que hace referencia a la metodología de diseño y recopilación de un corpus y de su plataforma de consulta y, por otra, a la de su explotación (tratamiento de los datos, sus análisis cuantitativos y estadísticos, etc.). El término básicamente define un proceso de almacenamiento y recuperación de datos, y una manera de considerar estos datos como una fuente de informaciones para comprobar empíricamente viejas teorías o como un banco de pruebas para establecer nuevas.

En nuestra investigación, la explotación del corpus es nuestra herramienta principal de interpretación de los datos:

en su segunda acepción, la referida a la explotación de corpus, lingüística de corpus equivale, tal como apunta Rojo (2015: 681), a la «lingüística basada en el análisis de corpus», que es una aproximación al análisis de los fenómenos lingüísticos tal como se presentan en un corpus textual o en diversos subcorpus textuales (Torruella, 2017: 23 y 24).

Esta interpretación de datos es la que permite defender resultados científicos a partir de pruebas empíricas que van más allá de las hipótesis con tintes más intuitivos, es decir, permite llevar a cabo «qualitative interpretations of quantitative patterns» (Biber, Conrad y Reppen, 1998: 5). La lingüística de corpus, por tanto, no se limita al recuento de elementos o características, sino a validar hipótesis sobre análisis a partir de material textual real. De acuerdo con Leech (1992: 112-113), la lingüística de corpus permite incorporar falsabilidad, exhaustividad, simplicidad, intensidad y objetividad a los estudios científicos en lingüística. Precisamente una de las ventajas más productivas de la lingüística de corpus es la «posibilidad de construcción de modelos teóricos» (Blecua, 1996: 15), puesto que su método de trabajo permite relacionar datos y teoría: este es nuestro interés principal en esta tesis doctoral.

1.1. Objeto de la lingüística de corpus

La lingüística de corpus recibe este nombre precisamente por su objeto de análisis (y, a la vez, la vía de donde extraer conclusiones): el corpus. La Expert Advisory Group on Linguistic Engineering Standards (EAGLES; 1996: 23), basándose en la definición de Sinclair, define el corpus como «a collection of pieces of language that are selected and ordered according to explicit linguistic criteria in order to be used as a sample of the language». Para Torruella (2017: 33 y 34), una definición más precisa y cercana al concepto actual de corpus textual es el que proporciona Rojo (2014: 371) en *The Routledge Handbook of Hispanic Applied Linguistics*:

[A corpus is] a set of natural texts (or pieces of texts), stored in electronic form, assumed to be jointly representative of a linguistic variety in some of its components, or in all of them, and grouped together so that they can be scientifically studied.

Ahora bien, un archivo informatizado o una biblioteca de textos electrónicos, que suelen contener una recopilación de textos en soporte informático, no constituye *per se* un corpus. De hecho, el objeto de la lingüística de corpus es un corpus informatizado, es decir, una «[r]ecopilación de textos en soporte informático, unificados en un formato estándar y seleccionados siguiendo unos criterios de equilibrio y representatividad según una final[i]dad específica» (Torruella, 2017: 35). Es decir, un corpus necesita haber sido construido siguiendo una relación de criterios que se ajusten a una finalidad investigadora concreta. En otras palabras: el objetivo investigador restringe la construcción del corpus, es decir, las muestras de texto que son seleccionadas sobre el total de la población. Esto deriva en reconocer los límites del objeto, puesto que, siguiendo a Parodi (2010: 24): «un corpus es sólo una colección finita de un universo infinito».

Este hecho permite debatir la representatividad de los corpus para estudiar un fenómeno lingüístico (*vid. infra*. III. 3.4., p. 186). La condición óptima para realizar estudios sobre corpus sería que el corpus contuviese toda la producción lingüística de una lengua. Sin embargo, esto no es posible ni lingüística ni técnicamente, por lo que «se debe confiar en el carácter de representatividad del corpus, ya sea este de la lengua en general, ya sea de un estado o nivel de lengua determinado» (Torruella, 2017: 36). De hecho, la propia definición de Rojo (*vid. supra*) advierte la condición de que el corpus es un conjunto de textos «assumed to be jointly representative of a linguistic variety».

1.2. Fundamentos teóricos de la lingüística de corpus

Otra de las ventajas del corpus se halla en que su utilidad como instrumento del que extraer pruebas lingüísticas se aplica a cualquier teoría, por lo que «no es necesariamente incompatible con ningún tipo de teoría» (Parodi, 2008: 94-95). Este carácter instrumental o, más bien, metodológico, es el defendido por multitud de investigadores como Leech (1992), Biber (1993), Stubbs (1996), Tognini-Bonelli (2001) o McEnery y Hardie (2012) y es el que adoptamos en este estudio. No obstante, también existen posturas de carácter más cognitivo, como la de Caravedo (2005: 96), que argumenta que la lingüística *del corpus* es «más bien un tipo de actitud epistemológica respecto de la relación entre teoría y realidad, independiente de una metodología o de un modelo teórico particular» (Torruella, 2017: 26). Torruella sintetiza las posturas que niegan el carácter metodológico de la lingüística de corpus afirmando que la tratan «como una teoría que pretende explicar el funcionamiento de la mente a partir de la concepción del lenguaje humano como una facultad probabilística o como un fenómeno estadístico» (Torruella, 2017: 26).

Centrándonos en su concepción como método, es posible identificar un cambio en el modo de trabajo con respecto al método descriptivo tradicional, que, según Rojo (2012: 436-437), se deja guiar por la sobrerrepresentación de «lo infrecuente, lo inesperado, lo llamativo» de un fenómeno a partir de «un conjunto reducido de textos o bien mediante una selección realizada con criterios variables de un conjunto más amplio de textos». No obstante, la lingüística descriptiva tradicional también sigue la observación de *corpus*. Usamos esta tipografía puesto que este corpus (o, mejor, conjunto o recopilación de textos) suele construirse a partir de los intereses y las intuiciones del investigador, que deja a un lado la exhaustividad que requieren los criterios de representatividad y equilibrio de las muestras, que son las que guían los procedimientos del método de compilación de la lingüística de corpus, y se centra en la selección de un número reducido de textos (frecuentemente de carácter literario). Tampoco la explotación de los datos que realiza el método descriptivo tradicional sigue técnicas estadísticas ni análisis cuantitativos o comparativos.

Siguiendo a Leech (1992: 107), la lingüística de corpus contrasta con el método propugnado por el enfoque generativo, puesto que pretende prestar atención a la actuación (en lugar de a la competencia), a la descripción (en lugar de a los universales lingüísticos), a los modelos cuantitativo-cualitativos y al carácter empírico (en lugar de al racionalista).

Así, los fundamentos de la lingüística de corpus quedan sintetizados en las palabras de Torruella (2017: 25):

La investigación que se hace siguiendo los métodos establecidos por la lingüística de corpus es de carácter inductivo, ya que pretende que, a partir de los análisis de los ejemplos reales que aparecen en los corpus, se generalicen los resultados, al contrario de lo que sucede siguiendo la metodología tradicional que trata de ilustrar con ejemplos hallados en los corpus el conocimiento que los usuarios ya tienen de la lengua. Así pues, la lingüística de corpus aspira a facilitar el estudio de la lengua basándose en ejemplos reales organizados en corpus y tiene un marcado enfoque empírico que pretende transformar los datos observables en evidencias científicas destinadas, más que a ilustrar con ejemplos hipótesis ya probadas, a comprobar las nuevas hipótesis que los investigadores proponen.

Ahora bien, esto no quiere decir que la función de un corpus sea exclusivamente inductiva, puesto que también pueden realizarse aproximaciones deductivas usando los corpus.

2. LA PERCEPCIÓN SOCIOCOGNITIVA DEL TÉRMINO

Desde el punto de vista del usuario no especialista, responder a la pregunta «¿qué es un término?» se vuelve una tarea cada vez más compleja cuando el grado de especialización o el tipo de emisor no está claro, puesto que el valor semántico-pragmático especializado no es discriminable. Uno de los criterios a los que queda acudir para responder a dicha pregunta es la percepción sociocognitiva del destinatario como medio de discriminación del valor especializado. En relación con esta percepción es conveniente revisar algunos conceptos como los de variación lingüística, cognición, valoración o actitud.

2.1. Fundamentos de variación lingüística

Si partimos de la consideración de la naturaleza social del lenguaje laboviano, debe defenderse la lengua como un ente necesariamente social y colectivo. De hecho, Labov (1972) se resistió durante muchos años a aceptar el término *sociolingüística*, puesto que consideraba inexistente una lengua no social y, por tanto, una lingüística no social. Labov se inclinaba por posicionarse más bien como un variacionista o un lingüista dedicado a la variación y al cambio (Wmffre, 2013: 82).

Precisamente la variación lingüística, en concreto su percepción, representa nuestro interés en este capítulo. La variación tiene su origen en el desajuste de la relación biunívoca entre significante y significado por el que un significado puede expresarse a través de más de un significante, siendo la equivalencia semántica el principio fundamental para identificar las variantes (Caravedo, 2014: 22-23). La variación constituye el paso previo necesario al cambio lingüístico, aunque este último no es resultado obligatorio de todo fenómeno de variación (ibíd.: 26). Por ejemplo, la palabra *bárbaro*, que deriva del latín y esta, a su vez, del griego, designaba en el período helenístico a un individuo extranjero. Durante la época medieval, la variación del significado derivó en un cambio de significado, siendo *bárbaro* la palabra que designaba la falta de educación o tosquedad.

Uno de los puntos débiles del modelo teórico sociolingüístico laboviano según Caravedo (2014: 30) se halla en la no contaminación del hecho lingüístico respecto del mundo externo, lo que implica que la variación se produce exclusivamente en el interior de la unidad:

Las diferencias de orden social entre las variantes, aun cuando supongan valores distintos para los hablantes, no entrarían en juego en la delimitación de variables, ya que se supone que no llegan a intervenir en su carácter simbólico ni a modificar el plano referencial/representativo, que es el tradicionalmente considerado propio de la esencia del lenguaje.

A esta crítica se suma la atención que han prestado los estudios sociolingüísticos canónicos sobre variación a la producción de los hablantes desde una dimensión grupal, dejando de lado el plano subjetivo (ibíd.: 36).

La percepción, que se halla en este plano subjetivo, estudiaría las valoraciones y actitudes de los hablantes respecto de los fenómenos percibidos, lo que forma parte de la cognición de los hablantes. Esta dimensión cognoscitiva se explicita mediante los siguientes enunciados (ibíd.: 36, 37, 46):

- a) La medición de probabilidades de los fenómenos forma parte de la competencia del hablante, y, por tanto, de su cognición.
- b) Las variantes y alternativas son percibidas desde el punto de vista del hablante y forman parte del conocimiento que el hablante tiene de su lengua.
- c) Los factores cognitivos permiten adquirir las unidades con función referencial, así como las propias variables sociales.

Por el protagonismo de lo cognitivo en la propuesta del modelo de percepción, creemos conveniente puntualizar brevemente algunos conceptos estrechamente relacionados.

2.2. La cognición lingüística

La cognición lingüística engloba tanto el conocimiento de la lengua con el que cuenta un hablante como el proceso que ejecuta para conocer y usar la lengua y, además, «incluye los contenidos derivados de la aproximación descriptiva de la lengua con fines científicos» (Caravedo, 2014: 47).

La cognición como pilar de la percepción del hablante permite defender que dicha percepción no tiene por qué identificarse de forma necesaria con la percepción científica. Dicho de otra forma, los criterios de segmentalidad de las unidades usados por los estudios científicos no son los mismos criterios usados por el hablante a la hora de categorizar la realidad. En efecto, «categorizar presupone percibir» (ibíd.: 35), lo que obliga a poner el foco de atención en la percepción y producción del hablante en contraste con los criterios de la percepción científica. Es decir, estos fundamentos teóricos nos deben servir para dar respuesta al criterio sociocognitivo necesario en el principio de activación del valor de especialización: ¿qué considera el hablante que es un término? Las actitudes lingüísticas de un hablante forman parte de su cognición, es decir, de su forma de entender el mundo, por lo que el problema de la percepción resulta de gran relevancia en la dimensión sociocognitiva del estudio discursivo.

Caravedo (1990, 1991) desarrolló el concepto de *espacio de variabilidad* para el plano sonoro, que es fácilmente adaptable a nuestros objetivos:

Desde el punto de vista del receptor, lo que este percibe no son los campos mismos, pues se trata de continuidades permisibles que el hablante no puede segmentar auditivamente porque no dirige su atención hacia ellas. Es la transgresión o la invasión de un elemento no contemplado dentro de los límites de los espacios de variabilidad, lo que captaría la atención del hablante (Caravedo, 2014: 36).

En otras palabras, el hablante capta los elementos no normales, en el sentido de permitidos o frecuentes por el espacio de variabilidad en el que se mueve.

De acuerdo con Jakobson (ibíd.: 48), la cognición funcionaría mediante la abstracción de un objeto a través de la selección y generalización de sus rasgos esenciales (i. e., los rasgos invariables), que son separados de los rasgos variables. Ciertos aspectos de

la lengua no formarían parte del sistema cognitivo primario del hablante hasta que se produjese un aumento de estímulos sociales o de su instrucción académica. Entre estos aspectos se encontrarían la variación diafásico-diastrática de los discursos especializados y su *continuum*, cuyos elementos adquiere el hablante tras su inserción y experiencia lingüística en sociedad y no por el contacto lingüístico con la lengua materna en las primeras fases de la adquisición lingüística.

2.3. Variación no funcional vs. variación funcional y variación intraindividual vs. variación extraindividual

Caravedo (2014: 28-29) distingue dos tipos de variación atendiendo a la posibilidad de generar cambio lingüístico:

- a) La **variación no funcional**: no produce cambio de significado; por ejemplo, la variación que existe en algunos sufijos diminutivos como *-ito* o *-illo* en español en el plano morfológico.
- b) La **variación funcional**: incluye modificaciones en el orden del significado primitivo de la función referencial y constituye la fase precedente del cambio lingüístico. Esta variación funcional puede tener dos efectos principalmente:
 - a) la ampliación o creación de nuevos significados, como la polisemia en el plano léxico; y b) la reducción de significados primitivos, como la neutralización o las indistinciones (por ejemplo, los casos de *-ra-* y *-se-* en el subjuntivo verbal del español actual).

La activación del valor especializado por parte de una unidad léxica estaría incluida dentro de esta variación funcional y, de alguna forma, sujeta a la percepción del hablante. De hecho, la percepción del hablante no coincide con las fronteras de la variación no funcional, que son establecidas de forma científica; es decir, un especialista puede considerar variación funcional lo que un lego determina como no funcional, puesto que no es consciente del cambio de significado. Por esta razón, Caravedo (2014: 44, 54-55) considera conveniente distinguir entre dos tipos de variación relacionadas con la cognición de una lengua, en la que se incluyen tanto los fenómenos variables como invariables:

- a) La **variación intraindividual**: de corte cognitivo y sujeta al hablante individual. En la variación intraindividual, una entidad de la lengua es variable si el hablante reconoce las posibilidades de actualización de dicha entidad lingüística. En otras

palabras, el hablante percibirá dicha entidad como variable si conoce las variantes o alternativas en las que se actualiza. Por ejemplo, en el caso de las variantes diatópicas *carro* y *coche*, un hablante del español peninsular, desde su cognición, consideraría *coche* como invariante en ciertos registros de la lengua, puesto que solo es intercambiable con *carro* en expresiones cotidianas o de tono coloquial como: «vaya *carro* te has comprado».

- b) La **variación extraindividual**: pertenece a la lengua histórica, pero no tiene su asiento en el individuo o grupo de individuos. En el caso de *carro* y *coche*, se consideran variantes extraindividuales, puesto que se trata de dos posibilidades diferentes de hacer referencia a la misma realidad. Se constituyen como formas de variación de la lengua independientemente de si carecen o no de valor cognitivo individual y son reconocidas por el hablante como variantes o no variantes.

2.4. Conceptos relacionados: captación, atención, fijación, subjetividad, valoración y actitud

La adquisición lingüística es fruto de un conjunto de estímulos lingüísticos sujetos al proceso de cognición, que regula, a su vez, procesos como los de captación, fijación y reproducción (creativa en el sentido generativo) de una variedad de la lengua, entendida como un patrón lingüístico de una comunidad. Cumplido el período de adquisición lingüística, el hablante no interrumpe el proceso cognitivo, sino que se sigue valiendo de él para seguir procesando nuevos estímulos lingüísticos.

Una vez iniciada la etapa de formación académica, el hablante puede llegar a percibir las diferencias entre la variación intraindividual y la extraindividual:

Generalmente, cuando el individuo no está expuesto directamente a una variedad distinta de la suya, la información proporcionada durante la escolaridad entra a formar parte de modo indirecto del conocimiento en general de su lengua. Tomando como base esta distinción, el niño que adquiere una variedad concreta de lengua (y desconoce las demás) percibirá lo variable como constante. Así, terminarán entrelazadas la invariación y la variación como si fueran una misma cosa (Caravedo, 2014: 57).

Así, prestaremos atención a continuación a conceptos que rigen el proceso cognitivo de la percepción y que se inician desde el primer contacto lingüístico.

En primer lugar, se produce la **captación** (*noticing* según Preston, 2013) del elemento lingüístico: el hablante identifica dicho elemento y lo asocia a un determinado hablante o grupo de hablantes. Esta captación activa un proceso de **atención** (*regard*) selectiva que se dirige a los fenómenos lingüísticos, cuyos elementos pertinentes acaban sufriendo el proceso de **fijación**, es decir, acaban siendo almacenados en la memoria (se convierten en *perceptos*; Caravedo, 2014: 76). Cuando el hablante se encuentra ante datos similares, es capaz de reconocerlos como relevantes y confrontarlos con las entidades seleccionadas anteriormente en el proceso de atención (ibíd.: 58-59).

Los procesos de captación, atención y fijación son, inevitable e intrínsecamente (al igual que el discurso), **subjetivos**¹²⁴, en tanto que forman parte de un fenómeno social (Searle, 1995). El discurso sería, por tanto, una entidad ontológicamente subjetiva (analizable también, no obstante, de forma epistémicamente objetiva), pues su existencia depende obligatoriamente de sus hablantes, por lo que su carácter convencional se halla vinculado a la subjetividad (Caravedo, 2014: 64). Es lógico pensar que la percepción, dependiente del hablante, es también inherentemente subjetiva, por lo que también lo son los perceptos o productos de la percepción. No obstante, es posible atisbar cierto convencionalismo y regulación en la identificación y parcelación de rasgos:

Sostener la subjetividad no significa minusvalorar o descartar las entidades mencionadas considerándolas como no susceptibles de conocimiento. Todo lo contrario: los hechos calificados de subjetivos o de relativos al sujeto pueden ser —o deben ser— abordados de modo objetivo; en los términos de Searle son pasibles de objetividad epistémica, y en este sentido constituyen materia del discurso científico [...]. Pero tal paradoja se disuelve cuando, identificados determinados rasgos subjetivos, se reconoce que tales rasgos están aceptados colectivamente y que presentan comportamiento regulares y motivados, esto es, que no constituyen producto del azar, de la imaginación o de la preferencia idiolectal del individuo (Caravedo, 2014: 125).

Las palabras de Caravedo justifican a la perfección nuestra intención de medir la percepción de varios grupos de informantes ante un corpus textual para identificar si existen patrones convencionales que superen las selecciones subjetivas (*vid. infra* III. 2.7., p. 179.).

A estos procesos hay que añadir la noción de *actitud lingüística*. La **actitud lingüística** de un hablante manifiesta la actitud psicosocial de un individuo respecto de una lengua o variedad de lengua (dialecto, sociolecto, fasolecto, etc.) y al uso social de esta

¹²⁴ La subjetividad puede entenderse como la relación de dependencia que existe entre un objeto respecto de un sujeto observador, mientras que, por contraste, la objetividad conlleva una relación de independencia del objeto respecto del observador (Searle, 1995).

(Moreno, 2008: 177-179). Dado que las marcas socioculturales de un grupo o comunidad se transmiten mediante la lengua, resulta complejo delimitar dónde comienza la actitud ante la variedad lingüística y dónde termina la actitud hacia el grupo social que emplea la variedad (ibíd.: 179). Siguiendo a Fasold (1996: 229 y ss.), es posible distinguir las actitudes lingüísticas desde dos puntos de vista:

- a) El **conductista**: considera la actitud como una conducta, es decir, como la respuesta ante un estímulo. Este estímulo, en este caso, sería la variedad de lengua o unas características sociolingüísticas determinadas.
- b) El **mentalista**: considera la actitud como un fenómeno de naturaleza psicosocial, como un estado mental hacia un hecho sociolingüístico determinado. Desde esta perspectiva, la actitud se encontraría a caballo entre el estímulo y la respuesta individual.

Evidentemente, un estado mental no puede ser observable de forma directa, por lo que es necesario recurrir a la información que pueda proporcionar el individuo consciente o inconscientemente (Moreno, 2008: 184). Para estudiar los datos desde una perspectiva mentalista, es posible usar tanto métodos indirectos como métodos directos, que incluyen el uso de cuestionarios o entrevistas (ibíd.: 185).

Las actitudes lingüísticas, de acuerdo con Moreno (2008: 181-182), implican la presencia de tres componentes: uno afectivo (la valoración), uno cognoscitivo (la creencia) y uno conativo (la conducta). El proceso se origina en el momento en el que se inicia la percepción de un mensaje, lo que pone en marcha una respuesta valorativa. Caravedo sintetiza a la perfección este hecho (la negrita es nuestra):

En una primera mirada, la **valoración** puede definirse como una expresión calificativa aplicada, sea de modo global a una lengua, a una variedad de ella o a un modo de hablar, sea de modo específico a un fenómeno en cualquiera de los rangos del sistema. Podría decirse que constituye una manifestación de la percepción, en la medida en que solo es posible valorar algo si primero ha sido percibido. No es ni siquiera imaginable una valoración sin que haya intervenido la percepción. En consecuencia, la percepción es la condición *sine qua non* de la valoración y constituye el instrumento por excelencia para pronunciarse, adoptar una actitud o una posición determinada respecto de lo observado (Caravedo, 2014: 95).

Entendida la valoración como una expresión calificativa (de cualquier objeto de la realidad; en nuestro caso, de un uso lingüístico; es decir, cumpliría una función metalingüística), es posible graduar dichas valoraciones entre lo negativo y lo positivo. La valoración, al igual que la percepción, es una actividad inherente al ser humano, que le

permite comparar, juzgar y contrastar hechos (también rasgos lingüísticos, por supuesto), por lo que va ligada a la lengua y, consecuentemente, debe tenerse en cuenta en los estudios sociocognitivos que impliquen la lengua.

Ahora bien, la relación que se establece entre valoración y percepción es doble, puesto que los valores establecidos por el individuo guían la percepción, y viceversa; lo percibido influye y reorienta los valores (Caravedo, 2014: 96). Cabe resaltar que la percepción y regulación de valores no se realiza de la misma forma durante la infancia o la edad adulta, precisamente porque en la edad adulta ya se cuenta con ciertos valores guiados por la percepción, lo que hace que el adulto no perciba de forma ingenua, sino que se concentra en los aspectos que ha aprendido a valorar (ídem).

A modo de ilustración, la percepción como instrumento de desarrollo de la cognición humana se podría esquematizar de la siguiente forma:

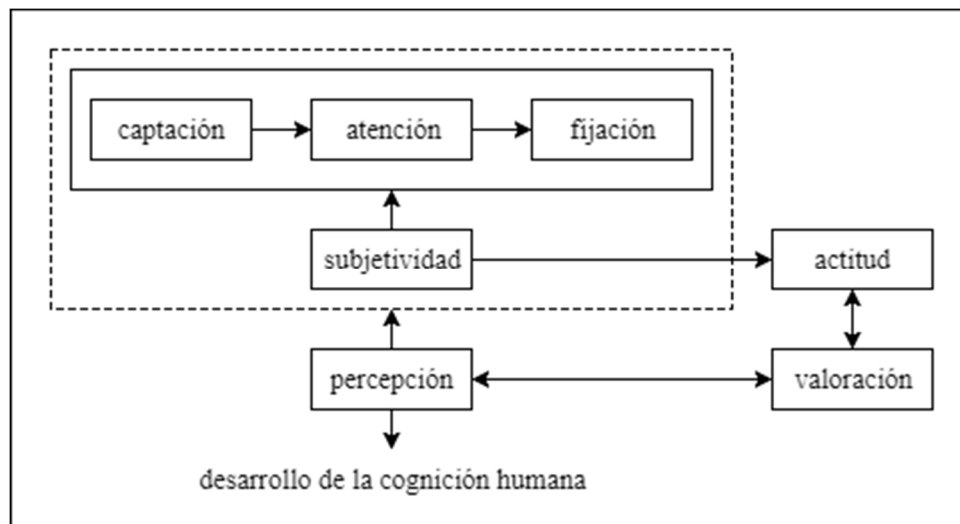


Ilustración 14: Esquema del funcionamiento de la percepción [elaboración propia]

2.5. Fases y funcionamiento de la percepción

Anteriormente hemos mencionado la interdependencia de la percepción y la valoración, pero las valoraciones que realizan los sujetos deben estar orientadas de alguna forma. Es decir, la relación entre percepción y valoración no es caótica ni azarosa, sino que está guiada por las primeras experiencias con agentes externos durante la fase de adquisición de la lengua, que llevan al interlocutor a reconocer elementos que sobresalen de un objeto complejo como es la realidad. Se trata, por tanto, de un fenómeno de orientación perceptiva no necesariamente premeditada (Caravedo, 2014: 105).

También avanzamos que la percepción no se desarrolla de la misma forma en la infancia y en la edad adulta, por lo que debe entenderse que existen diferentes fases cronológicas en la percepción (ibíd.: 80):

- a) Fase primaria o periodo adquisitivo, que se inicia con el nacimiento y acaba en la etapa preescolar.
- b) Fase secundaria o periodo de aprendizaje escolar.
- c) Fase terciaria o periodo estable.

Hasta el final de la fase secundaria el hablante adquiere los elementos esenciales de su variedad vernácula. Por otra parte, durante la fase terciaria, la percepción no modifica profundamente su producción, sino que sirve al hablante como recurso discriminatorio (analítico, según veremos posteriormente) para contrastar con otras variedades. Durante esta fase el hablante termina su formación escolar (entendida como nivel de enseñanza secundaria) e inicia su formación postobligatoria o profesional, lo que origina el contacto con otros grupos o actividades sociales.

Serán las variedades lingüísticas usadas por estas comunidades las que serán objeto de valoración y contraste con el vernáculo del hablante durante su interacción profesional y social cotidiana. Dicho de otra forma, al convivir el individuo con otras variedades lingüísticas (también otras lenguas) la percepción le permite marcar las diferencias principales entre variedades (fonéticas, léxicas, discursivas, etc.). El período estable recibe este nombre porque durante la etapa adulta los estímulos de lo desconocido normalmente se reducen y los procesos perceptivos *se estabilizan*.

Las diferencias entre variedades y elementos sobresalientes son objeto de atención tras la captación cognitiva. En este sentido, una característica de la percepción sería su carácter selectivo. El percepto, que constituye la representación de la realidad percibida y almacenada en la memoria tras el proceso de fijación, resulta, pues, fruto de esta selección de características: el hablante no construye una representación totalmente fiel a la realidad, sino que «enfoca unos elementos mientras que desenfoca otros» (Caravedo, 2009b: 25) mediante el mecanismo cognitivo de la atención y depende, como resulta lógico, del contacto social y las valoraciones subjetivas que se hayan llevado a cabo.

2.6. Tipos de percepciones y selección de información

Caravedo (2014: 107) justifica esta selección de información con las siguientes palabras:

Y aquí entra la dimensión individual del perceptor, que se fusiona con la colectiva. Existe una intervención intelectual del individuo que le permite seleccionar los rasgos salientes o pertinentes de lo observado. Ahora bien, como la realidad no se absorbe de modo absolutamente mimético (en el sentido sensorial) ni exhaustivo (en el sentido nemónico), la percepción individual de alguna manera deforma el objeto en cuestión tal como este se presenta en su materialidad o fisicidad transmitido por los mayores; lo parcela y hasta cierto punto lo falsifica, pero no de modo absolutamente arbitrario, sino dirigido o —como lo acabamos de proponer— orientado.

Este comportamiento nos hace preguntarnos qué información resulta relevante en el proceso de percepción y cómo se selecciona dicha información, puesto que el hablante individual, de alguna forma, se ve guiado por la dimensión social de la comunidad con la que tiene contacto. Así, respecto del objeto percibido, se distinguen tres percepciones (Caravedo, 2014: 108, 110, 111):

- a) La **percepción interna**: corresponde con el desarrollo de la identificación y observación de la variante local adquirida durante la fase primaria.
- b) La **percepción externa**: conlleva la captación de variedades ajenas durante las fases que involucran la socialización. Emplea como referencia la primera variedad adquirida (resultado de la percepción interna) como medio de comparación con las otras variedades.
- c) La **autopercepción**: permite la conciencia y evaluación de la propia variedad del individuo o de algunos de los fenómenos que se asignan a ella.

En cuanto a la segunda pregunta (¿cómo se percibe el objeto lingüístico?), Caravedo distingue dos recursos lógicos primarios que cumplen el proceso de percepción a través de la conversión de lo sensorial en nemónico y, tras esto, en conceptual (2003: 555-556; 2014: 113):

- a) La **percepción analítica**: permite aprehender rasgos discretos (fónicos, morfosintácticos, léxicos, etc.) del patrón variable o invariable de una lengua.
- b) La **percepción sintética**: permite distinguir modalidades lingüísticas de forma global sin individualizar entidades discretas.

La percepción analítica constituye una operación lógica en la que el individuo descompone un objeto complejo en cada una de sus partes mediante la operación de

abstracción. Esta descomposición se realiza siguiendo ciertos criterios de pertinencia, es decir, la mente del individuo identifica las partes más relevantes que integran el todo. Para realizar esto, uno de los criterios empleados es el criterio de frecuencia de *ocurrencia* de las entidades, que facilita la percepción, puesto que la repetición constituye un auxilio memorístico (Caravedo, 2014: 115). Con todo, el fenómeno contrario, es decir, el de *prominencia*, implica que el hablante pueda hiperpercibir de forma subjetiva elementos que, aunque no son frecuentes, cuentan con características sobresalientes, que el individuo es capaz de vincular a «determinados referentes, contextos, interlocutores o situaciones discursivas» (Caravedo, 2014: 116), es decir, los asocia a un determinado grupo o circunstancia comunicativa.

Con objeto de captar la recurrencia y la prominencia se ponen en juego dos operaciones cognitivas: la generalización y la singularización (Caravedo, 2003:556):

- a) La **generalización**: consiste en una operación que capta las entidades más frecuentes entre los enunciados recibidos por los hablantes. Hay que tener en cuenta que los hablantes no cuentan con un contador científico de probabilidades, sino que su percepción es intuitiva, y, por tanto, es falible con respecto a lo realmente recurrente.
- b) La **singularización**: es el proceso que actualiza el criterio de prominencia. Al discriminar fenómenos poco frecuentes, la singularización identifica rasgos relevantes o representativos en la entidad, que, a su vez, pueden ser representativos de un grupo o comunidad. El criterio para discriminar esta representatividad en el rasgo implica la subjetividad del hablante y «el sistema de valores transmitido generacionalmente, así como la capacidad de asociar metonímicamente un rasgo a un individuo, a un grupo, a un estilo (entendido en el sentido individual y social), a una situación, y asignarle algún tipo de valor» (Caravedo, 2014: 118). En el caso de nuestra tesis doctoral, si se piensa en un texto como una sentencia, un informe clínico o un informe técnico, es posible identificar ciertos elementos en los diferentes niveles de análisis lingüístico que se desvían del uso lingüístico cotidiano que podría haber en un sitio web, un periódico, un informativo, etc. Estos elementos singulares y prominentes podrían ser la mayor cantidad de terminología, las estructuras sintácticas más complejas o formas léxicas especiales (formas grecolatinas, sigladas o alfanuméricas), y son los que acaban siendo asociados a un acto comunicativo propio de especialistas.

La percepción sintética representa la capacidad para identificar un objeto de la realidad de forma global o integral, como un todo. Tiene lugar de forma complementaria a la percepción analítica y no necesariamente de forma aislada. En otras palabras, se refiere a que un hablante sea capaz de percibir una variedad lingüística en su conjunto. El percepto sobre la variedad en bloque es construido por el individuo a través de la orientación que suponen las operaciones analíticas (que son imprecisas, como mencionamos anteriormente). Al ser las operaciones analíticas observaciones parciales, estas resultan a veces análisis hipercharacterizadores cuyos elementos seleccionados sirven de generalización para identificar una variedad: «[e]n esta línea podría haber la formación de estereotipos, extendiendo la propuesta de Putnam (1975), como producto de percepciones analíticas imperfectas que sobredimensionan ciertos rasgos sobre otros y llevan a la percepción sintética de una variedad o de varias» (Caravedo, 2014: 121). En nuestro caso, la suma de los diferentes rasgos prominentes en los textos mencionados (sentencia, informe clínico e informe técnico).

Putnam (1975), además, distinguía el discurso propio del especialista del discurso del hablante común (noción de división del trabajo lingüístico), lo que supone contar con percepciones también diferentes: la del especialista y la del hablante lego. Esta percepción científica (Caravedo, 2014: 123) cuenta con el mismo funcionamiento que la percepción del hablante lego, aunque la percepción analítica científica presta atención, selecciona y valora otros elementos distintivos que serán convertidos en perceptos.

2.7. Implicaciones teóricas y metodológicas de la percepción en nuestro estudio

Si aplicamos este concepto a nuestros objetivos, es posible defender que los aspectos relacionados con el grado de especialización textual no se captan hasta el final de la fase secundaria o inicios de la fase terciaria, donde las experiencias lingüísticas con los medios de comunicación de masas aumentan y comienza la exposición a material lingüístico profesional o científico diverso. Ahora bien, este nuevo contacto con estas variedades lingüísticas a través de material textual profesional normalmente se restringe a la actividad académico-profesional del hablante, mientras que otras variedades lingüísticas solo son accesibles mediante los medios de masas. Teniendo en cuenta que los textos especializados constituyen un grupo de medios de comunicación restringidos, en el sentido

de que sus emisores y destinatarios se agrupan en una comunidad específica (los especialistas), es lógico pensar que la percepción cognitiva del hablante se ejecuta en forma de prominencia: la ausencia de ocurrencia de textos especializados en la vida cotidiana de los usuarios conlleva que existan elementos que, aun siendo poco frecuentes, se constituyan como rasgos característicos y, por ende, protagonistas del proceso de atención cognitiva, siendo, por tanto, los elementos asignables a grupos concretos, como especialistas, o a discursos específicos, como, por ejemplo, los relativos a la medicina.

Esta falta de exposición al texto especializado contrasta con el contacto lingüístico que el hablante lego o instruido tiene con el texto divulgativo o semiespecializado, si bien consideramos que la valoración de estas variedades realizada por el hablante es más compleja (que, por ejemplo, la valoración de variedades fónicas o léxicas), pues la complejidad del objeto lingüístico (que es discursivo e integrador) es mayor.

A nuestro juicio, los elementos que un lego o hablante común podría valorar en un texto especializado y sus variantes del grado serían, entre otros: a) las unidades léxicas poco frecuentes y de difícil comprensión (normalmente términos); b) las relaciones conceptuales diversas entre dichas unidades (relaciones partitivas, lógicas y otras, cuya falta de comprensión impide la comprensión global del texto); y c) las estructuras textuales y modos de organización de la información extrañas (como las que existen en las secuencias textuales de los artículos de investigación). El especialista, obviamente, también puede distinguir variedades dentro del propio discurso especializado. Por otra parte, los elementos de un hablante instruido en traducción y lingüística (que será uno de los perfiles con los que cuente uno de nuestros grupos de informantes; *vid. infra* apartado III. 6., p. 216) serán similares, en tanto que participan de las convenciones discursivas de los hablantes comunes si bien presta especial atención al tratamiento de la información en relación con la traducción, por lo tanto, se interesa por: a) la significación y relación de unidades léxicas poco frecuentes (sobre todo si activan valor especializado; es decir, si son terminológicas); y b) cómo y quién organiza la información lingüística y conceptual de dicha forma y a quién se dirige el texto.

La clasificación de los textos especializados y de las unidades léxicas especializadas parte del criterio semántico-pragmático desde el punto de vista de la TCT. Consideramos que la percepción de los usuarios es muy relevante en el análisis de estos elementos, por lo que: (i) desde el punto de vista léxico-semántico, la percepción de los usuarios nos permitirá conocer qué consideran que es un término y su ocurrencia o prominencia en el discurso; y (ii) desde el punto de vista pragmático, los informes de los usuarios podrían permitir medir

la percepción sintética, mediante las decisiones de clasificación de emisores, destinatarios y clases de texto por parte de los informantes.

3. CORPUS PARA ANALIZAR EL GRADO DE ESPECIALIZACIÓN

Contamos con hipótesis, objetivos para validar estas hipótesis y con un enfoque teórico y metodológico para llevar a cabo el estudio, pero es necesario trabajar sobre un objeto de estudio que permita extraer conclusiones. Sus requisitos corresponden con que sea un corpus representativo en español que cuente con gran número de textos de todo el grado de especialización y que sea diverso en tipos de usuarios, tipos de textos y temáticas. Actualmente, el corpus disponible más similar a estas características es el Corpus Técnico del Institut Universitari de Lingüística Aplicada de la Universitat Pompeu Fabra. Se trata de un corpus etiquetado que cuenta con cinco bloques temáticos y más de nueve millones de palabras en español (Cabré y Bach, 2004). Permite realizar búsquedas contextuales y aplicar diferentes filtros para explotar el corpus. El principal problema es que solo es accesible mediante su programa de explotación (BwanaNet) y no nos facilita información esencial en nuestro estudio, como el tipo de texto, ni tampoco es explotable más allá de las opciones que incorpora dicho programa. En otras palabras, no podemos medir el grado de reformulación, el número de términos, el número de formas léxicas, etc.

Por tanto, es necesario compilar un corpus *ad hoc* para este trabajo que supla las carencias del Corpus Técnico con respecto a nuestros objetivos investigadores.

3.1. Criterios de confección del corpus

El corpus de trabajo de esta investigación está construido a través de una articulación multicriterio, en la que atendemos, precisamente, a algunas de las variables que serán analizadas y que constituyen los elementos más relevantes de nuestras hipótesis de partida. En efecto, dado que pretendemos caracterizar el grado de especialización atendiendo tanto a elementos temáticos (campo en el que se inscriben los textos), como lingüísticos (sobre todo densidad terminológica¹²⁵ y reformulación) y pragmáticos

¹²⁵ A pesar de que estudios empíricos como el de Domènech (2006: 207) sostienen que la densidad terminológica de un texto no es un instrumento suficiente para discriminar el grado de especialización, nuestra hipótesis considera que es uno de los elementos básicos para la discriminación del grado. En esta investigación se emplea un método diferente al empleado sobre Domènech (2006) que trabajaba, a partir del Corpus Tècnic del IULA, con Mercedes, un sistema de reconocimiento de términos que cuenta con un diccionario de referencia, y con YATE, un sistema de extracción de

(interlocutores, tipo textual...), el corpus adopta un criterio inicial y, a partir de este, se suman nuevos criterios a partir de los cuales elegir los objetos textuales específicos que sirvan de material textual real. En todo momento se ha intentado asegurar, en la medida de lo posible, no solo la representatividad del corpus a través del número de individuos textuales, sino la homogeneidad de las muestras seleccionadas y, dentro de esta, la heterogeneidad de realizaciones posibles. Así, los criterios de compilación del corpus son los siguientes:

- a) Partiendo del repaso en torno a las reflexiones que han tenido lugar sobre la producción, recepción, función y contexto del discurso oral y escrito que realiza San Vicente (2017: 25-31), destacamos que nuestro corpus cuenta con **textos planificados** en mayor o menor medida (los dos tipos textuales con menor planificación son la entrevista, que se trata de un diálogo preparado por uno de los interlocutores; y el foro, que se trata de un debate entre dos o más personas sobre un hecho); se trata de **textos escritos** cuyo objetivo es ser leídos integra o parcialmente; y son de tipo descriptivo, expositivo o argumentativo.
- b) Los textos están redactados de forma original en **lengua española** con el objetivo de intentar evitar las interferencias lingüísticas de otras lenguas, sobre todo del inglés. A pesar de esto, consideramos que no es posible garantizar la nula interferencia de otras lenguas, dado que tratamos, en ciertos casos, con producciones textuales derivadas de la comunicación internacional, y el contacto interlingüístico suele ser inevitable.
- c) La selección de tres **disciplinas** constituye el punto de partida. Se seleccionan los textos atendiendo a que traten objetos temáticos del ámbito médico, informático y jurídico (*vid. infra* apartado III. 3.2., p. 183). Aunque en un principio planteamos la posibilidad de seleccionar un objeto temático concreto de cada disciplina¹²⁶, finalmente optamos por no restringir el corpus a fenómenos concretos, sino abarcar todas las posibilidades que permite la realidad textual.
- d) Por último, en el marco de cada temática y atendiendo a cada categoría emisor-destinatario, se toma como criterio adicional un **tipo textual** que sea posible dentro de dicho bloque. Así, por ejemplo, sería posible adoptar el criterio tipo de

candidatos a término. Nuestra propuesta incluye el aspecto de la percepción sociocognitiva, ya que no medimos el valor especializado de las unidades léxicas a partir de los criterios de frecuencia en corpus o diccionarios de referencia.

¹²⁶ Trabajar con un corpus de un objeto temático concreto tiene grandes ventajas al tratar con extractores automáticos de terminología, cuyos resultados indicarían candidatos a términos más fiables. No obstante, la desventaja principal supone reducir el estudio a un fenómeno específico dentro de una disciplina, lo que puede afectar a las conclusiones del análisis, ya que se corre el riesgo de que no sean extrapolables a la disciplina como un todo.

texto artículo de revista científica para la relación especialista-especialista o especialista-instruido pero no para el lego-especialista o el especialista-lego, entre otros, puesto que el artículo científico sirve para vehicular conocimiento a interlocutores con competencia cognitiva en el tema (es decir, con conocimiento sobre el tema). Además, para garantizar la homogeneidad de dicha muestra en el corpus, es obligatorio que la realización del tipo textual seleccionado como criterio sea posible en cada disciplina. Por ejemplo, un artículo científico puede tratar tanto temas médicos como informáticos y jurídicos. Por el contrario, un prospecto médico no puede tratar un tema que no se corresponda exclusivamente con la medicina, por lo que no es un texto seleccionable como tipo textual para construir el corpus.

En resumen, el esquema de selección de criterios para el corpus sería el siguiente:

Medicina	Derecho	Informática
		Tipo textual 1
		Tipo textual 2
		Tipo textual 3
		Tipo textual 4
		Tipo textual 5
		Tipo textual 6
		Tipo textual 7

Tabla 32: Esquema de selección de criterios para compilar el corpus de análisis

En definitiva, el corpus estaría constituido por siete tipos textuales comunes a tres disciplinas. De cada tipo textual se recopilan 15 textos, de forma que el corpus se constituye como un conjunto de 315 textos. La justificación de dicho número de textos puede encontrarse en el apartado III. 3.6. (p. 190).

En los siguientes apartados se indican las estrategias específicas para identificar los elementos necesarios que nos permiten adoptar dichos criterios en la compilación del corpus.

3.2. Disciplinas y campos temáticos

Dado que las diferentes disciplinas científicas se aproximan a la realidad desde perspectivas muy diferentes, sería posible atreverse a considerar que las tipologías discursivas en relación con el grado de especialización no se actualicen de forma homogénea en todas las disciplinas: «una de las consecuencias de esta variedad de

tendencias posibles es que la diferencia entre discurso general y discurso especializado no necesariamente ha de ser la misma en todos los campos temáticos» (Alonso y DeCesaris, 2006: 270).

A pesar de que las investigaciones consideran que la temática (Monterde Rey, 2002: 124; Cabré, 2002: 23, 28, 29 y 33; Rodríguez-Tapia, 2016a: 995) no es criterio determinante del grado de especialización textual (sino el tratamiento de esta temática), consideramos que sí que pueden existir diferencias entre disciplinas a la hora de graduar la especialización: en el número de términos que emplean en sus textos, en los procedimientos de reformulación que emplean, los tipos textuales que seleccionan, etc., incluso condicionado por la frecuencia en la que se aborda dicho tema en los medios.

De esta forma, proponemos trabajar con tres disciplinas: medicina, derecho e informática, de manera que sea posible estudiar tres áreas de conocimiento acotadas y bien diferenciadas científicamente y socialmente. Se han seleccionado estas tres disciplinas por tratarse de campos con reconocimiento y relevancia social, cuya repercusión informativa puede medirse cada día en diferentes fuentes de divulgación, principalmente los medios de masas.

Las disciplinas constituyen el punto de partida de selección de textos para compilar el corpus. El tema de cada texto es identificable a través de sus títulos y del análisis del campo semántico del léxico del texto, lo que puede analizarse de forma manual o automática.

3.3. Tipos textuales

Los tipos textuales¹²⁷ se constituyen como productos textuales resultado de convenciones sociales según las necesidades comunicativas de los interlocutores. Estas necesidades comunicativas son las que nos permiten pensar que los grados de especialización tienden a realizarse en tipos textuales específicos, que pueden ser más o menos compartidos, pero que dependen de factores como la relación emisor-destinatario o la función lingüística. La identificación de los tipos textuales parte, necesariamente, de dichos elementos como factores de construcción del tipo textual, de acuerdo con nuestra experiencia lingüística con la realidad.

¹²⁷ Véase la nota al pie número 102 (p. 142) para recordar la diferencia conceptual que hacemos entre tipo textual y clase textual en esta investigación.

Así, la pregunta que debe guiar la compilación del corpus es: ¿qué tipos textuales deben incluirse en el corpus y por qué? Recordemos que uno de los problemas teóricos de los que parte esta tesis doctoral es la falta de vinculación de tipos textuales con el grado de especialización textual y, por tanto, la ausencia de repertorios de estrategias de redacción. Estas estrategias de redacción son el objeto de interés para las fuentes de divulgación. Por tanto, a nuestro juicio, nuestro corpus debe contar con aquellos tipos textuales que sean más propensos a ser divulgados o que se constituyen como formas divulgativas.

En definitiva, partimos de una caracterización apriorística y prototípica de los tipos textuales pero reconocemos que las características que a continuación presentamos constituyen tan solo un ideal de la variabilidad del objeto textual. Los textos que son contabilizados y analizados en el corpus son¹²⁸:

- a) **Artículo en revista científica:** texto académico en el que se reflexiona sobre conocimiento especializado desde una perspectiva especialista. Está emitido por especialistas y se dirige a especialistas o instruidos. Puede obtenerse a partir de revistas científicas.
- b) **Artículo de revista o web de divulgación especializada**¹²⁹: texto informativo no dirigido a especialistas, sino a un público menos restringido interesado en el tema objeto del artículo, normalmente el público general, aunque existen textos dirigidos a instruidos. Puede obtenerse a partir de revistas de divulgación científica.
- c) **Consultas y respuestas en foro en línea:** texto informativo en un canal no restringido a ningún perfil cognitivo-comunicativo, normalmente emitido por especialistas, instruidos o legos y dirigidos a legos o instruidos que tratan de resolver dudas o problemas. Puede obtenerse a partir de foros digitales.
- d) **Entrada de blog:** texto informativo emitido por y dirigido a cualquier tipo de usuario en el que se trata de forma panorámica o en mayor profundidad, temas diversos de actualidad que resultan de interés a diferentes grupos de usuarios, ya sea por su motivación personal, académica o profesional.

¹²⁸ En una primera versión del corpus decidimos incluir también los siguientes tipos textuales, que fueron descartados por considerar que se alejaban ligeramente del criterio de tipo textual común y prototípico de la divulgación científica: folleto o tríptico informativo o publicitario, ensayo universitario redactado, artículo de opinión en periódico generalista, editorial de periódico, capítulo de novela o documento de presentación de diapositivas. Los argumentos que excluyen estos tipos textuales son: que el capítulo de novela cuenta con una función principal diferente al resto de textos, es decir, la estética o lúdica; que el folleto o la presentación de diapositivas presenta una estructura organizativa claramente diferente al resto de tipos textuales; que el ensayo universitario es similar al artículo en revista científica y que el artículo de opinión y la editorial de periódico son análogos a la noticia.

¹²⁹ Cabré (2002b: 31) usa la denominación *textos de divulgación especializada* para hacer referencia a aquellos textos «de amplia difusión destinados al público interesado, pero sin competencia específica en la materia».

- e) **Entrevista:** texto informativo emitido normalmente por instruidos y dirigido a especialistas, instruidos o legos; resultado del diálogo entre dos usuarios en el que el entrevistado es la fuente del conocimiento que se desarrolla en el texto. Puede obtenerse a partir de revistas de divulgación científica o periódicos.
- f) **Noticia:** texto que informa de un hecho o suceso de la actualidad en medios de divulgación (normalmente divulgación de masas) y que está emitido por un especialista o instruido y se dirige al público general.
- g) **Tesis doctoral:** texto académico que persigue la reflexión y argumentación en torno al conocimiento especializado desde una perspectiva especialista. Está escrito por especialistas y se dirige a especialistas. Puede obtenerse a partir de los repositorios de tesis doctorales.

Existen otros tipos textuales que podrían incluirse entre estos criterios pero consideramos que estos representan los ejemplos más frecuentes y más representativos para ilustrar el grado de especialización, por mencionarse o incluirse diariamente en los medios de masas. Además, como indicamos anteriormente, se trata de tipos textuales que garantizan la homogeneidad del corpus, puesto que es posible que en cada tipo textual se traten las tres temáticas que proponemos como criterio principal. Dado que nuestro objetivo es contrastar las características según las disciplinas, necesariamente tenemos que trabajar con tipos textuales idénticos en dicho punto del eje horizontal.

3.4. Representatividad, selección de muestras y problemas de compilación

El universo o población de textos de nuestro estudio engloba todos aquellos textos que cumplen todos nuestros criterios de compilación (es decir, cualesquiera de los tipos textuales señalados anteriormente, en español, que versen sobre medicina, derecho o informática). Dado que es imposible compilar un corpus que englobe la totalidad de dicha población, necesitamos trabajar con una muestra, con un subconjunto de elementos de dicha población. Nuestra muestra es una muestra probabilística, pues se trata de un «subgrupo de la población en el que todos los elementos de ésta tienen la misma posibilidad de ser elegidos» (Hernández, Fernández y Batista, 2015: 176).

Una vez seleccionados los criterios de compilación del corpus y previstas las posibles complicaciones que pueden afectar a la selección de la muestra, se inicia el debate en torno a la representatividad del subconjunto que vamos a analizar. Debe recordarse que

nuestro interés se halla en la finalidad predictiva del corpus, es decir en la posibilidad de generalizar los resultados al total de la población a partir de las observaciones particulares de la muestra, por lo que asegurar la representatividad tanto cualitativa como estadística es un aspecto esencial de la tesis. Como menciona Kabatek (2013: 14), a partir del valor de la representatividad en la sociología:

La representatividad sociológica deriva de la identificación de parámetros correlacionables que condicionan el comportamiento humano. Tales parámetros permiten reducir la investigación de un fenómeno supuestamente relevante para una población exhaustiva a una porción reducida de la población, la cual refleja porcentualmente los parámetros relevantes.

Torruella (2017: 129) afirma al respecto que:

para que un corpus sea el modelo adecuado del universo lingüístico que se quiere describir, su composición debe cubrir todas las variedades lingüísticas en función al tipo de corpus que se quiera crear (representatividad), y ha de hacerlo en la proporción adecuada (equilibrio), dos aspectos básicos dentro de la lingüística de corpus.

Para garantizar la representatividad de un corpus debe cumplirse necesariamente la característica de que la muestra sea equilibrada, sobre todo en corpus pequeños como el nuestro, que no alcanzan los diez millones de palabras. Ahora bien, son numerosos los autores (Caravedo, 1999; Parodi, 2008; o Kabatek, 2013, entre otros) que cuestionan la posibilidad de alcanzar la representatividad en un corpus. Entre ellos se encuentra Baker (1995: 239), que desde el punto de vista teórico y metodológico, sostiene que:

in theoretical terms then, it is untenable to assume that a perfectly balanced and representative corpus can ever be achieved. [...] Any corpus, whether of whole texts or text fragments, is essentially a sample of a particular domain of language, or of the general core of everyday language. First, it is impossible to delimit in any rigorous way the population of texts which itself constitutes the domain in question. This makes the application of methods of statistical sampling irrelevant, given that statistical sampling is generally done on clearly defined populations. Second, almost every unit of language that we might want to consider as the basis for collecting a 'sample' raises its own problems of definition. Even 'text' is not a well-defined unit. Similarly, most of the criteria that we have to use in selecting our sample, for example informal vs. formal, has a high subjective element to it and further complicates the question of representativeness. And finally, the sheer size of the population we need to sample (namely language) and, by comparison, the severe restrictions on available resources (whether physical, financial or human), all mean that it is virtually impossible to ensure coverage of every feature of the population being sampled.

La representatividad se observa, por tanto, como una especulación acerca de lo que no se puede ver (Kabatek, 2016: III), puesto que «el grado de representatividad en los corpus [...] no puede resultar nunca del cien por cien» (Torruella, 2017: 136). En este sentido, las conclusiones que se proyectan en los estudios sobre lingüística de corpus deben tener en cuenta los límites que alcanza la muestra analizada. Nuestro corpus sigue, por tanto, los principios esbozados por Moreno (2008: 310) para los estudios sociolingüísticos:

Principio de la cuantificación:

para establecer relaciones entre variables lingüísticas y variables sociales se requieren métodos cuantitativos [...]

Principio de la responsabilidad:

los datos de aparición de una variante de una variable lingüística deben ir acompañados de los datos de la no aparición de esa variante [...]

Principio de la representatividad:

el habla de una muestra debe ser representativa de un grupo social o de una comunidad.

3.5. Representatividad cualitativa

La representatividad cualitativa hace referencia a la calidad de las muestras seleccionadas¹³⁰. Como menciona Torruella (2017: 138), deben cumplirse cuatro condiciones para que los textos de un corpus se consideren representativos desde el punto de vista cualitativo:

1. Cada rasgo definitorio de un texto debe ser claro y diferenciado de los otros rasgos definitorios.
2. Todos los rasgos definitorios de los textos deben haber sido establecidos a partir de los mismos criterios.
3. Los rasgos definitorios de los textos deben coincidir con los rasgos definitorios de un apartado del corpus.
4. Cada texto del corpus debe encajar totalmente en los rasgos definitorios de su apartado y solo en los de su apartado.

En nuestro caso, cada texto forma parte de un estrato diferenciado del corpus, resultado de la suma de criterios de compilación, de forma que, aunque todos los textos compartan parte de sus características, siempre contarán con una característica esencial que los diferencia del resto. Por ejemplo, el texto MED ART07 comparte característica con el texto DER ART08, ya que ambos son textos extraídos de artículos científicos, siendo su

¹³⁰ La relación de textos y sus fuentes se recoge en el anexo 4.

característica esencial (y, por tanto, diferenciadora), el tema que tratan (medicina y derecho, respectivamente). En el caso del texto DER DIV02, su característica diferenciadora con respecto al texto DER ART08 es el tipo de texto (se trata de un texto extraído de una revista de divulgación).

Debido a la complejidad del estudio, que combina el análisis automático con el análisis manual (*vid. infra* apartados III. 6. y III. 7., pp. 216 y 228), ha sido necesario proponer un conjunto de criterios para seleccionar el contenido de los textos analizables en el corpus. En otras palabras, existen textos que no son analizados en su totalidad, sino que se ha seleccionado una muestra de estos, normalmente, por secciones diversas a lo largo de los textos (sobre todo en artículos científicos y tesis doctorales).

En relación con este problema existen tanto investigaciones que se posicionan a favor de seleccionar textos de forma íntegra, puesto que consideran que el estilo de lengua varía según las diferentes secciones del texto (como en los apartados de introducción y conclusión de trabajos académicos; Kennedy, 1998: 68; Biber, Conrad y Reppen, 1998: 249), y aquellos que prefieren trabajar sobre partes de texto de un abanico mayor de muestras. Por tanto, una de las soluciones propuestas para los textos más amplios en extensión consiste en una postura intermedia: «recoger muestras del principio, del medio y del final de cada obra para tener representadas las características de cada parte» (Torruella, 2017: 146).

La selección del contenido de todos los tipos textuales ha sido aleatoria en aquellos textos de gran extensión (artículos de revistas especializadas y tesis doctorales, principalmente). A todos los tipos textuales se han aplicado una serie de criterios de exclusión de cierta información, que no es analizada ni computa de forma estadística, es decir, información que es eliminada de cada texto:

Tipo de texto	Información eliminada
a) Artículo de revista científica	Autores, referencias bibliográficas, datos institucionales, imágenes ¹³¹ , <i>abstract</i> en inglés, títulos en inglés, notas al pie ^{132, 133} .
b) Artículo de revista o web de divulgación	Autores, información geográfica y temporal, referencias bibliográficas, imágenes.
c) Consultas y respuestas en foro en línea	Autores, fechas, metainformación.
d) Entrada de blog	Autores, información geográfica y temporal, referencias bibliográficas, imágenes, índices de contenido.

¹³¹ Dado que en nuestra investigación, una de las variables es la reformulación lingüística parafrástica, las imágenes y gráficas, que ejercen una reformulación no lingüística, no entran dentro de nuestro objeto de estudio.

¹³² Aunque las notas al pie a veces tienen función reformulativa o aclarativa, hemos decidido excluirlas del análisis por considerarlas textos complementarios al texto principal, puesto que a veces solo incluyen digresiones, comentarios o referencias bibliográficas.

¹³³ En este tipo de texto se han seleccionado fragmentos aleatorios de todas sus secciones (introducción, estado de la cuestión, método, resultados, conclusiones, etc.).

Tipo de texto	Información eliminada
e) Entrevista	Autores, información geográfica y temporal, referencias bibliográficas, imágenes.
f) Noticia	Autores, fechas, metainformación.
g) Tesis doctoral	Autores, títulos, referencias bibliográficas, datos institucionales, imágenes, <i>abstract</i> , notas al pie ¹³³ .

Tabla 33: Información eliminada de los textos del corpus

3.6. Tipo de corpus compilado

Para ilustrar la complejidad clasificatoria de nuestro corpus, usaremos los parámetros que menciona Torruella (2017: 41 y 42).

Criterio	Tipos y subtipos de corpus		
Modalidad	Oral		
	Escrito		
	Mixto		
Temática	General		
	Especializado	Genérico	Canónico
Época	Contemporáneo		
	Histórico		
Temporalidad	Sincrónico		
	Diacrónico		
Magnitud	Grande		
	Restringido		
Evolución	Abierto		
	Cerrado		
	Monitor		
Distribución	Proporcional		
	Equivalente		
Número de ediciones	Monoedición		
	Pluriedición	Comparable	Alineado
Número de lenguas	Monolingüe		
	Plurilingüe		
Tipo de edición	Reproducción fotográfica		
	Edición paleográfica		
	Texto normalizado		
	Texto crítico		
	Etc.		
Muestras	Textual		
	De referencia		
Marcaje	Léxico		
	Simple		

Criterio	Tipos y subtipos de corpus		
			Codificado
			Anotado
	Etiquetado		morfológicamente
		Anotado	Lematizado
			Parentizado
			Analizado

Tabla 34: Criterios de clasificación y tipos de corpus según Torruella (2017)

Modalidad

Se trata de un *corpus escrito* o corpus textual que contiene una colección de textos en lengua escrita en formato electrónico. La característica común de todos ellos es que se trata de textos escritos y planificados.

Temática

Nuestro corpus recoge exclusivamente algunas variedades lingüísticas concretas que atienden a la suma de criterios de campo temático y tipo textual, por lo que hablamos de que se trata de un *corpus especializado*. En concreto, sería un corpus especializado *genérico*, puesto que la recopilación que realizamos atiende al tipo textual.

Época

El corpus se centra en la lengua sincrónica, puesto que nuestro interés se halla en el estudio de las distintas variedades lingüísticas de la actualidad. En concreto, las muestras seleccionadas cubren los años de 2007 a 2017. Por ello, hablamos de que se trata de un *corpus contemporáneo*.

Temporalidad

Puesto que nos centramos en una única etapa temporal, nuestro corpus es un *corpus sincrónico*.

Magnitud

Los corpus grandes quedan reservados a aquellos que superan los cien millones de palabras. Dado que nuestro corpus alcanza tan solo el medio millón de palabras, se concibe como un *corpus restringido*. En concreto, se definiría como un corpus pequeño desde el punto de vista de Guerreau y de Torruella (2017: 47). Más allá de la propia

representatividad de un corpus pequeño respecto a la totalidad de la lengua o del fenómeno que se estudia, el tamaño del corpus no tiene ninguna implicación en la fase de análisis de esta tesis doctoral.

Evolución

Una vez alcanzado el número necesario que permite considerar el corpus como un corpus equilibrado y representativo, la inclusión de textos adicionales finalizó y no se espera aumentar el corpus en ninguna otra fase. Por ello, hablamos de *corpus cerrado*.

Distribución

Las cantidades de palabras, similares en gran medida en todos los subcorpus, en todas sus secciones (aunque no necesariamente entre tipos textuales diferentes) nos permiten hablar de un *corpus equivalente*.

Número de ediciones

Contamos con un *corpus monoedición*, ya que el corpus está constituido exclusivamente por un texto único, sin versiones, para cada instancia analizable.

Número de lenguas

Se trata de un *corpus monolingüe*, aunque incluye variedades tanto peninsulares como del continente americano.

Muestras

Las muestras seleccionadas clasifican el corpus como un *corpus textual*, ya que, normalmente, seleccionamos el contenido completo de un texto (es decir, el texto se analiza íntegramente). Debido a que parte de análisis es manual y, para evitar un análisis más complejo, las obras más extensas en número de palabras han sido fragmentadas, de forma que solo partes de ellas se han incluido en el texto final que es analizado. Por esta razón, puede hablarse de *corpus de referencia*.

Marcaje

Dado que nuestro corpus es un corpus *ad hoc*, no ha sido sometido a marcaje automático o manual, puesto que la fase de marcaje conlleva una fase de investigación acorde con el objetivo del corpus. Como el análisis del corpus se realiza mediante grupos

de informantes, nuestro corpus no está marcado lingüística, formal ni estructuralmente. Se trata de un *corpus simple*.

3.7. Representatividad cuantitativa: características estadísticas

La finalidad de la representatividad cuantitativa es alcanzar el equilibrio estadístico del corpus. Este equilibrio puede alcanzarse tanto de manera externa como de manera interna.

El **equilibrio externo** establece la relación entre el número de muestras que se seleccionan y el total de la población. Dado que, a nuestro juicio, no es posible contabilizar la totalidad de la población de nuestra investigación, el grado de satisfacción que presenta la representatividad es debido al nivel de confianza del investigador. En nuestro caso, contamos con un corpus de 315 textos de un total poblacional de, seguramente, cientos de millones de textos teóricamente seleccionables.

En cuanto al **equilibrio interno**, hace referencia al número de muestras para cada apartado clasificador del corpus. Su finalidad es que sea posible comparar los datos de los distintos apartados mediante técnicas estadísticas. El equilibrio interno puede conseguirse mediante un reparto proporcional sobre el total de la población o mediante un reparto equivalente de las muestras. Este último no requiere conocer el total de la población, puesto que las cantidades de palabras de cada apartado son similares y se prescinde de la «correspondencia porcentual entre las diferentes partes en el corpus y entre las diferentes partes y el todo» (Torruella, 2017: 141). Nuestro corpus alcanza el equilibrio mediante un reparto equivalente y la distribución de los textos atiende a las recomendaciones que conlleva dicho reparto: es preferible contar con un número considerable de textos de extensión reducida a contar con pocos textos de gran extensión, pues «no es lo mismo, por ejemplo, que cien ocurrencias de un apartado compuesto por cincuenta obras se repartan entre tres de ellas a que se repartan entre treinta [y] cinco» (Torruella, 2017: 141).

En nuestra tesis no solo nos interesa el número de palabras¹³⁴ del corpus, sino también el número de instancias o patrones¹³⁵ (i. e. número de textos) por cada variable de

¹³⁴ *Vid. supra* apartado II. 5.3. la distinción que se realiza sobre el término *palabra* aplicado a esta tesis doctoral, que reside en su valor para la estadística y la cuantificación de cadenas de caracteres.

¹³⁵ En relación con la representatividad cuantitativa de un corpus, Corpas y Seghiri (2009 y 2010) investigan sobre el umbral mínimo de representatividad de un corpus empleando para ello un algoritmo de análisis de la densidad léxica en relación con el aumento incremental de un corpus (N-Cor). El resultado de esta investigación es el *software* ReCor que determina el número de textos que sirven de umbral mínimo para determinar la representatividad de un corpus, dado que, una vez alcanzado cierto número de textos, el número de tipos de unidades no aumenta en proporción al número de

análisis, ya que la base de datos que resulte del análisis del corpus será explotable mediante aprendizaje automático, lo que nos obliga a contar con un mínimo de 30 instancias por variable. Para ilustrar los tipos textuales que se seleccionan siendo comunes a las tres disciplinas con las que trabajamos, esbozamos la siguiente tabla, que incluye el número de textos que seleccionamos por cada tipo:

Tipo textual	Medicina	Derecho	Informática
a) Artículo de revista científica (ART)	15 textos	15 textos	15 textos
b) Artículo de revista o web de divulgación (DIV)	15 textos	15 textos	15 textos
c) Consultas y respuestas en foro en línea (FOR)	15 textos	15 textos	15 textos
d) Entrada de blog (BLOG)	15 textos	15 textos	15 textos
e) Entrevista (ENT)	15 textos	15 textos	15 textos
f) Noticia (NOT)	15 textos	15 textos	15 textos
g) Tesis doctoral (TES)	15 textos	15 textos	15 textos
TOTAL	105 textos	105 textos	105 textos
			315 textos

Tabla 35: Relación de número de textos por tipos textuales en el corpus

Biber, en su trabajo de 1988, parte de un corpus de un millón de palabras repartidas en 23 géneros diferentes y considera que se trata de un corpus extrapolable a la totalidad de la lengua inglesa. En cuanto al número de palabras por muestra, Biber (1990) señala que pueden ser suficientes muestras de entre 2000 y 5000 palabras; incluso muestras de entre 1000 y 5000 palabras según Torruella (2017: 146) para estudios de tipo léxico. Como nuestro interés se halla en compilar un corpus que cumpla con equilibrio interno mediante reparto equivalente, tenemos en cuenta las palabras de Sinclair (1992: 379-497), quien defiende que alcanzar el grado óptimo de representatividad deriva de recoger un número de palabras suficientemente elevado de cada texto para que sea posible analizar sus características particulares. Ahora bien, la selección de textos y palabras debe contar con ciertos límites, tanto por eficiencia en el análisis como por la propia rentabilidad del corpus, ya que el número de formas o fenómenos lingüísticos nuevos que proporciona un corpus disminuye a medida que el tamaño del corpus crece.

Así, en lugar de seleccionar un menor número de patrones y aumentar el número de palabras en cada patrón, hemos decidido aumentar el número de patrones y reducir el número de palabras en cada uno, de forma que pudiésemos contar con más patrones diferentes y otorgar heterogeneidad al corpus. El corpus provisional que compilamos en mayo de 2017 cuenta con 315 textos: 105 por temática, 45 por tipo de texto; o, en otras palabras, 15 textos por tipo y temática. En cuanto al total de textos, debe mencionarse que

palabras que contiene el corpus (Biber, 1993; Sánchez y Cantos, 1998). Este software sería verdaderamente útil en nuestra investigación si nuestros objetivos estuviesen relacionados con la investigación de patrones léxicos o sintácticos, así como con la frecuencia léxica. No obstante, dado que nuestro interés se halla en cuestiones tipológicas, extralingüísticas o perceptivas del léxico, consideramos que los resultados que arroja ReCor no son del todo determinantes en nuestro trabajo.

se trata de una cifra que permite contar con distintas variedades discursivas por cada variable, lo que asegura la heterogeneidad de las muestras: 15 textos por tipo y tema constituyen un conjunto lo suficientemente amplio para estudiar fenómenos divergentes para una misma variable y, a su vez, ser estudiados mediante análisis manual por diferentes tipos de informantes.

El total de palabras alcanza las 525 875 palabras¹³⁶ contabilizadas con WordSmith Tools, la mitad de palabras que defiende Biber para la totalidad de la lengua inglesa. A nuestro juicio, dado que estudiamos en concreto una variedad lingüística (el grado de especialización), medio millón de palabras resulta una muestra representativa del total de la población. A continuación, se presenta la relación provisional de número de palabras y de formas léxicas (FL) por tipo de texto y temática.

		ART	ENT	BLOG	NOT	FOR	DIV	TES	TOTAL
Medicina	n.º palabras	42 394	24 572	12 699	8 295	12 678	13 495	49 688	163 821
	% total	8 %	5 %	2 %	2 %	2 %	3 %	9 %	31 %
	n.º FL	22 905	12 395	6 560	4 327	6 536	7 187	26 786	86 696
	% total	8 %	5 %	2 %	2 %	2 %	3 %	10 %	32 %
Informática	n.º palabras	62 578	16 872	10 972	13 622	14 316	10 330	40 095	168 785
	% total	12 %	3 %	2 %	3 %	3 %	2 %	8 %	32 %
	n.º FL	33 063	8 695	5 613	7 272	7 371	5 442	21 159	88 615
	% total	12 %	3 %	2 %	3 %	3 %	2 %	8 %	32 %
Derecho	n.º palabras	63 682	25 843	18 202	11 398	13 617	12 899	47 628	193 269
	% total	12 %	5 %	3 %	2 %	3 %	2 %	9 %	37 %
	n.º FL	32 427	13 175	9 085	6 733	6 733	6 831	24 052	99 036
	% total	12 %	5 %	3 %	2 %	2 %	2 %	9 %	36 %
Total palabras									525 875
Total FL									274 347

Tabla 36: Relación provisional de número de palabras y de formas léxicas por tipos textuales en el corpus antes del análisis estadístico

Artículos y tesis doctorales (aun habiendo sido seleccionado solo parte de su contenido) cuentan con un mayor número de palabras por la propia naturaleza del texto, más amplio que el resto. Todas las muestras son similares porcentualmente y la mayor diferencia se halla en el campo del derecho. Tras el primer muestreo se identificó que el subcorpus de derecho contaba con unas 20 000 palabras más que cada uno de los otros dos subcorpus, por lo que se decidió balancear el corpus reduciendo la selección de contenido en aquellos grupos muestrales que estadísticamente superaban muy por encima la media (es decir, los subcorpus de artículos en revistas científicas, entradas de blog, artículos en

¹³⁶ Recuértese que en nuestro trabajo consideramos palabra toda aquella cadena de caracteres entre dos espacios en blanco.

revistas o webs de divulgación especializada, entrevistas y tesis doctorales). Tras volver a realizar la operación de conversión de archivos y realizar el recuento de la totalidad del corpus, seguimos identificando una leve descompensación en el subcorpus de artículos en revistas científicas sobre medicina (8 % frente a 12 % en los otros subcorpus). Por tanto, se decidió aumentar el número de palabras de la sección de artículos científicos del subcorpus de medicina en 20 000 palabras para garantizar que cada corpus fuese equivalente estadísticamente en cuanto a cada tipo de texto. No obstante, al revisar la selección de contenido que se incluía en el subcorpus, se llegó a la conclusión de que los artículos de temática médica cuentan con menor extensión que los de temática informática o jurídica, por lo que no era posible aumentar el contenido de los textos.

Por tanto, existía la posibilidad de reducir en 20 000 palabras tanto el subcorpus de artículos científicos en informática como en derecho (lo que reduciría el tamaño total del corpus a 470 000 palabras aproximadamente). No siendo esto una opción viable y teniendo en cuenta que la descompensación estadística es fruto de la propia realidad textual, concluimos que la mejor opción sería mantener el corpus con las características que presentaba.

Así, el resultado final del corpus definitivo es el siguiente:

		ART	ENT	BLOG	NOT	FOR	DIV	TES	TOTAL
Medicina	n.º palabras	62 498	23 845	14 284	11 555	13 845	12 056	44 754	182 837
	% total	12 %	5 %	3 %	2 %	3 %	2 %	9 %	35 %
	n.º FL	31 843	12 353	7 237	5 832	6 949	6 105	23 070	93 389
	% total	12 %	5 %	3 %	2 %	3 %	2 %	9 %	35 %
Informática	n.º palabras	43 905	12 795	13 650	24 691	12 794	8 347	51 679	167 861
	% total	8 %	2 %	3 %	5 %	2 %	2 %	10 %	32 %
	n.º FL	22 777	6 776	7 320	12 791	6 754	4 389	26 707	87 514
	% total	8 %	3 %	3 %	5 %	3 %	2 %	10 %	32 %
Derecho	n.º palabras	63 506	11 006	10 363	16 987	14 372	13 691	40 447	170 372
	% total	12 %	2 %	2 %	3 %	3 %	3 %	8 %	33 %
	n.º FL	33 960	5 819	5 482	7 536	7 536	7 227	21 656	89 216
	% total	13 %	2 %	2 %	3 %	3 %	3 %	8 %	33 %
Total palabras									521 070
Total FL									270 119

Tabla 37: Relación definitiva de número de palabras y de formas léxicas por tipos textuales en el corpus después del análisis estadístico

4. TIPO DE ESTUDIO Y MÉTODO

Nuestro estudio adopta un enfoque cuantitativo mixto con alcance eminentemente correlacional pero también descriptivo (Hernández, Fernández y Batista, 2015: 80, 81 y 546):

- a) Es un estudio **cuantitativo-mixto** (Johnson y Onwuegbuzie, 2004) porque usamos a nuestro favor los beneficios de los enfoques cuantitativo (que es el que tiene mayor peso en nuestra tesis) y cualitativo, a fin de poder percibir una panorámica mucho mayor del problema que abordamos. Esto supone, en cierta medida, usar tanto el método hipotético-deductivo como el método inductivo en fases diferentes de la investigación. En efecto, el diseño de un modelo teórico útil para clasificar el grado de especialización resulta de la generalización de análisis de un corpus representativo (método inductivo) y de su posterior validación sobre un corpus de control (método hipotético-deductivo), que permita confirmar o refutar hipótesis. En sociolingüística este planteamiento se concibe como una aproximación sociológica, que se fundamenta en el análisis cuantitativo de grupos humanos amplios y complejos (Moreno, 2008: 309). La estadística descriptiva, que incluye pruebas sencillas de cuantificación, como el recuento de frecuencias absolutas, relativas, medias, modas, etc. nos permite contabilizar los datos y ordenarlos, mientras que la estadística de inferencias nos permite aplicar las conclusiones de los análisis de estos datos a entidades sociales mayores que no han sido estudiadas en su totalidad (Moreno, 2008: 315).
- b) Es un estudio **correlacional**, desde el punto de vista del enfoque cuantitativo, puesto que pretendemos identificar la relación que existe entre dos o más variables a partir de «patrones predecibles para un grupo o población» (Hernández, Fernández y Batista, 2015: 81). En palabras de Moreno (2008: 309):

[...] prestar atención a los fenómenos lingüísticos y sociales como elementos característicos de una estructura general o de una configuración supraindividual: se trata de conseguir información suficiente para hacer una caracterización adecuada de esa estructura supraindividual, llámese ciudad o barrio, llámese comunidad de habla o llámese clase social.

En nuestro caso, queremos conocer la relación que existe entre el uso de cierto tipo de términos en ciertos tipos textuales, los usuarios que emiten o reciben el texto y su correspondencia con una clase textual según el grado de

especialización. Puede que existan correlaciones claras que siempre se establezcan y otras que nunca lo hagan.

- c) **Descriptivo**, desde el punto de vista del enfoque cuantitativo, porque nos centramos en buscar las tendencias de una población. Este alcance «busca especificar propiedades, características y rasgos importantes de cualquier fenómeno que se analice» (Hernández, Fernández y Batista, 2015: 80).

Se sigue un **diseño no experimental**, puesto que no se varían intencionalmente las variables independientes, es decir, se observan los fenómenos en su contexto natural. Las ventajas de este diseño no experimental se centran en su mayor validez externa, ya que permite generalizar las conclusiones a toda la población. Este diseño no experimental es transversal, de tipo descriptivo y correlacional-causal (Hernández, Fernández y Batista, 2015: 151, 152 y 155).

- a) **Transversal**, desde la perspectiva cuantitativa, porque los datos recogidos en el corpus pertenecen exclusivamente a la actualidad y el objetivo es describir las variables en un momento específico.
- b) **Transversal descriptivo**, porque se orientan a la descripción de pronóstico y tiene como objetivo conocer la relevancia de diferentes variables en un objeto (el texto) y su situación.
- c) **Transversal correlacional**, porque pretendemos conocer las relaciones entre variables, aunque no necesariamente precisando relaciones de causalidad.

5. VARIABLES DE ESTUDIO

Moreno (2008: 297) explica las relaciones entre la actualización discursiva y los actores de la comunicación:

Al relacionar directamente la organización de niveles de la sociedad y de la lengua se observa que la sociolingüística ha proporcionado fundamentos empíricos de la conexión que existe entre los niveles más concretos de ambas estructuras (se ha explicado qué usos lingüísticos se dan en las relaciones entre individuos), entre estos niveles y los de abstracción media (se ha explicado qué usos lingüísticos distinguen a unas clases o grupos de otros) e incluso se ha intentado incorporar los factores sociales como elementos de una competencia sociolingüística.

Para explicar estas relaciones, Moreno (ídem) se sirve de los componentes que expone Searle (1969):

- a) Componente regulador: existe independiente de las formas de conducta, es decir, de los propios sistemas o códigos lingüísticos y la propia organización social. No está sujeto a manipulación individual.
- b) Componente constitutivo: crea una actividad, es decir, corresponde con la conducta verbal y la comunicación lingüística. Puede ser más o menos manipulable por el individuo.

Así, nos encargaremos de estudiar estos componentes constitutivos de los discursos fruto de los actos comunicativos especializados, por lo que nos centraremos en el estudio de diversas variables.

Una variable es una característica, propiedad o atributo de una población (como la edad de una persona, su color de ojos, el tipo de texto, el número de palabras, etc.) que está sujeta a variación según los elementos de dicha población y puede medirse u observarse (Lavid, 2001: 302; Hernández, Fernández y Batista, 2015: 93). En palabras de Herrera, Martínez y Amengual (2011: 20), entendemos por *variable*:

cada rasgo o característica que se registra en una muestra de individuos, que siendo común a un cierto grupo de individuos u objetos tiene distintos grados de magnitud (e.g. peso, CI, puntuaciones obtenidas, etc.) o las diferentes categorías en que se presentan nuestros datos (e.g. edad, género, religión, etc.).

Es posible distinguir dos tipos de variables¹³⁷ principalmente según su modo de medición (Lavid, 2001: 302 y ss.):

- a) **Variables cualitativas** o categóricas: corresponden con características que no son cuantificables, como el color de ojos de una persona. Son las que varían en tipo (Herrera, Martínez y Amengual, 2011: 22). En estas variables las categorías se excluyen mutuamente. Dentro de estas, distinguimos las variables nominales y las variables ordinales:
 - (i) **Variables nominales**: clasifican los datos según si existe o no un valor para las características, como, por ejemplo, el sexo de una persona (hombre o mujer). En nuestro caso, contamos con dos variables categóricas nominales: el **tema del texto** y el **tipo de texto**. Estas variables, a su vez, son variables de control, pues pueden regular los resultados del estudio, aunque no constituyen el objeto central de la investigación. El tema y el tipo textual nos

¹³⁷ Todas las variables que mencionamos son variables independientes, a excepción de la clase de texto, que es la variable dependiente de nuestro estudio. La variable dependiente es la que pretende medirse en la investigación (es la causante); la variable independiente es aquella que se sospecha que puede afectar a la variable dependiente (es la consecuente).

permiten controlar el estudio incluyendo exclusivamente miembros de la población que cumplan las características que dictan estas variables de control. Siguen los criterios de compilación del corpus, indicados en los apartados III. 3.2. y III. 3.3. respectivamente (pp. 183 y 184).

(ii) **Variables ordinales:** permiten identificar una gradación o establecer cierto orden sobre el conjunto de valores que puede adquirir la variable. Las variables categóricas ordinales que manejamos en este estudio son: **emisor** y **subemisor**, **destinatario** y **subdestinatario** y la **clase de texto**. La clase de texto, a su vez, es la variable dependiente que investigamos.

b) **Variables cuantitativas:** pueden medirse de forma numérica, es decir, varían en número (Herrera, Martínez y Amengual, 2011: 22). Entre ellas, encontramos el **número de formas léxicas**, la **ratio estandarizada** entre tipos de formas léxicas y el total de unidades léxicas, el **índice de densidad terminológica**, el **número de procedimientos de reformulación formal** y sus **tipos** y el **índice de reformulación**. Seleccionamos estas variables como cuantitativas puesto que nos permiten trabajar con densidades de diferentes fenómenos que pueden ser contrastables numéricamente y, además, ser descriptivas e incluso explicativas de los componentes constitutivos del acto comunicativo.

De acuerdo con Moreno (2008: 39), las variables extralingüísticas pueden determinar la variación hasta donde el sistema de la lengua lo permite, mientras que las variables lingüísticas son capaces de «incidir en la aparición de tales o cuales variantes de una variable determinada» (ídem). En nuestro caso, nuestra intención se halla en correlacionar las variables extralingüísticas con las lingüísticas teniendo en cuenta el hecho de que la variable de competencia cognitivo-comunicativa según el emisor y el destinatario que estudiamos aquí no tiene por qué manifestarse lingüísticamente a través de las mismas variantes:

los factores sociales actúan sobre la lengua de una forma irregular, es decir, en dos comunidades de habla diferentes la variación sociolingüística de un mismo fenómeno no tiene por qué manifestarse de la misma manera; en segundo lugar, porque los factores sociales no están configurados de forma idéntica en todas las comunidades, aunque en ellas se hablen modalidades cercanas de una misma lengua (Moreno, 2008: 40).

5.1. Diversidad léxica: la relación tipos de formas léxicas y formas léxicas

La variedad léxica de un texto constituye un dato que, en relación con otras variables (como el tipo, el tema o la clase del texto y, especialmente, con variables numéricas como el índice de densidad terminológica o de reformulación) permite conocer el uso del léxico en el acto discursivo.

Existen multitud de métricas para calcular la variedad o diversidad léxica de un texto¹³⁸ (véanse los trabajos de Granger y Wynne, 2000; y de Roberto, Marti y Salamó, 2012: 99 para conocer algunas de estas métricas). Una de las fórmulas más empleadas (Pérez, 2002) es la fórmula propuesta por Templin (1957) que calcula la ratio *types/tokens* (TTR por sus siglas en inglés) o relación entre tipos de palabras y palabras totales de un texto. Los *tokens* son las palabras que contiene un texto, considerándose palabra cualquier cadena de caracteres separada por espacios. En caso de estar separada por guiones, la hemos considerado como dos palabras. Los *types* son el número de tipos, es decir, las palabras diferentes de dicho texto. Cuando se divide el número de unidades de tipos de palabras entre el número total de palabras (es decir *types* entre *tokens*) y se expresa el resultado como valor porcentual o decimal, se obtiene la ratio de tipos de unidades y palabras. Previamente se aplica una lista de exclusión de palabras gramaticales, de forma que las preposiciones, conjunciones, etc. quedan fuera del cómputo. Es una forma de expresar la diversidad o variedad léxica de un texto. Cuanto mayor es el porcentaje, mayor es la diversidad léxica.

La TTR depende inevitablemente de la longitud del texto. Para evitar esto, Yule (1944) diseñó la fórmula denominada *K de Yule* para el cálculo del índice de diversidad léxica, que calcula la proporción de tipos con respecto al total de palabras independientemente de la longitud del texto (Miranda-García y Calle, 2005).

$$K = \frac{10^4 \times (\sum i^2 T_i - N_{lex})}{N_{lex}^2}$$

T_i = número de *types* léxicos que ocurren i veces.

N_{lex} = *tokens* léxicos

Tabla 38: Fórmula K de Yule para calcular el índice de diversidad léxica

No obstante, teniendo en cuenta el tamaño de nuestro corpus, que los textos son comparables, en mayor o menor medida, por el número de palabras y que la fórmula de

¹³⁸ Laufer (2003) considera que la diversidad léxica es una combinación de la densidad de palabras poco frecuentes (*rare word density*), también denominada *sofisticación léxica (lexical sophistication)* y la variación léxica, también denominada *ratio type/token* o *densidad léxica (lexical density)* (Carter, 1987) (Malvern *et al.*, 2004: 192).

densidad terminológica y de reformulación tiene semejanzas con la fórmula para la TTR, usaremos una modificación de la fórmula de Templin (1957): la TTR estandarizada.

Para obtener el porcentaje de TTR de cada texto se ha usado WordSmith Tools. En nuestro caso, tanto los *types* como los *tokens* solo tienen en cuenta las unidades léxicas, habiendo sido excluidas previamente mediante una lista de exclusión todas las palabras gramaticales del texto (preposiciones, conjunciones, demostrativos, pronombres, locuciones diversas, etc.). Por tanto, para ser más precisos desde el punto de vista denominativo, la ratio que obtenemos establece la relación entre tipos de unidades y unidades léxicas, puesto que todas las formas lingüísticas gramaticales han sido omitidas del cómputo.

La TTR estandarizada se calcula cada n palabras en cada archivo individual, siendo n el valor que el investigador seleccione, por ejemplo, 1000. De esta forma, WordSmith Tools calcula la TTR de las 1000 primeras palabras; tras esto, realiza el cálculo de las 1000 siguientes, y así sucesivamente. Después, realiza la media de todos los cálculos del total de palabras del texto y proporciona el valor de la ratio estandarizada. Cuando un texto, en este caso, contiene menos de 1000 palabras, su ratio es 0. Para evitar esto, en nuestros cálculos hemos seleccionado un cálculo cada 500 palabras y, en caso de que la ratio estandarizada fuese 0, se ha asignado el valor de la TTR de Templin (1957).

5.2. Índice de densidad terminológica: cálculo y problemas

A la hora de realizar el cálculo del índice de densidad terminológica (IDT) resurgen dos problemas, uno ontológico y otro metodológico. El ontológico tiene que ver con el problema de activación del valor especializado de las unidades léxicas en discursos de difícil clasificación. El metodológico tiene que ver con el cómputo de las unidades, palabras o elementos léxicos o gramaticales y la operación matemática que habría que realizar para obtener los valores de la variable cuantitativa.

El primer problema pretendemos resolverlo mediante la teoría de la percepción y el análisis de las respuestas recibidas por los informantes (*vid. infra* III. 6.7., p. 222). No obstante, las respuestas perceptivas de los grupos de informantes se basan en un conjunto de criterios que servirán para calcular el IDT.

Partimos del concepto saussureano de signo como unión de significante y significado, siendo la palabra el signo por excelencia. Precisamente el concepto de *palabra*

es el que constituye el problema ontológico principal, dado que se trata de un concepto viciado desde el punto de vista científico, dadas todas las perspectivas científicas que han abordado dicho signo. Ya en el apartado II. 5.3. (p. 77) mencionamos el problema, considerando la *palabra* como la unidad hiperonímica bajo la cual se hallan las *unidades léxicas* y las *unidades no léxicas*; y la propuesta de algunas teorías que también usan la denominación de *palabra* como opuesta a *término* (nosotros oponemos *término* o *unidad léxica especializada* a *unidad léxica no especializada*).

Pottier en 1972 [1967] (pp. 54-56) distinguía desde un punto de vista estructuralista entre unidades mínimas construidas (las palabras), compuestas por morfemas, y las unidades de comportamiento (es decir, unidades funcionales): las *lexías*, que están compuestas por palabras. Las *lexías* pasan a ser, así, las unidades de análisis del nivel lexicológico, distinguiéndose las *lexías simples* (con términos como *cáncer*, *delito* o *software*) y las *lexías compuestas* (con términos como *enlaces RDF* o *mal de Chagas*). En nuestro trabajo, emplearemos la denominación *forma léxica* para hacer referencia a toda aquella forma lingüística con *lexema* (esto es, cualquier unidad, ya sea simple o compuesta). Esto incluye las categorías gramaticales de sustantivo, adjetivo, verbo y adverbio.

En principio, el cómputo del índice de densidad terminológica ha seguido el siguiente esquema (Rodríguez-Tapia, 2016a, 2016b):

$$\text{Índice de densidad terminológica} = \frac{\text{Unidades de significación especializadas}}{\text{Unidades léxicas}}$$

Tabla 39: Fórmula original del índice de densidad terminológica

Sin embargo, esto implica un problema teórico, que tiene que ver con aquellos términos fraseológicos y con el debate existente en torno al concepto de *unidad léxica* (y relacionado con el de *unidad fraseológica*). Para evitar dicho problema metodológico (y también el ontológico), en un primer momento se consideró emplear alguno de los extractores automáticos de candidatos a término del IULA (como WikiYATE). No obstante, como su propia denominación indica, los resultados obtenidos se corresponderían con candidatos a término, es decir, un conjunto de elementos lingüísticos que, a nivel estadístico, son frecuentes u ocurrentes en los textos relacionados con la temática del corpus. Dado que nuestro interés se halla en la dimensión sociocognitiva del texto especializado y su gradación, decidimos aproximarnos a un método más manual, lo que implica reflexionar sobre el método de cómputo. Así, las primeras ideas propuestas para contabilizar el denominador de la fórmula antes mencionadas contaban con: a) el cómputo

del total de unidades léxicas del texto; b) el cómputo del total de formas léxicas del texto; o c) el cómputo del total de palabras del texto, tanto léxicas como gramaticales.

Para ilustrar la investigación que nos ha guiado a tomar nuestra decisión metodológica, nos serviremos del siguiente ejemplo.

Los *síntomas diagnósticos* son: **disnea** (ahogo), **expectoración anormal**, *tos crónica* y disminución de la *capacidad pulmonar*. También pueden presentarse **sibilancias** (un sonido de silbido al respirar) y **presión torácica**. Era más frecuente en hombres, pero con el aumento del consumo de tabaco y la exposición a la contaminación, actualmente afecta a ambos sexos por igual. Para *diagnosticarla* basta una **espirometría simple**, es una sencilla prueba que mide el *volumen respiratorio*. Otras exploraciones a realizar son la **rx de tórax** y **gasometría arterial** (medición de *oxígeno* y *CO2* en sangre)¹³⁹.

Tabla 40: Ejemplo de texto para cómputo de IDT

En este ejemplo, de forma apriorística y guiada por la intuición como hablante-oyente ideal, desde nuestra aproximación perceptiva, usamos la negrita para marcar las unidades léxicas que consideramos que cualquier usuario clasificaría como términos y la cursiva aquellas unidades léxicas que puede que algunos usuarios considerasen como terminológicas. Dependiendo de los usuarios (y he ahí nuestro principal interés en usar grupos de informantes), el valor especializado de algunas formas léxicas podría ponerse en duda, como podrían ser *síntomas*, *diagnósticos*, *tos*, *diagnosticarle*, *volumen respiratorio*, *oxígeno* o *CO2*.

Para ilustrar el problema metodológico previamente mencionado, a continuación, contabilizamos el total de las marcadas en negrita y en cursiva en los tres casos que mencionamos anteriormente. Esto es: (a) el cómputo del total de unidades léxicas del texto; (b) el cómputo del total de formas léxicas del texto; o (c) el cómputo del total de palabras del texto, tanto léxicas como gramaticales:

		Numerador y denominador	IDT
A	Cómputo de unidades léxicas	Unidades léxicas = 44 Unidades léxicas especializadas = 14	14/44 = 0,31
B	Cómputo de formas léxicas	Formas léxicas = 53 Formas léxicas especializadas = 23	23/53 = 0,43
C	Cómputo de palabras totales	Palabras totales = 89 Formas léxicas especializadas = 24	24/89 = 0,27

Tabla 41: Casos posibles en el cómputo de IDT

¹³⁹ Fuente: PARICIO, Jorge (10 de mayo de 2017): «En España, uno de cada diez ciudadanos de más de 40 años padece EPOC, pero el 73% de las personas que la sufre no lo sabe», *Salud y Medicina*. Disponible en: <http://www.saludymedicina.org/post/en-espana-uno-de-cada-diez-espanoles-de-mas-de-40-anos-padece-epoc-pero-el-73-de-las-personas-que-la-sufre-no-lo-sabe>.

La opción A conlleva demasiados problemas teóricos, puesto que debería realizarse un estudio teórico de todas las unidades léxicas especializadas desde la perspectiva de la fraseología especializada, lo que implicaría un estudio demasiado exhaustivo para los resultados que pretenden conseguirse. El concepto o categoría de unidad fraseológica, frasema o expresión formulaica, entre sus muchas otras denominaciones (Wray, 2000: 465), incluye las colocaciones, locuciones y enunciados fraseológicos (Corpas, 1997). Engloba desde expresiones idiomáticas fijas (Dubois, 1973) a combinaciones frecuentes pero no fijas dentro de la lengua (Benson *et al.*, 1986), entre las cuales se incluye también las empleadas en el discurso especializado (en el «vocabulario técnico», según Ruiz Gurillo, 1998: 12). Algunos de los rasgos distintivos (Corpas, 1998: 167) de la fraseología son la familiaridad perceptiva por parte del hablante, la alta frecuencia de aparición y coaparición, y la fijación o estabilidad morfológica y la especialización semántica. Dado que nuestro interés investigador no se halla en medir el grado de coaparición, fijación o especialización de las unidades lingüísticas, no profundizaremos en si las unidades son parte de una estructura fraseológica o no.

La opción C implica tener que contabilizar las palabras gramaticales en aquellos términos que sean fraseológicos o sintagmáticos (por eso, contabilizamos 22 términos, ya que la preposición *de* en *rx de tórax* de la Tabla 40 se incluye dentro del cómputo) y, desde el punto de vista cuantitativo, al aumentar el número de palabras, el intervalo de valores numéricos sería más reducido.

A nuestro juicio, la opción B es la óptima para reflejar la realidad léxica del texto. En primer lugar, solo se contabilizarían las formas léxicas (por ejemplo, de forma automática mediante una lista de exclusión de palabras gramaticales en WordSmith Tools, teniendo en cuenta el margen de error que puede tener dicho programa). Las formas léxicas son las que representan el significado referencial de la realidad y las que codifican el conocimiento del mundo. Aunque existen unidades terminológicas que están compuestas por más de un elemento léxico y por más de un elemento gramatical, los elementos gramaticales no serían computados en el análisis, por lo que el cálculo se realizaría teniendo en cuenta los elementos léxicos con valor terminológico (o que forman parte de una unidad terminológica) sobre el total de elementos o formas léxicas del texto. Al reducirse el denominador (el total de formas léxicas), aumenta el intervalo de posibles resultados, como ilustramos a continuación usando como base el texto de la Tabla 40 y un número de casos supuestos que nos permiten comparar los resultados:

A/B = IDT				
A	11	23	39	44
B	53	53	53	53
IDT	0,20	0,43	0,74	0,83

A: Número de elementos léxicos especializados (formas léxicas especializadas)

B: Número de formas léxicas totales

Tabla 42: Cálculo de IDT con número de formas léxicas totales y número de elementos léxicos especializados

A/C = IDT				
A	11	23	39	44
C	89	89	89	89
IDT	0,12	0,26	0,44	0,49

A: Número de elementos léxicos especializados (formas léxicas especializadas)

C: Número de palabras totales (léxicas y gramaticales)

Tabla 43: Cálculo de IDT con número de palabras totales y número de elementos léxicos especializados

La Tabla 42 cuenta con un intervalo de valores más amplio (varían de 0,20 a 0,83), lo que facilita comparar resultados entre textos (los valores de la Tabla 43 solo varían de 0,12 a 0,49).

Por estas razones, el cómputo del índice de densidad terminológica cuenta con dos factores:

- El numerador y divisor: consistente en el número de elementos léxicos especializados, entendidos como unidades léxicas, y, específicamente, como formas lingüísticas con lexema, identificados mediante el criterio de percepción del grupo de informantes. No computan las palabras gramaticales. Los tipos de formas léxicas contabilizadas se indican en el apartado III. 6.7.1. (p. 223).
- El denominador y dividendo: el número total de formas léxicas totales, resultado de aplicar una lista de exclusión gramatical a cada texto del corpus mediante WordSmith Tools.

$$\text{Índice de densidad terminológica} = \frac{\text{Formas léxicas especializadas}}{\text{Formas léxicas totales}}$$

Tabla 44: Fórmula del índice de densidad terminológica

5.3. Índice de reformulación: cálculo y problemas

El índice de reformulación (IR) ha sido empleado en trabajos anteriores (Rodríguez-Tapia, 2016b) para mostrar de forma numérica cómo un texto reformula aquellos segmentos que codifican conocimiento especializado. Se basa en el índice de denominación desarrollado por Freixa (2005), que hemos alterado y modificado para emplearlo en nuestro trabajo. Constituye:

aquel valor resultante de dividir el valor del IDT entre el número total local de ocurrencias de los procedimientos de reformulación y multiplicar el valor resultante por un ponderador que nos permita trabajar con cifras cercanas a la unidad (en este caso, mil (1000)) (Rodríguez-Tapia, 2016b: 236).

En otras palabras, el IR muestra la relación que existe entre el número de formas léxicas especializadas por formas léxicas totales del texto con el número de procedimientos empleados para reformular dichos términos. En la fórmula siguiente se usa un multiplicador con valor 1000 para aproximar el resultado de las divisiones a la unidad, ya que, de lo contrario, los valores serían centesimales o milésimales.

$$\text{Índice de reformulación} = \left(\frac{\text{Formas léxicas especializadas}}{\frac{\text{Formas léxicas totales}}{\text{Procedimientos de reformulación formal}}} \right) \times 10^3$$

Tabla 45: Fórmula del índice de reformulación en Rodríguez-Tapia (2016b)

Esta fórmula conlleva que el valor resultante establezca una relación inversamente proporcional, de forma que cuanto mayor sea el valor del índice, menor reformulación existe en el texto; y cuanto menor sea dicho valor, mayor será la reformulación que se observe en el texto. Estos valores no tienen mínimos ni máximos (*cf.* Rodríguez-Tapia, 2016b: 236) si bien cabe realizar varias puntualizaciones:

- a) Cuando no existen procedimientos de reformulación en el texto (es decir, los PRF son igual a 0), el valor del IR también es 0, puesto que no es posible dividir el IDT por 0. Este valor nulo indica que no existe reformulación en el texto.
- b) Cuando existen procedimientos de reformulación, los valores que son muy cercanos al 0, sin poder llegar a ser 0, significan que la terminología del texto ha sufrido mucha reformulación. Cuanto más se aleje el IR de 0, menor reformulación existirá. El valor máximo de IR para cada texto es igual a su IDT (puesto que la operación implicaría dividir el IDT del texto por un único PRF). Ningún valor puede ser negativo.

La combinación de estas dos puntualizaciones conlleva que el cálculo de la media del índice de reformulación sea un valor poco representativo de la realidad reformulativa, ya que, al corresponder 0 con la ausencia de procedimientos de reformulación, y los valores muy cercanos a 0 corresponder con una reformulación muy elevada, el cálculo del valor promedio haría que el valor 0 acercase la media hacia una correspondencia con una reformulación muy elevada, cuando el 0, en realidad, corresponde justo con lo contrario.

Para resolver este problema la fórmula del índice de reformulación debe ser alterada, de forma que puedan invertirse numerador y denominador y eliminar el

ponderador que en principio servía para trabajar con valores cercanos a la unidad. De esta forma, la relación que se establece entre terminología y reformulación es directamente proporcional: cuanto mayor sea el valor del IR, mayor será la reformulación en un texto, y viceversa. La fórmula que usaremos en esta tesis doctoral para calcular el índice de reformulación es la siguiente:

$$\text{Índice de reformulación} = \left(\frac{\text{Procedimientos de reformulación formal}}{\text{Índice de densidad terminológica}} \right)$$

Tabla 46: Fórmula del índice de reformulación

Así las cosas, las puntualizaciones antes mencionadas quedarían modificadas según la siguiente relación:

- a) Cuando no existen procedimientos de reformulación en el texto (es decir, los PRF son igual a 0), el valor del IR también es 0. Este valor nulo indica que no existe reformulación en el texto.
- b) Cuando existen procedimientos de reformulación, los valores que son muy cercanos al 0 significan que la terminología del texto ha sufrido escasa reformulación. Cuanto más se aleje el IR de 0, mayor reformulación existirá en el texto. Ningún valor puede ser negativo.

Los procedimientos de reformulación contabilizados pertenecen a las categorías anteriormente mencionadas para la reformulación parafrástica (*vid. supra* apartado II 10.1.1. y II 10.4., pp. 152 y 159): expansión (explicación y precisión), reducción (denominación y resumen) y variación. Las marcas que sirven para identificar dichos procedimientos son las indicadas en la Tabla 31 del apartado II. 10.4. (p. 159), que repetimos aquí:

PROCEDIMIENTOS DE REFORMULACIÓN PARAFRÁSTICA			
	Expansión	Reducción	Variación
Función	<div style="border: 1px solid black; display: inline-block; padding: 2px;">A</div> » <div style="border: 1px solid black; display: inline-block; padding: 2px;">A'</div>	<div style="border: 1px solid black; display: inline-block; padding: 2px;">A</div> » <div style="border: 1px solid black; display: inline-block; padding: 2px;">A'</div>	<div style="border: 1px solid black; display: inline-block; padding: 2px;">A</div> » <div style="border: 1px solid black; display: inline-block; padding: 2px;">A'</div>
Procedimientos	Definición Explicación Precisión Ejemplificación	Denominación Siglas y acrónimos UT diferente (incluidos epónimos y unidades fraseológicas) Neologismos (incluidos préstamos) Ejemplificación	Variación denominativa ¹⁴⁰ Siglas y acrónimos UT diferente (incluidos epónimos y unidades fraseológicas) Neologismos (incluidos préstamos)

¹⁴⁰ Recordemos que la variación denominativa es un fenómeno de sinonimia.

PROCEDIMIENTOS DE REFORMULACIÓN PARAFRÁSTICA			
Función	<p>Expansión</p>	<p>Reducción</p>	<p>Variación</p>
Marcadores metaling.	<p>ser, denominar(se), llamar(se), significar, recibir el nombre de, referirse a, aludir, decir, querer decir, conocer como, es decir, esto es, también, lo que es lo mismo, dicho de otro modo/de otra manera, (dicho) en otras palabras/en otros términos, o sea, o, i.e., etc.</p>		<p>conocer como, es decir, esto es, también, lo que es lo mismo, dicho de otro modo/de otra manera, (dicho) en otras palabras/en otros términos, o sea, o, i.e., etc.</p>
Mecanismos tipográf.	<p>Coma, dos puntos, raya, paréntesis...</p>		

Tabla 31: Herramienta de análisis de procedimientos de reformulación parafrástica

Para ilustrar el procedimiento de análisis del IR nos serviremos de dos de los ejemplos anteriormente empleados. De nuevo, empleamos la negrita para destacar las formas léxicas que cualquier usuario clasificaría como términos y la cursiva para aquellas formas léxicas que puede que algunos usuarios considerasen como terminológicas, así como el subrayado amarillo para marcar los procedimientos de reformulación:

(34) La insuficiencia cardiaca diastólica **(también denominada «insuficiencia cardiaca con función sistólica conservada»)** [...] ¹⁴.

Tabla 47: Ejemplo de texto para cómputo de IR (1)

En este primer caso, contamos con 8 formas léxicas especializadas de un total de 10 formas léxicas y un único procedimiento de reformulación de variación que emplea tres marcadores reformulativos, que son el paréntesis, las comillas y el segmento *también denominada*. Este ejemplo contaría con un IR de 1,25, por lo que sería un texto poco reformulado.

(35) Los *síntomas diagnósticos* son: **disnea** (ahogo), **expectoración anormal**, *tos crónica* y disminución de la *capacidad pulmonar*. También pueden presentarse **sibilancias** (un sonido de silbido al respirar) y **presión torácica**. Era más frecuente en hombres, pero con el aumento del consumo de tabaco y la exposición a la contaminación, actualmente afecta a ambos sexos por igual. Para *diagnosticarla* basta una **espirometría simple**, **es** una sencilla prueba que mide el *volumen respiratorio*. Otras exploraciones a realizar son la **rx de tórax** y **gasometría arterial** (medición de *oxígeno* y *CO2* en sangre) ¹⁴¹.

Tabla 48: Ejemplo de texto para cómputo de IR (2)

¹⁴¹ Fuente: PARICIO, Jorge (10 de mayo de 2017): «En España, uno de cada diez ciudadanos de más de 40 años padece EPOC, pero el 73% de las personas que la sufre no lo sabe», *Salud y Medicina*. Disponible en: <http://www.saludymedicina.org/post/en-espana-uno-de-cada-diez-espanoles-de-mas-de-40-anos-padece-epoc-pero-el-73-de-las-personas-que-la-sufre-no-lo-sabe>.

En el segundo caso, contamos con un total de 53 formas léxicas en total, 23 formas léxicas especializadas y 4 procedimientos de reformulación (dos de variación y dos de expansión; cada uno con un único marcador de reformulación), lo que conlleva un IR de 1,74, que implica que el texto ha sido más reformulado que el de la Tabla 47.

Existen algunas cuestiones metodológicas que conviene tener en cuenta. A continuación, se relacionarán estas cuestiones empleando los códigos de color empleados por los informantes en los análisis¹⁴², junto con ejemplos extraídos del corpus analizado por parte del grupo de informantes de estudiantes. Se usa el subrayado para destacar los procedimientos de reformulación y segmentos reformulados:

- a) Solo se computan las reformulaciones que afectan a los términos. Aquellos procedimientos que no estén ligados a segmentos terminológicos no se consideran en el cómputo. Por ejemplo, en los ejemplos (36), (37) y (38), los términos a los que afecta la reformulación son *suplementos nutricionales*, *ransomware* y *VPN*. Si en el ejemplo (36) *suplementos nutricionales* no fuese marcado como terminológico, el segmento (SN) no computaría como procedimiento de reformulación formal.

(36) Los **suplementos nutricionales** (SN) se han convertido en una industria que mueve anualmente billones de dólares.

Tabla 49: Ejemplo del texto MED ART09 sobre procedimiento de reformulación que afecta a término

(37) El comunicado añade que no puede evitarse una mayor propagación de este *ransomware*, que limita o impide a los usuarios el acceso al **ordenador** o **ficheros** a menos que paguen un rescate.

Tabla 50: Ejemplo del texto INF NOT01 sobre procedimiento de reformulación que afecta a término

(38) Una **VPN (Virtual Private Network)**, nos permite conectar una o más computadoras a una **red privada** utilizando **Internet**.

Tabla 51: Ejemplo del texto INF BLOG13 sobre procedimiento de reformulación que afecta a término

- b) Aunque exista más de un marcador, solo se tiene en cuenta un único procedimiento (véase ejemplo (34)).
- c) Si el procedimiento vuelve a repetirse, vuelve a computarse. Sin embargo, si el procedimiento se aplica al uso de un acrónimo o sigla, solo computa el procedimiento que vincule el segmento reformulado con el segmento

¹⁴² Véase anexo 2, anexo 6 en papel o apartado III. 6.7.1. (p. 220) para conocer la leyenda de colores.

reformulador y no cuentan las siguientes ocasiones que aparezca dicho procedimiento aislado. Es decir, si en un texto se usa una sigla y se repite a lo largo del texto constantemente, solo se cuenta 1 procedimiento. En el ejemplo (39) se observan dos PRF por variación, que son (*ERC*) y (*FG*), que reformulan respectivamente a *enfermedad renal crónica* y *filtrado glomerular*. En secciones más avanzadas del mismo texto se vuelve a emplear *ERC*, esta vez en su forma abreviada, y, aunque marcada como terminológica, dicho segmento no computa como reformulación.

(39) La enfermedad renal crónica (*ERC*) se define como la disminución de la función renal, expresada por un filtrado glomerular (*FG*) inferior a 60 ml/min/1,73 m² o como la presencia de daño renal de forma persistente durante al menos 3 meses [1].

[...]

La *ERC* tiene una alta prevalencia, y afecta a un 5-10% de la población mundial [3]. De acuerdo con el estudio *EPIRCE* publicado en 2010, la prevalencia de la *ERC* grados 3-5 en España se sitúa en el 6.8% de la población [...].

Tabla 52: Ejemplo del texto MED TES02 sobre procedimiento de reformulación que afecta al término

d) Aquellos segmentos que reformulan un concepto como valor porcentual de un fenómeno, aunque sea terminológico, no son computados, como en el caso del ejemplo (40):

(40) Los consumidores declaran un total de 575 productos diferentes, de los cuales podemos encontrar 38 categorías, donde existe una marcada preferencia por los batidos proteicos en polvo (39.7%), polivitamínicos (16.5%) y agentes lipolíticos (13.4%).

Tabla 53: Ejemplo del texto MED ART09 sobre IR

La relación de procedimientos de reformulación formal es ligeramente diferente en cada informante, puesto que los PRF están ligados a la terminología. El número de PRF viene determinado por el investigador de la tesis cuando este identifica que un procedimiento está asociado a alguna forma léxica precedente o no.

5.4. Relación discursiva emisor-destinatario

La distinción de perfiles se basa en la competencia cognitivo-comunicativa de los interlocutores, lo que supone un importante problema metodológico, puesto que existe gran dificultad para evaluar dicha competencia en los instruidos y legos (incluso, a veces, en los especialistas) exclusivamente a través de material textual. De nuevo, es necesario adoptar

una perspectiva apriorística e idealizada de cada perfil, puesto que pueden existir leves variaciones que esbocen un grado de posibilidades. De esta forma, puede comprobarse que los criterios *tipo de texto y relación emisor-destinatario* son interdependientes, pues es necesario contar con información sobre uno para poder catalogar el otro. Para identificar tanto el emisor como el destinatario, es necesario recurrir a la firma del texto o a las justificaciones sobre el tipo de emisor y destinatario en la fuente de este (en caso de que existan: por ejemplo, las que pueden encontrarse en las revistas especializadas o de divulgación, los artículos de opinión, las solicitudes, etc.). De esta forma, es necesario recurrir a información explícita que permita discriminar los tipos de interlocutores. Sin embargo, pueden encontrarse casos en los que la información lingüística y textual no sea suficiente para distinguir el tipo de destinatario, sobre todo cuando el emisor es un lego.

La categoría de lego como destinatario supone algunos problemas teóricos, puesto que implica que un texto sea comprensible por el resto de categorías, es decir: un texto orientado a un lego es comprensible por un especialista y por un instruido pero un texto orientado a un especialista es solo comprensible por un especialista (de ahí que defendamos que se trate de una variable cualitativa ordinal en el apartado III. 5., p. 198). Como menciona Monterde Rey (2007: 42) (la negrita es nuestra):

El grupo de los legos es muy heterogéneo, pues revistas de este tipo son adquiridas por un público que puede abarcar desde el que no sabe absolutamente nada sobre el tema, pasando por el aficionado que tiene algunos conocimientos sobre él, hasta el profesional del campo. Estas revistas, sin embargo, como me informaron sus redactores, **se adaptan al nivel de conocimiento del que sabe menos sobre el tema y presentan artículos que puedan interesar al lego [...]**.

Así, por ejemplo, ¿cómo es posible determinar que un artículo de opinión de un periódico va dirigido a un lego y no a un instruido o a un especialista? A nuestro juicio, si se explicita el destinatario, se toma dicho perfil; si no se explicita, se considera que se dirige al público general, adoptando un perfil lego¹⁴³. Por ejemplo, el artículo de opinión de un lector en un periódico que trata un tema médico y se dirige al ministro de sanidad podría considerarse un texto lego-especialista¹⁴⁴ pero un artículo de opinión emitido por un interlocutor del público general en el que se elabora una queja sobre el ministerio de sanidad

¹⁴³ Debe tenerse en cuenta que el público general no necesariamente se ajusta o corresponde con el perfil de usuario lego.

¹⁴⁴ En este caso, igual que en muchos otros, hay que tener en cuenta que, a veces, la persona encargada del ministerio de sanidad no pertenece a este sector, es decir, no es especialista en medicina o sus campos relacionados, pero sí puede considerarse especialista en los asuntos que afectan al ministerio de salud.

puede entenderse que va dirigido al público general, por lo que sería considerado un texto lego-lego.

A continuación, presentamos una propuesta *ad hoc* de ejemplos de tipos de textos y los procedimientos empleados para discriminar los tipos de perfiles:

	Estrategias de identificación de:	
	Emisor	Destinatario
Especialista-especialista <i>un artículo científico</i>	A través de la firma del artículo, normalmente acompañado de la institución para la que trabaja/investiga.	A través de la justificación y motivación científica de la revista, que normalmente suele dirigirse a profesionales o iniciados de la materia, es decir, especialistas o instruidos. El conocimiento codificado en el texto solo es decodificable por alguien que domine el sistema conceptual de la materia.
Especialista-instruido <i>un artículo divulgativo en un libro</i>		A través de la justificación y motivación científica del libro donde aparece el artículo, que suele dirigirse a instruidos en la materia, puesto que suelen constituirse como introducciones a dichos temas. El conocimiento codificado en el texto solo es decodificable por alguien que domine los conceptos básicos de la materia.
Especialista-lego <i>un folleto informativo</i>		A través de la justificación y motivación de la divulgación científica de la revista, manual, guía o folleto donde aparece el texto. El conocimiento codificado en el texto es comprensible para el gran público.
Instruido-especialista <i>un ensayo o proyecto universitario</i>	A través de la firma del documento. Suele ser un especialista en formación, un interesado en el tema o un profesional de la comunicación especializada.	A través del emisor, del tipo de texto y del conocimiento codificado en el texto, que solo es comprensible por otros instruidos o especialistas.
Instruido-instruido <i>un artículo divulgativo</i>		A través del tipo de texto en el que se inserta el contenido, normalmente una revista o periódico de divulgación especializada. El conocimiento que codifica el texto es comprensible por un especialista o por un instruido.
Instruido-lego <i>una entrevista</i>		A través del emisor, del tipo de texto y del conocimiento codificado en el texto, que es comprensible para el público general.
Lego-especialista <i>una consulta en revista divulgativa</i>	A través del tipo de texto, de la firma del texto y del conocimiento codificado en él.	A través de la función del texto y del tipo de texto en el que se inserta el contenido, normalmente una revista o periódico de divulgación especializada. El conocimiento que codifica el texto es comprensible por el público general.
Lego-instruido <i>una consulta en foro</i>		A través de la función del texto y del tipo de texto en el que se inserta el contenido, (en un foro suele participar cualquier persona del público general, entre los que se incluyen especialistas e instruidos). El conocimiento que codifica el texto es comprensible por el público general.
Lego-lego <i>Un artículo de opinión en periódico generalista</i>		A través de la función del texto y del tipo de texto en el que se inserta el contenido, normalmente una revista o periódico generalista o de divulgación especializada. El conocimiento que codifica el texto es comprensible por el público general.

Tabla 54: Estrategias de identificación de la relación emisor-destinatario en los textos del corpus [elaboración propia]

Hay casos, como los textos emitidos por legos, en los que es difícil delimitar el tipo de destinatario incluso recurriendo a la función textual o el conocimiento codificado, por lo que a veces solo cabe recurrir a la percepción psicológica, o, mejor aún, intuición técnica,

del investigador. En esta investigación, analizaremos dicha percepción a través de los resultados que esbozen los grupos de informantes (*vid. infra* apartado III. 6.7., p. 222.).

5.5. Las clases de los textos del corpus

La clase del texto es la variable dependiente protagonista de nuestra tesis doctoral. Usamos la denominación *clase* desde la perspectiva de las ciencias de la computación y no desde la lingüística. Constituye el tipo de texto según el grado de especialización y puede adquirir la forma de tres variables ordinales diferentes: el texto especializado, el texto semiespecializado y el texto no especializado. En trabajos anteriores (Rodríguez-Tapia y Camacho-Cañamón, 2018) nos acercamos a la definición de estas clases textuales:

Texto especializado:

productos predominantemente verbales de registros comunicativos específicos, que se refieren a temáticas propias de un dominio de especialización, y que responden a convenciones y tradiciones retóricas específicas; por lo tanto, en dependencia del tipo de disciplina pueden ser más o menos dependientes de la cultura y la época dada [...]. Los textos especializados se realizan en clases textuales específicas del discurso de especialización (artículo de investigación, ponencia, artículo de divulgación científica, comunicados científicos a la prensa, etc.). Concebimos el ámbito de los textos especializados en términos de *continuum* (Ciapuscio y Kuguel, 2002: 43).

Texto semiespecializado:

aquel texto redactado por un especialista y dirigido indistintamente a un público lego o especialista, en el que predominan, de forma independiente o en conjunto, las funciones representativa y comunicativa, cuyo índice de densidad terminológica no es especialmente alto (0,12 de media) y que se asocia a la superestructura de los glosarios, artículos divulgativos y artículos científicos (Rodríguez-Tapia, 2016b).

Texto divulgativo: de forma opuesta al texto especializado, se trata del texto que no codifica conocimiento especializado, es producto de los intercambios comunicativos ordinarios entre interlocutores y se configura discursivamente hacia un receptor no especialista.

Estas caracterizaciones parten de estudios teóricos previos o de aproximaciones empíricas a la identificación del *continuum* que han usado como criterio métodos matemáticos principalmente. Descubiertos los principales problemas teóricos y metodológicos de dichos métodos (Rodríguez-Tapia, 2016a, 2016b; Rodríguez-Tapia y Camacho-Cañamón, 2018), es decir, la determinación apriorística del grado de

especialización textual sin conocer con exactitud la realidad de las características esenciales del *continuum*, tomamos la decisión de emplear los fundamentos de la teoría de la percepción. Así, consideramos que la clase textual es un valor que determina el hablante según su contacto y experiencia con el mundo. Por ello, la clase textual es una de las variables que proporcionan los grupos de informantes. Se constituye como una decisión teórica fruto del conocimiento lingüístico y del análisis textual de cada objeto del corpus por parte del informante; en definitiva, constituye el percepto tras el resultado de la percepción sintética. De esta forma resolvemos dos problemas:

- a) Asignar de forma matemática la clase textual, lo que impide describir con precisión los límites difusos entre clases.
- b) Asignar de forma apriorística la clase textual, que supone, de alguna forma, arriesgarse a confundir las clases. En otras palabras, decidir de antemano si un texto pertenece a una clase textual según su grado de especialización implica que, a la hora de analizar los resultados y proyectar conclusiones, exista el riesgo de atribuir elementos característicos a una clase a la que el conjunto de textos asignados en realidad no pertenece.

Para justificar esta decisión es muy interesante contar con la cita de Caravedo (2014: 125), que ya avanzamos en el apartado II. 2.4. (p. 172):

Sostener la subjetividad no significa minusvalorar o descartar las entidades mencionadas considerándolas como no susceptibles de conocimiento. Todo lo contrario: los hechos calificados de subjetivos o de relativos al sujeto pueden ser —o deben ser—abordados de modo objetivo; en los términos de Searle son pasibles de objetividad epistémica, y en este sentido constituyen materia del discurso científico [...]. Pero tal paradoja se disuelve cuando, identificados determinados rasgos subjetivos, se reconoce que tales rasgos están aceptados colectivamente y que presentan comportamiento regulares y motivados, esto es, que no constituyen producto del azar, de la imaginación o de la preferencia idiolectal del individuo.

Desde nuestro punto de vista, contar con el análisis de informantes multiplica en gran número los resultados asignados a clases textuales y no dependen, en este caso, del autor de la tesis, por lo que identificar cierto grado de homogeneidad en los resultados permitiría reconocer ciertos convencionalismos a partir de las decisiones subjetivas del grupo de informantes.

6. GRUPOS DE INFORMANTES

6.1. Criterios de selección de los grupos de informantes

El análisis del corpus se basa en los fundamentos sociocognitivos de la valoración y la percepción (*vid. supra* apartado I. 2., p. 17). Por ello, contamos con tres grupos de informantes (más un cuarto grupo, formado por el análisis del autor de la tesis, que representa el hablante oyente ideal) que, a partir del margen y discreción que permite su percepción lingüística, se encargan de extraer diferentes datos (*vid. infra* apartado III. 6.7., p. 222.) sobre el corpus. El perfil del informante viene determinado por la finalidad del modelo teórico de descripción del fenómeno del grado de especialización, que puede tener repercusión en la capacidad de previsión y solución de problemas de la lingüística aplicada. A continuación, se esbozan las variables que permitirán describir a cada grupo de informantes.

6.1.1. Género

Las investigaciones sociolingüísticas han determinado que el sexo (o, con mayor precisión, el género¹⁴⁵), si bien puede influir en la variación lingüística, lo hace subordinado a otros factores sociales diferentes (como el nivel sociocultural o el estilo) (Fasold, 1990: 223; Martín Zorraquino, 1993: 115-126; Moreno, 2008: 41). De hecho, por ejemplo, las diferencias lingüísticas que se basan en el género van variando a lo largo de la vida del individuo, por lo que la variable del género se subordina a la de la edad (Moreno, 2008: 47).

6.1.2. Edad

La edad es uno de los factores sociales que determinan con más fuerza el uso lingüístico y la variación (Moreno, 2008: 47). Con el transcurso del tiempo, la edad «va determinando y modificando los caracteres y los hábitos sociales de los individuos, incluidos los comunicativos y puramente lingüísticos» (*ídem*), lo que hace posible distinguir diferentes etapas en la vida lingüística de un hablante (aunque sin que exista

¹⁴⁵ Es conveniente tener en cuenta el género como interpretación sociocultural del texto: «el género sociocultural se opone al sexo en tanto en cuanto el sexo es una característica biológica que viene dada prácticamente desde el momento de la concepción del nuevo ser, mientras el género es una dimensión sociocultural que el individuo adquiere al ser socializado» (Moreno, 2008: 44).

consenso sobre sus características ni límites). Esto conlleva la distinción de grupos generacionales, que influyen en el «el uso de ciertas variables o rasgos lingüísticos que sirven para marcar distancias entre niños y jóvenes, entre jóvenes y adultos» (ibíd.: 51), siendo estos rasgos lingüísticos el léxico, la fraseología o el discurso principalmente (ibíd.: 52).

Por ello, es conveniente tener en cuenta que los informantes deben formar parte de un grupo lo suficientemente estable en lo que se refiere a su variedad idiolectal y sociolectal, es decir, que cuenten con madurez en el uso social de la lengua. Labov (1966) explica la adquisición de un conjunto de normas de habla por parte de los hablantes: en la adolescencia empiezan a desarrollar normas evaluadoras y al alcanzar la edad de 17 o 18 años comienzan a ser capaces de discriminar la significación social de los modos de hablar (propios y externos) y de los usos prestigiosos de la lengua.

6.1.3. Clase: nivel educativo y profesión

La clase social o sociocultural es una de las variables más difíciles de definir en un estudio sociolingüístico debido a la necesidad de emplear enfoques multidimensionales o multicriterio y a que no existen límites nítidos entre los diferentes estratos que pueden diseñarse para medir esta variable (Moreno, 2008: 53-54), por ejemplo, para delimitar los ingresos de un individuo, su nivel sociocultural, etc. La distancia social ha sido vinculada proporcionalmente a diferencias sociolingüísticas.

Para evitar estos problemas, Sankoff y Laberge (1978) propusieron el concepto de *mercado lingüístico*, que considera que las conductas lingüísticas dependen de las actividades socioeconómicas de los individuos, si bien «resulta complicado demostrar la objetividad de los procedimientos seguidos para asignar a cada hablante un índice de integración en el mercado lingüístico» (Moreno, 2008: 57). Es innegable la vinculación que existe entre la formación, el estatus, el poder o la profesión, que, a nuestro juicio, integran el concepto de clase (social). Por esta razón, y dado que nuestro grupo de informantes comparte todas las características, nos centraremos en dos aspectos concretos para definir la clase a la que pertenecen: su nivel educativo y su profesión.

Cuando nos referimos al nivel educativo hacemos referencia al tipo de formación académica (analfabetismo, enseñanza primaria, enseñanza secundaria o enseñanza universitaria) del informante, que determina de forma directa la variación lingüística:

cuanta más formación posea un hablante, mayor uso de las variantes prestigiosas o normativas hará (Moreno, 2008: 61).

La profesión también influye de forma directa sobre la variación lingüística (Alvar, 1976: 114), siendo «el léxico donde encuentra su expresión más evidente: los rasgos que caracterizan el habla de ciertos grupos profesionales son rasgos mayoritariamente léxicos» (Moreno, 2008: 67).

6.2. Materiales de trabajo

El material analizado por los cuatro grupos de informantes (es decir, tres grupos más el autor de la tesis) corresponde con el corpus textual compilado, que reúne 315 textos. Los textos han sido asignados a los informantes de forma totalmente aleatoria mediante un algoritmo diseñado en MO Excel. Todos los textos han sido asignados y analizados. Para el análisis de cada texto es esencial la lectura atenta, exhaustiva y perceptiva de la totalidad del texto. El período de trabajo abarca desde septiembre de 2017 a febrero de 2018. Los resultados proporcionados por estos informantes pueden encontrarse en el anexo 1A y en el anexo 1B.

El análisis del corpus ha seguido un protocolo específico con un marco teórico concreto relacionado en un libro de instrucciones que ha recibido cada grupo de informantes y que atiende a los criterios que pueden guiar la percepción para contabilizar las unidades. Dichas instrucciones pueden consultarse en el anexo 2.

Cada informante cuenta con los textos que se les han asignado y un libro de Excel que sirve de base de datos en la que debe introducir la información que indicamos en este apartado. Esta base de datos requiere la información siguiente: identificación del informante, número del texto asignado al informante, tipos de términos según su percepción, tipos de emisor y destinatario y subcategorías correspondientes y clase de texto.

Esta base de datos no proporciona cierta información al grupo de informantes, de forma que no pueda influir en su toma de decisiones. Esta información censurada se corresponde con el tema del texto y el tipo textual, que, recordemos, constituyen las variables de control de nuestra investigación y los criterios principales de compilación del corpus, así como el número de palabras del texto, la TTR, el código del texto, el número de procedimientos de reformulación y los índices.

A continuación, incluimos una tabla que resumen las características de los grupos de informantes. En el anexo 5 se muestra un ejemplo de consentimiento de uso de datos estadísticos para esta investigación.

Grupo de informantes	Número de informantes	Número de hombres	Número de mujeres	Rango de edad del grupo	Número de textos asignados por informante
Estudiante	12	1 (8,3 %)	11 (91,6 %)	20-22	26-27
Usuario medio de la lengua	25	11 (44 %)	14 (56 %)	24-53	12-13
Traductor	35	8 (22 %)	27 (77 %)	27-45	9
Investigador	1	1	0	25	315

Tabla 55: Resumen de grupos de informantes

Estos cuatro grupos están formados por hablantes nativos de español y guardan relación con la población real de la que se obtienen los datos:

- a) La población de estudiantes matriculados en el Grado de Traducción e Interpretación de la Universidad de Córdoba durante el curso 2016-2017 corresponde con un 17 % de hombres y un 83 % de mujeres.
- b) Según los datos proporcionados por el Instituto Nacional de Estadística (INE) para el segundo semestre de 2016 (período en el que se realizó el trabajo de análisis con los informantes), la población española cuenta con un 49 % de hombres y 51 % de mujeres. En esta población también se incluyen inmigrantes y hablantes no nativos; no obstante, consideramos el dato del INE representativo de los hablantes nativos de español en España.
- c) Según los datos proporcionados mediante correo electrónico por la Asociación Española de Traductores, Correctores e Intérpretes (ASETTRAD), el número de asociados es representativo de la totalidad de traductores en activo de España, correspondiendo con un 25 % de hombres y un 75 % de mujeres.
- d) El investigador adopta la forma de hablante oyente ideal generativo, que es representativo de todos y cada uno de los hablantes de una lengua.

6.3. Grupo 1: estudiantes

El perfil del primer grupo de informantes corresponde con un potencial traductor que pueda enfrentarse a la traducción de un texto sobre las temáticas de las que versa el corpus. Desde el punto de vista profesional, el traductor no tiene por qué ser especialista ni instruido en la materia de la que versa un texto, pero sí, al menos, ser capaz de reconocer el valor especializado de ciertas unidades léxicas, los procedimientos de reformulación y los tipos de usuarios emisores y destinatarios a partir de la información lingüística que proporciona el texto.

Los criterios que determinan el perfil del grupo son la necesidad de no ser especialista o instruido en medicina, informática o derecho, tener conocimiento o estar formándose en traducción y, por último, contar con conocimientos sobre análisis lingüístico y sobre tipologías textuales, para lo cual se han seguido las instrucciones del anexo 2.

El grupo de informantes está constituido por 12 personas, seleccionadas entre el alumnado del tercer y cuarto curso del Grado de Traducción e Interpretación y del Doble Grado de Estudios Ingleses y Traducción e Interpretación de la Universidad de Córdoba, que hayan superado la asignatura de Terminología con más de un 8 de calificación final y que colaboren con el área de Lingüística General de la mencionada universidad.

A continuación, analizaremos en concreto algunas variables del grupo de informantes, como son el sexo, la edad o la clase social. Debe recordarse que el grupo de informantes no es objeto de estudio sociolingüístico en sí, sino que nuestro interés se halla en la percepción sociocognitiva de dicho grupo sobre un corpus textual.

Así, el género predominante de nuestro grupo de informantes es el femenino (11 mujeres). Desde nuestro punto de vista, el género del grupo de informantes no determina en ninguna forma los resultados en cuanto a la percepción terminológica que aquí buscamos. Además, esta muestra corresponde proporcionalmente con la población de la que obtenemos los datos (amplia mayoría de mujeres y reducido número de hombres).

La valoración y conciencia de estos usos prestigiosos, cercanos al modelo normativo, es alcanzada por las generaciones más jóvenes y mejor instruidas según Moreno (2008: 47), de ahí que tomemos la decisión de que este grupo de informantes esté formado por estudiantes universitarios de entre 20 y 22 años.

En este caso, nuestro grupo de informantes se integra en el grupo de enseñanza universitaria.

6.4. Grupo 2: usuario medio de la lengua

El segundo grupo de informantes incluye a varias personas sin formación específica en análisis lingüístico: se trata de usuarios medios de la lengua. El principal interés investigador para incluir a dicho grupo de informantes tiene que ver con la forma de clasificar los términos y los grados de especialización en diferentes tipos textuales a partir de la información lingüística que proporciona cada texto. Este grupo de informantes está formado por personas que han finalizado la enseñanza secundaria postobligatoria y pertenecen a ámbitos profesionales variados, como la enfermería, la medicina, el derecho, la administración y dirección de empresas, el arte dramático o la historia del arte.

Este grupo está compuesto por 25 informantes, de los cuales la mayoría son mujeres (14 informantes) con un rango de edad de los 24 a los 53 años. Esta proporción también se ajusta a la población española, donde las mujeres son más numerosas que los hombres.

6.5. Grupo 3: traductor profesional

El tercer grupo de informantes es similar al primero: los sujetos han recibido formación en análisis lingüístico y sus integrantes se dedican profesionalmente a la traducción. El contacto diario con la terminología y con diversos tipos textuales de diferentes temas nos permiten considerar que la percepción del grado de especialización, de los tipos de usuarios emisores y destinatarios y de los tipos de términos percibidos en cada texto pueden divergir a los de los grupos de informantes de estudiantes y de usuarios medios de la lengua, ya que el aumento de la frecuencia de la terminología puede afectar a la capacidad perceptiva de estos usuarios de la lengua.

Este grupo está compuesto por 35 informantes, de los cuales la mayoría son mujeres (27 informantes) con un rango de edad que oscila entre los 27 y los 45 años.

6.6. Grupo 4: investigador

El cuarto grupo de informantes corresponde con el autor de la tesis doctoral. En este caso, se sigue el punto de vista generativo (Chomsky, 1965: 18, 24) en el que contrastaremos la percepción de un hablante-oyente ideal (el autor) con el resto de los informantes.

La gramática generativa tiene por objetivo explicar la competencia lingüística. Se entiende por competencia lingüística la capacidad mental de un nativo ideal (un hablante-oyente ideal) para *generar* oraciones gramaticales, es decir, bien formadas. En el paradigma generativo, el hablante nativo se considera ideal porque conoce la lengua de forma perfecta, se inserta dentro de una comunidad de habla homogénea y no se ve afectado por limitaciones, como la memoria o las distracciones, de manera que es representante de una comunidad de habla *ideal*¹⁴⁶. Este hablante-oyente ideal no corresponde con ningún hablante concreto, sino que se infiere que se trata de cualquier hablante de la lengua que pueda generar oraciones bien formadas (Gil, 2007: 36), «algo así como una entidad platónica que no denota a nadie pero que también denota a todos» (Gil, 2007: 86).

En 1957, Chomsky usa la intuición del hablante-oyente ideal («his latent, unconscious knowledge of universal grammar», Chomsky, 2006 [1968]: 112) para determinar si las oraciones son gramaticales o no (Chomsky, 1985 [1957]: 2, 56, 81). Los datos de la intuición, que son introspectivos (Alcaraz Varó, 1990: 91), son los datos que permiten explicar la competencia, de ahí que hablemos de que el paradigma generativista es un paradigma mentalista.

Así, en conjunción, puede defenderse que «la competencia es la posesión intuitiva que tiene todo hablante nativo del sistema implícito de su lengua. En palabras de Chomsky es el «sistema finito de reglas flexibles constitutivas del lenguaje humano» (Alcaraz Varó, 1990: 98).

En esta tesis doctoral, el cuarto grupo de informantes está compuesto por el autor de la investigación, que adopta el perfil de hablante-oyente ideal, guiado por su intuición como hablante nativo, que servirá como usuario de control, con el que también se comparará el resto de grupos de informantes.

6.7. Datos que proporcionan los grupos de informantes

En el apartado III. 2.7. (p. 179) justificábamos que nuestro interés se hallaba en trabajar con la percepción sociocognitiva de varios grupos de informantes diferentes para simular cómo el hablante no especialista percibe los textos según su grado de

¹⁴⁶ Literalmente, según Chomsky (1965: 3): «linguistic theory is concerned with an ideal speaker-listener in a completely homogeneous speech community, who knows its language perfectly and is unaffected by such grammatically irrelevant conditions as memory limitations, distractions, shifts of attention and interest, and errors (random or characteristic) in applying his knowledge of the language in actual performance».

especialización (en este caso, con una aproximación a la traducción). Dado que contamos con 315 textos y cuatro grupos de informantes, contaremos con un total de 1260 textos analizados. A continuación, señalamos la información que proporcionan los grupos de informantes.

6.7.1. Tipos de términos y cómputo

La identificación de los términos sigue la concepción teórica de la TCT, matizada por la propia percepción subjetiva del informante desde la teoría sociocognitiva. Los informantes no solo enumeran el conjunto de términos que identifican, sino que clasifican los términos encontrados en tres categorías diferentes¹⁴⁷ según cuenten con:

- a) **FLE a)**: Conocimiento de la denominación y del concepto, no necesariamente con precisión, sino posiblemente de forma laxa. El informante lo reconoce como frecuente en el discurso común. Corresponde con los denominados *términos délficos*.
- b) **FLE b)** Conocimiento de la denominación, pero no del concepto o conocimiento muy pobre del concepto. El informante lo reconoce como poco frecuente en el discurso común. Corresponde con lo que nosotros denominaremos *términos semicrípticos*.
- c) **FLE c)** Conocimiento nulo de la denominación o del concepto. El informante lo reconoce como poco frecuente en el discurso común. Corresponde con los denominados *términos crípticos*.

El siguiente es un ejemplo extraído del texto INF FOR04 de un informante del grupo de estudiantes, que muestra la clasificación realizada: **turquesa** para el tipo a), **verde azulado** para el tipo b) y **gris** para el tipo c).

(41) Aunque tu procesador se pueda **overclockear**, tu **placa base** no permite **OC**, con lo que no podrás hacer **OC** ni al **procesador** ni a la **RAM**, como mucho a la **GPU**, pero eso bajo tu responsabilidad, las **GPU's** son mucho más sensibles que los procesadores o la **RAM** con el **OC**, y yo por lo menos no lo suelo recomendar, puedes dejar inservible la **gráfica** si te pasas mucho con las **frecuencias** o el **voltaje**, y encima tu **gráfica** se suele calentar bastante sin hacerle **OC** ni nada...

Tabla 56: Ejemplo del texto INF FOR04 sobre selección y clasificación de términos por informantes

¹⁴⁷ Esta clasificación es una adaptación de la estrategia de indagación de Caravedo (2014: 238 y 254) en torno al conocimiento del informante sobre significante y significado.

Esta información es codificada en cada texto por parte de los informantes siguiendo las instrucciones facilitadas según el anexo 2.

Cabe destacar la relevancia del campo temático del texto, que consideramos que actúa como condicionante, en ciertos casos, del valor de especialización de las unidades léxicas. Así, los cuatro grupos de informantes coinciden en no clasificar como terminológicas unidades como *ferrocarriles*, *servicios de seguridad de carreteras* o *sistemas*, subrayadas en amarillo en el ejemplo siguiente, que habrían sido seleccionadas como unidades léxicas especializadas (no por todos los informantes) en textos de otra temática, como, por ejemplo, la relacionada con los transportes:

(42) En declaraciones a la cadena británica BBC, el joven informático aseguró que nuevas versiones del "malware", cuya propagación él ayudó a desactivar, comenzarán difundirse de forma "inminente", "con bastante probabilidad el lunes".

"La primera versión de 'WannaCrypt' se pudo detener, pero la versión 2.0 probablemente corregirá ese fallo", afirmó en Twitter el experto, conocido como "MalwareTech".

En Reino Unido, el ciberataque obligó a la Sanidad británica a cancelar operaciones, retrasar citas y desviar ambulancias, según el Ministerio del Interior, que ha asegurado que los historiales de los pacientes no se vieron comprometidos. En tanto, en Rusia, uno de los países más golpeados por el ataque, los equipos en los principales bancos del país, la compañía estatal de ferrocarriles, los servicios de seguridad de carreteras y los sistemas del Ministerio de Interior quedaron inutilizados por horas.

Grandes empresas como la española Telefónica y las automotrices Nissan, en el Reino Unido, y Renault, en Francia, también se vieron afectadas por el ciberataque del viernes pasado.

Tabla 57: Ejemplo del texto INF NOT11 sobre selección y clasificación de términos por informantes

6.7.2. Perfil de emisor y destinatario

Una vez finalizadas las fases de lectura de cada texto y de identificación de tipos de términos, los informantes indican, de acuerdo con su criterio perceptivo sobre la totalidad del texto, qué tipo de emisor y destinatario (y sus correspondientes subperfiles) según su competencia cognitivo-comunicativa observa en el texto. Recordemos que los valores que pueden adquirir estas variables las encontramos en la Tabla 1 y fueron debatidos en el apartado II. 2.4. (p. 46):

Especialista	Investigador
	No investigador
Instruido	Semiespecialista de otra disciplina
	Aprendiz de especialista

Lego formado	
Lego	Lego no formado

Tabla 1: Perfiles y subperfiles de tipos de interlocutores según su competencia cognitivo-comunicativa

6.7.3. Clase textual

Como conclusión de su análisis, es decir como resultado de su percepción sintética, los informantes determinan a qué clase consideran que pertenece el texto analizado, a saber, texto especializado, texto semiespecializado y texto no especializado, según la información debatida en el apartado III. 5.5. (p. 214).

6.7.4. Resumen

En resumen, esta es la información que proporcionan los grupos de informantes:

Variable por cada texto	Valores		
Número de tipos de término a)	Numéricos		
	Valor mínimo	Valor promedio	Valor máximo
	0	89	897
Número de tipos de término b)	Numéricos		
	Valor mínimo	Valor promedio	Valor máximo
	0	59	511
Número de tipos de término c)	Numéricos		
	Valor mínimo	Valor promedio	Valor máximo
	0	55	698
Emisor y destinatario	Cualitativos ordinales Especialista Instruido Lego		
Subemisor y subdestinatario	Cualitativos ordinales Especialista investigador Especialista no investigador Instruido semiespecialista Instruido aprendiz de especialista Instruido lego formado Lego no formado		
Clase textual	Cualitativos ordinales Texto especializado Texto semiespecializado Texto no especializado		
Tipo de informante	Cualitativos nominales Estudiante Usuario medio Traductor Investigador		

Tabla 58: Resumen de información proporcionada por los grupos de informantes

6.8. Datos que proporciona el investigador

Existe un conjunto de datos que queda oculto a los grupos de informantes, bien porque podría condicionar sus decisiones de percepción y clasificación, bien porque es resultado de operaciones matemáticas que usan los datos proporcionados por los informantes.

El tema y el tipo textual, que son las variables de control del corpus, no son comunicadas a los informantes. En cuanto a los valores numéricos básicos, como el número de palabras, el número de formas léxicas, el porcentaje de formas léxicas con respecto al total de palabras, el número de tipos de formas léxicas y la TTR son calculados *a priori* mediante WordSmith Tools.

El índice de densidad terminológica de cada texto se calcula atendiendo a las operaciones indicadas en III. 5.2. (p. 202), para lo cual se usa la suma de formas léxicas de los distintos tipos en cada texto. Además, también se calcula el IDT para cada uno de los tipos de término, para comprobar el grado de densidad terminológica de los términos más conocidos o frecuentes en cada tipo de texto, tema o clase.

Los diferentes procedimientos de reformulación vinculados a la terminología (es decir, según correspondan con procedimientos de reducción, expansión o variación) según los criterios esbozados en el apartado II. 5.2. y en la Tabla 31 son contabilizados por el investigador en los 1260 textos. Tras ello, se calcula el índice de reformulación según las operaciones en el apartado III. 5.3. (p. 206), tanto global como por cada tipo de reformulación.

En resumen, esta es la información que proporciona el investigador y que no conocen los grupos de informantes:

Variable por cada texto	Valores		
Tema	Cualitativos nominales		
	Derecho	Informática	Medicina
Tipo de texto	Cualitativos nominales		
	Artículo en revista científica		
	Blog		
	Entrevista		
	Noticia		
	Participación en foro		
	Revista o web de divulgación especializada		
	Tesis doctoral		
Número de palabras	Numéricos		
	Valor mínimo	Valor promedio	Valor máximo
	220	1654	8343
Número de formas léxicas	Numéricos		
	Valor mínimo	Valor promedio	Valor máximo
	102	862	4424
	Numéricos		

Variable por cada texto	Valores		
Porcentaje de formas léxicas con respecto al total de palabras	Valor mínimo	Valor promedio	Valor máximo
	43 %	52 %	60 %
Tipos de formas léxicas	Numéricos		
	Valor mínimo	Valor promedio	Valor máximo
	80	462	1669
TTR	Numéricos		
	Valor mínimo	Valor promedio	Valor máximo
	30,809	61,248	89,542
TTR estandarizada	Numéricos		
	Valor mínimo	Valor promedio	Valor máximo
	42,400	65,554	89,542
Total de términos (total de FLE)	Numéricos		
	Valor mínimo	Valor promedio	Valor máximo
	2	203	1656
Índice de densidad terminológica	Numéricos		
	Valor mínimo	Valor promedio	Valor máximo
	0,008	0,217	0,576
Índice de densidad terminológica para los términos de tipo a)	Numéricos		
	Valor mínimo	Valor promedio	Valor máximo
	0	0,100	0,452
Índice de densidad terminológica para los términos de tipo b)	Numéricos		
	Valor mínimo	Valor promedio	Valor máximo
	0	0,062	0,301
Índice de densidad terminológica para los términos de tipo c)	Numéricos		
	Valor mínimo	Valor promedio	Valor máximo
	0	0,055	0,422
Número de procedimientos de reformulación formal a) por expansión	Numéricos		
	Valor mínimo	Valor promedio	Valor máximo
	0	5	41
Número de procedimientos de reformulación formal b) por reducción	Numéricos		
	Valor mínimo	Valor promedio	Valor máximo
	0	1	22
Número de procedimientos de reformulación formal c) por variación	Numéricos		
	Valor mínimo	Valor promedio	Valor máximo
	0	2	29
Número total de procedimientos de reformulación formal	Numéricos		
	Valor mínimo	Valor promedio	Valor máximo
	0	8	64
Índice de reformulación	Numéricos		
	Valor mínimo	Valor promedio	Valor máximo
	0	43,668	796,533
Índice de reformulación a) por expansión	Numéricos		
	Valor mínimo	Valor promedio	Valor máximo
	0	25,235	397,409
Índice de reformulación b) por reducción	Numéricos		
	Valor mínimo	Valor promedio	Valor máximo
	0	0,016	1
Índice de reformulación c) por variación	Numéricos		
	Valor mínimo	Valor promedio	Valor máximo
	0	9,120	113,545
	0	0,081	1

Tabla 59: Resumen de información proporcionada por el investigador¹⁴⁸

¹⁴⁸ Los valores para IR aparecen tanto crudos (valor superior en la celda) como escalados (valor inferior en la celda).

7. APLICACIÓN DE LA INTELIGENCIA ARTIFICIAL

La aplicación de las ciencias de la computación a la lingüística supone no solo un aumento de la eficiencia en el trabajo lexicométrico o terminológico, sino una contribución a la propia lingüística teórica cuando la interpretación de resultados permite crear un conjunto de criterios teóricos. En el caso concreto de la terminología, la ingeniería lingüística se ha encargado de crear bases de datos, ontologías, de diseñar *software* para detectar y extraer unidades candidatas a término, etc. Algunos ejemplos de estos trabajos son el *Sistema d'Extracció Automàtica de Candidats a Unitats de Significació Especialitzada* (SEACUSE) de Estopà (1999), *Yet Another Term Extractor* (YATE) de Vivaldi (2001), WikiYATE (Vivaldi y Rodríguez, 2012) o la aplicación de algoritmos estadísticos supervisados en la investigación de Nazar y Cabré (2011) o de Nazar (2014). De hecho, este último trabajo, titulado *Desarrollo de un algoritmo estadístico de extracción terminológica basado en coocurrencia léxica* tiene por objetivo diseñar un «estudio teórico para explorar las bases de una teoría cuantitativa de la terminología, que tendría por objeto definir de manera formal la diferencia entre lo que es y lo que no es un término» (Rodríguez-Tapia y Camacho-Cañamón, 2018: 14).

Relacionado con el grado de especialización textual, en 2014 Cabré *et al.* publican un trabajo con el título «Automatic specialized vs. non-specialized text differentiation» que contribuye a la discriminación automática de los dos polos opuestos del grado de especialización mediante la distinción de características gramaticales sobre un corpus etiquetado. El grado intermedio queda excluido en este trabajo. En cuanto a clasificación de textos, el trabajo de Guantiva, Cabré y Castellà (2008) utiliza análisis factorial discriminante para clasificar patrones usando variables cuantitativas.

En esta tesis doctoral, la contribución de las ciencias de la computación a la lingüística tiene un carácter teórico, ya que tiene por objetivo permitir distinguir las características más relevantes o predictivas en el grado de especialización textual desde un punto de vista descriptivo. La Ilustración 15 esquematiza la relación entre ingeniería informática (especialmente aprendizaje automático; *vid infra*) y lingüística.

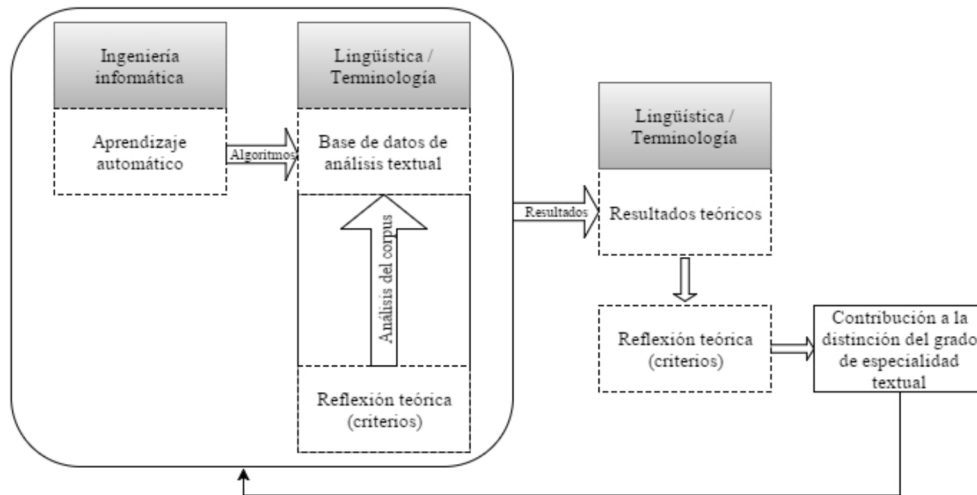


Ilustración 15: Contribución de las ciencias de la computación a la reflexión teórica de la lingüística (Rodríguez-Tapia y Camacho-Cañamón, 2018)

7.1. El aprendizaje automático

Como se mencionaba anteriormente (y en la Ilustración 15), en esta tesis hacemos uso del aprendizaje automático (*machine learning* en inglés). Se trata de una rama de la inteligencia artificial cuyo principal propósito es desarrollar técnicas que permitan que los ordenadores *aprendan*, entendiéndose por *aprender* la «capacidad de generalizar comportamientos a partir de información no estructurada suministrada como ejemplo» (Rodríguez-Tapia y Camacho-Cañamón, 2018: 16). En otras palabras, mediante diferentes algoritmos, con el aprendizaje automático los ordenadores son capaces de reconocer patrones similares dentro de un conjunto determinado (Bishop, 2010), así como clasificar en grupos, determinar rasgos relevantes o determinar probabilidades, las cuales son empleadas en multitud de campos científicos, como el diagnóstico de enfermedades, los movimientos de la bolsa financiera o el procesamiento del lenguaje natural, entre muchos otros (Witten y Frank, 2005). En el caso de las aplicaciones en el campo lingüístico, el aprendizaje automático ha sido empleado en la categorización automática de textos (Sebastiani, 2002) o en la medición de la complejidad de las lenguas (Becerra-Bonache y Jiménez-López, 2015), entre otros.

Los tipos de algoritmos y enfoques que usa el aprendizaje automático son variados y tienen diferentes usos. Con el objetivo de describir las formas de representación del conocimiento, en este caso, nos basamos en investigaciones anteriores donde relacionamos dicha información (Rodríguez-Tapia y Camacho-Cañamón, 2018: 16) basándonos en el trabajo de Witten y Frank (2005: 61-82):

a) Las **tablas de decisión** usan patrones por filas (textos) y atributos (variables) por columnas (*vid. infra* Tabla 60). Se trata de una forma rudimentaria de representar el conocimiento y clasificar en clases.

	Atributo1 <i>IDT</i>	Atributo2 <i>Emisor</i>	Atributo3 <i>Tipo texto</i>	Atributo n <i>Variable n</i>	Decisión <i>Clase</i>
Patrón1 <i>Texto1</i>					
Patrón2 <i>Texto2</i>					
Patrón3 <i>Texto3</i>					
Patrón n <i>Texto n</i>					

Tabla 60: Ejemplo de tabla de decisión [elaboración propia]

b) Los **árboles de decisión** permiten conocer la clase en la que se clasifica un patrón recorriendo las diferentes ramas del árbol (valores de los atributos) hasta llegar hasta la hoja final (la clase asignada al patrón). En la intersección de las diferentes ramas (es decir, en cada nodo) se evalúan uno o varios atributos para tomar dicha decisión. La importancia de los árboles de decisión reside en su capacidad de predicción, puesto que generan reglas a partir de condiciones (la conjunción de los valores de los atributos) para determinar una clase.

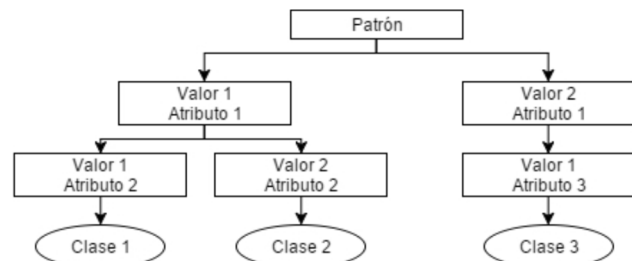


Ilustración 16: Ejemplo de árbol de decisión (Rodríguez-Tapia y Camacho-Cañamón, 2018: 17)

c) Los **clusters** o agrupaciones concentran de forma gráfica aquellos patrones que cuentan con características comunes, de forma que la conjunción de patrones similares permite identificar las características prototípicas (el centroide) del grupo.

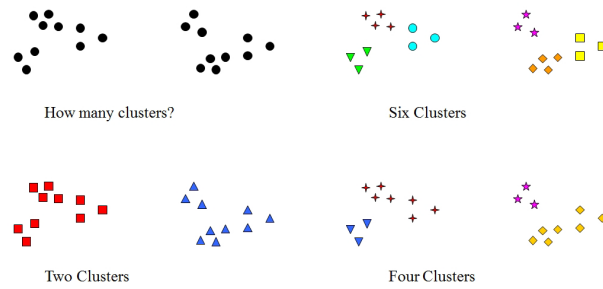


Ilustración 17: Ejemplo de agrupamiento en diferente número de clusters¹⁴⁹

El software de aprendizaje automático empleado en esta tesis doctoral ha sido la versión 3.8.0 de WEKA (*Waikato Environment for Knowledge Analysis*). Constituye una de las aplicaciones de escritorio más extendidas y empleadas actualmente para el tratamiento de datos y extracción de conocimiento, pues cuenta con una colección avanzada de algoritmos de aprendizaje automático y herramientas de preprocesado.

En las secciones siguientes se justifican los cuatro algoritmos empleados para interpretar los datos recopilados. Todos tienen en común que, sobre el conjunto de algoritmos alternativos que pueden emplearse con el mismo fin, son los que mayor porcentaje de éxito tienen en la clasificación o agrupamiento. En trabajos previos nos hemos aproximado a la explotación de bases de datos sobre el grado de especialización textual mediante algoritmos supervisados y no supervisados y hemos comprobado algunas de las vías más eficientes en la interpretación de los datos (Rodríguez-Tapia y Camacho-Cañamón, 2016, 2017, 2018, en prensa).

Con el objetivo de validar los modelos generados, todos los algoritmos empleados utilizan validación cruzada (*cross-validation* en inglés). La validación cruzada es una técnica que permite evaluar los resultados de un análisis estadístico y asegurar que dichos resultados son independientes de la partición entre datos de entrenamiento y datos de prueba. La validación cruzada consiste en dividir los datos de muestra en dos conjuntos complementarios: un subconjunto constituye los datos de entrenamiento y es objeto del primer análisis; el segundo subconjunto constituye los datos de prueba y es sobre el cual se valida el análisis del primer conjunto. Sobre cada una de estas particiones se repite y calcula la media aritmética que se obtiene de las medias de evaluación de dichas particiones, como muestra la Ilustración 18.

¹⁴⁹ Fuente: PANDRE, A. (s. f): «Cluster Analysis: see it 1st», *Data Visualization*. Disponible en: <https://apandre.wordpress.com/visibledata/cluster-analysis/>.

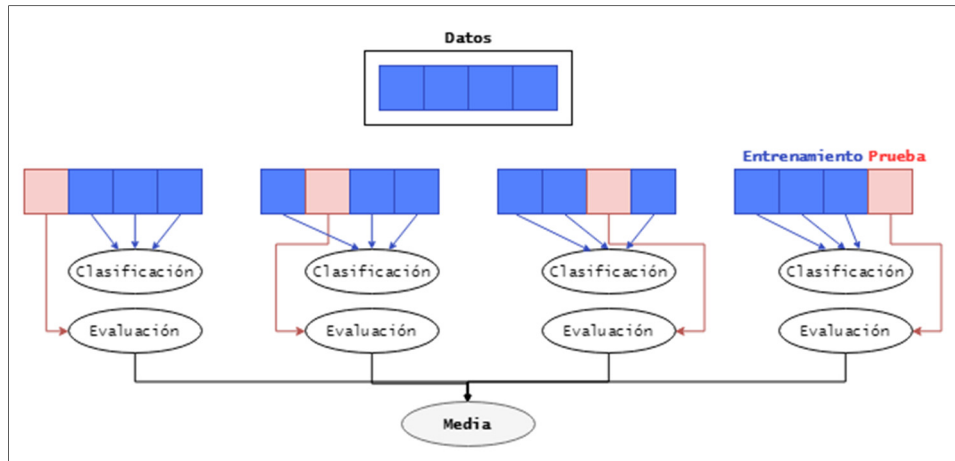


Ilustración 18: Esquema de validación cruzada [elaboración propia]

En nuestra tesis se emplea una validación cruzada de k iteraciones (*k-fold cross-validation*). Este tipo de validación consiste en dividir los datos que sirven de muestra en k subconjuntos: uno de estos subconjuntos es empleado como datos de prueba y el resto de los subconjuntos sirve como datos de entrenamiento. El proceso de validación cruzada se repite durante las k iteraciones con todos los posibles subconjuntos de datos de prueba, como ejemplifica la Ilustración 19.

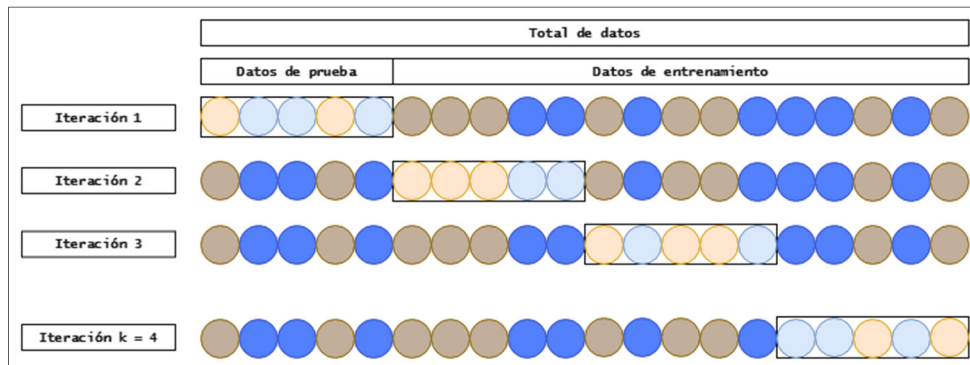


Ilustración 19: Esquema de validación cruzada de k iteraciones [elaboración propia]

7.2. Red bayesiana

Una red bayesiana, red de Bayes o red de creencia, entre otros (*Bayesian network*, *Bayes net* o *belief network* en inglés), es un modelo grafo probabilístico que permite conocer la probabilidad de que un patrón (es decir, un texto) sea clasificado dentro de una clase teniendo en cuenta los diferentes valores de los diferentes atributos (Ben-Gal, 2007: 1). La red bayesiana proporciona una tabla de decisión que indica el valor numérico que posee el valor de un atributo dada una clase.

Una red bayesiana se basa en el teorema de Bayes (1763), que expresa la probabilidad condicional de un evento aleatorio A dado B. En otras palabras, dicho teorema calcula la probabilidad de un suceso A sabiendo que dicho suceso A cumple cierta característica que condiciona su probabilidad¹⁵⁰. La fórmula del teorema de Bayes relaciona la probabilidad condicionada en el numerador y la probabilidad total en el denominador, como se indica en la Tabla 61:

$$P[A_n|B] = \frac{P[B|A_n] \cdot P[A_n]}{\sum P[B|A_i] \cdot P[A_i]}$$

A_n : sucesos condicionados

B: suceso del que se tiene información previa

Tabla 61: Fórmula del teorema de Bayes

Al aplicar el algoritmo de red bayesiana, WEKA facilita el valor para $P[B|A_n]$, por lo que puede calcularse la probabilidad de clasificación en una clase usando dicho valor. Si reformulamos dicho teorema y lo aplicamos a nuestro trabajo, obtenemos la siguiente fórmula:

$$P(C|At) = \frac{\#At \cap \#C}{\#At}$$

C: valor de la clase

At: valor del atributo

#: número de patrones que cumplen con el valor

Tabla 62: Fórmula del teorema de Bayes aplicado al corpus

Así, una de las aplicaciones de este teorema es, por ejemplo, establecer la probabilidad de que un texto se clasifique como texto especializado cuando el emisor es especialista: esto viene determinado por la división entre el número total de patrones del corpus que cumplen conjuntamente con la condición de contar con un emisor especialista y ser clasificados como textos especializados entre el número total de patrones que cuentan con un emisor especialista (*vid.* Tabla 63):

$$P(TE|Emisor especialista) = \frac{\#Emisor especialista \cap \#TE}{\#Emisor especialista}$$

Tabla 63: Ejemplo de la fórmula del teorema de Bayes para determinar la probabilidad de clasificación de un texto como especializado cuando el emisor es especialista

¹⁵⁰ El teorema de Bayes interpreta la probabilidad de manera inversa al teorema de la probabilidad total ($P(B) = \sum_{i=1}^n P(A_i) \cdot P(B|A_i)$), ya que el segundo infiere sobre un suceso B a partir de los resultados de los sucesos A y el teorema de Bayes calcula la probabilidad de un suceso A condicionado a un suceso B.

La configuración del algoritmo empleado en WEKA corresponde con la que muestra la Ilustración 20. Dicho algoritmo ha seguido una validación cruzada de 10 iteraciones y ha obtenido un 80,95 % de instancias correctamente clasificadas (1020 textos).

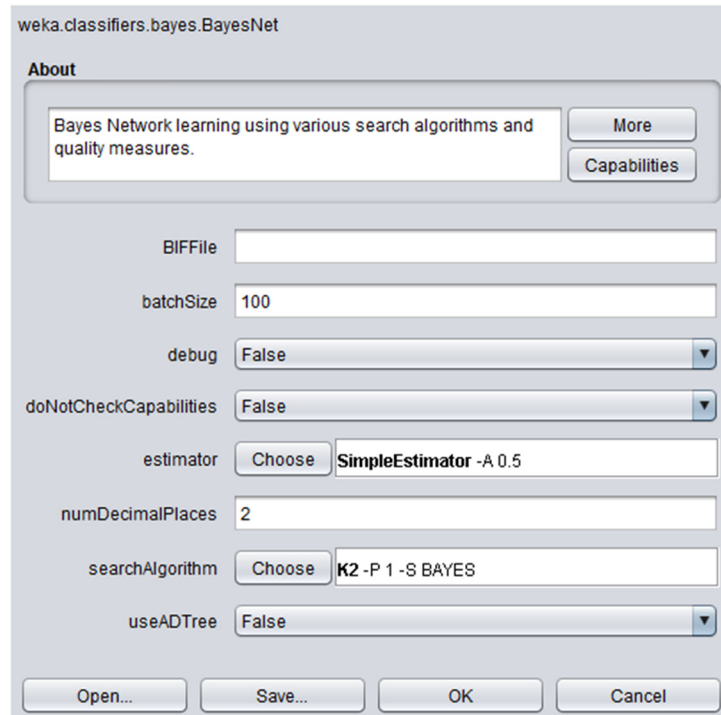


Ilustración 20: Configuración de red bayesiana en WEKA

7.3. Regresión logística

La regresión logística (también conocida como modelo logit o clasificador de máxima entropía) es un tipo de análisis de regresión que permite, a partir de variables independientes (o predictoras), predecir el resultado de una variable categórica (es decir, una variable con un número limitado de valores, especialmente, los valores dicotómicos) (Cornfield, Gordon y Smith, 1961; Walker y Duncan, 1967; Hosmer y Lemeshow, 2000).

La regresión logística permite conocer los valores de los atributos que son más relevantes en la clasificación de un patrón dentro de una clase u otra (en nuestro caso, cuál es el atributo más relevante para el texto especializado y cuál es el más relevante para el texto no especializado, puesto que el modelo trabaja con valores dicotómicos): la regresión logística devuelve una tabla de datos en los que se relacionan todos los valores de los atributos por filas y las clases por columnas. En cada intersección se obtiene un valor que determina la relevancia del atributo en la clasificación. Así, cuanto mayor es dicho valor numérico, más importante es dicho valor del atributo en la correspondiente columna (es

decir, en la clase). Es importante subrayar que el valor que devuelve el modelo de regresión logística para cada una de las variables de la base de datos no establece una relación directa entre el valor de dicha variable y la clase, sino que determina que dicho valor es muy relevante (o, por el contrario, muy poco relevante) para la clasificación. Y esta relevancia también puede interpretarse de forma negativa; es decir, un atributo puede ser muy relevante para *no determinar* un grado de especialización concreto. Con la regresión logística podemos conocer si la terminología es importante en la discriminación del texto especializado (en concreto, qué unidades léxicas son las más importantes), si el tipo textual permite discriminar las clases o si el tema es condicionante de algún grado de especialización.

El algoritmo que empleamos ha sido modificado con un método de regularización de arista (regularización de Tikhonov o *ridge regression* en inglés) de 1×10^{-5} , con el que se consigue que los valores de la ratio de cocientes (u *odds ratio*, como es ampliamente conocido por todos los campos de la comunidad científica) se aproximan a la unidad. Así, los valores que resultan en la tabla final no son tan dispersos y lejanos a cero, sino todo lo contrario.

La configuración del algoritmo empleado en WEKA corresponde con la que muestra la Ilustración 21. Dicho algoritmo ha seguido una validación cruzada de 10 iteraciones y ha obtenido un 80,24 % de instancias correctamente clasificadas (1011 textos).

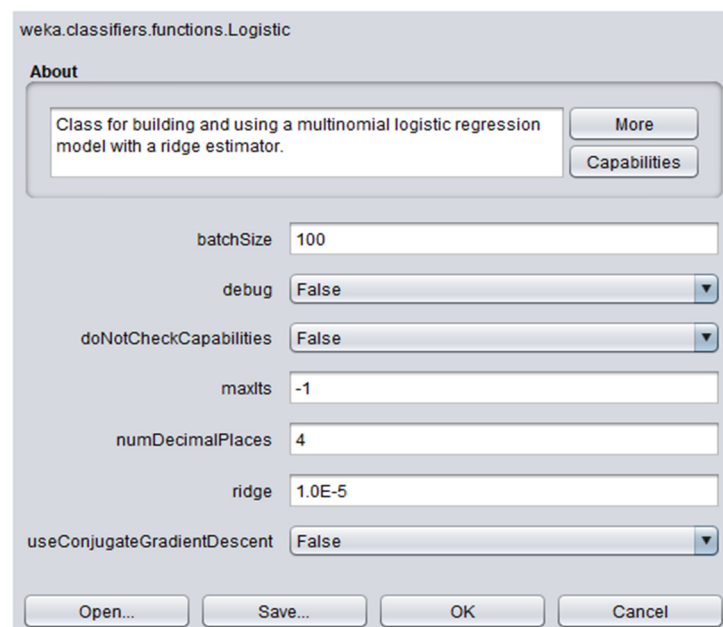
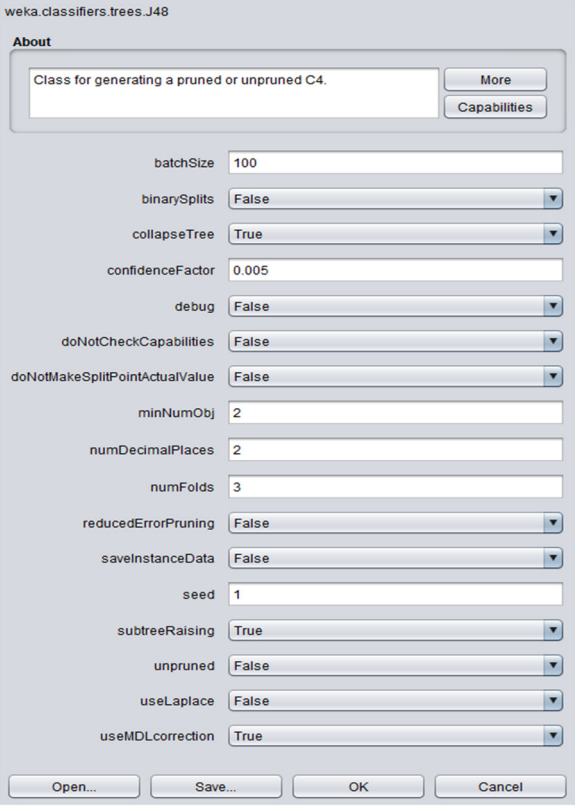


Ilustración 21: Configuración de regresión logística en WEKA

7.4. Árbol de decisión J48

El algoritmo J48 es una implementación *open source* en Java del algoritmo C.45 de WEKA. El algoritmo C.45 es un clasificador estadístico diseñado por Quinlan (1993) que permite diseñar árboles de decisión podados y no podados (*pruned* o *unpruned*). Se habla de un árbol de decisión podado cuando el algoritmo ha eliminado parte de las ramas del árbol hasta llegar a la hoja de la clase, generalizando los datos que existen en dichas ramas. Cuando un árbol de decisión no se poda, los datos desde el tronco del árbol (el atributo de partida) y las hojas finales (la clase), no generalizan ni eliminan ningún dato.

En esta investigación, se ha reducido el factor de confianza del algoritmo (*confidenceFactor*; *vid.* Ilustración 22) para obtener el mejor resultado entre número de hojas del árbol, tamaño del árbol y patrones correctamente clasificados. Así, con la configuración mostrada en la Ilustración 22 y una validación cruzada de 10 iteraciones obtenemos un árbol de decisión con 22 hojas, 32 nodos y un 77,46 % de instancias correctamente clasificadas (976 textos).



The image shows a screenshot of the WEKA J48 classifier configuration dialog box. The window title is "weka.classifiers.trees.J48". It has an "About" section at the top with a text box containing "Class for generating a pruned or unpruned C4." and buttons for "More" and "Capabilities". Below this are various configuration options:

- batchSize: 100
- binarySplits: False
- collapseTree: True
- confidenceFactor: 0.005
- debug: False
- doNotCheckCapabilities: False
- doNotMakeSplitPointActualValue: False
- minNumObj: 2
- numDecimalPlaces: 2
- numFolds: 3
- reducedErrorPruning: False
- saveInstanceData: False
- seed: 1
- subtreeRaising: True
- unpruned: False
- useLaplace: False
- useMDLcorrection: True

At the bottom, there are buttons for "Open...", "Save...", "OK", and "Cancel".

Ilustración 22: Configuración de árbol de decisión J48 en WEKA

7.5. Agrupamiento mediante clasificación no supervisada: *simple k-means*

K-means es un método de agrupamiento de minería de datos no supervisado (MacQueen, 1967). Este algoritmo divide los datos de muestra que pertenecen a un conjunto que cuenta con n patrones en k agrupaciones y clasifica cada patrón en el grupo cuyo valor medio (o centroide) es más cercano a dicho grupo. El centroide es el patrón que, debido a sus características, mejor representa a un grupo.

En la Ilustración 23 se esquematiza el funcionamiento de k-means de forma gráfica: a) en primer lugar se parte de diferentes centroides generados aleatoriamente (marcados en verde, rojo y azul en el primer cuadro); b) en segundo lugar, se generan k grupos con los patrones más cercanos al centroide; c) en tercer lugar se recalculan y actualizan los centroides para cada agrupación; d) los pasos de asignación y actualización se repiten las veces necesarias hasta que se alcance la convergencia en cada grupo. Finalmente, cada agrupación cuenta con características comunes o más o menos similares y un centroide que representa de forma óptima dicha agrupación.

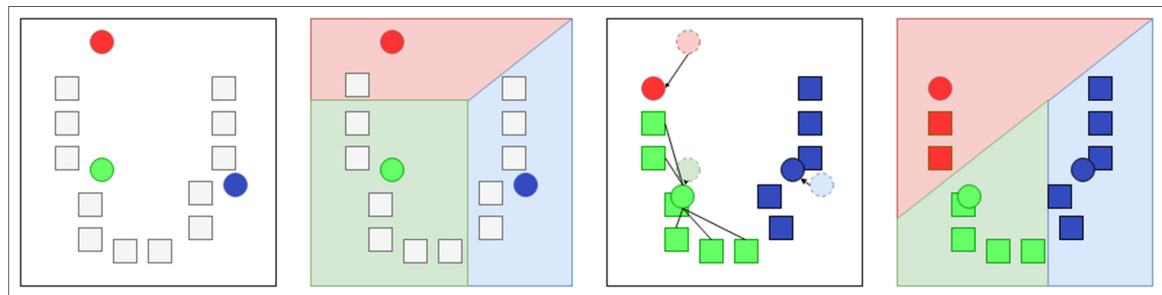


Ilustración 23: Ejemplo de funcionamiento de k-means [elaboración propia]

K-means es un algoritmo no supervisado porque no emplea la etiqueta de clase para diseñar el modelo, sino que utiliza solamente los atributos facilitados. Así, el algoritmo no conoce la clase que se ha asignado a cada patrón. A veces, aun contando con la etiqueta de clase, esta se ignora en el proceso de modelado y solo se emplea en la evaluación del modelo para comprobar que existe similitud entre los resultados del experto o facilitador de la base de datos y el algoritmo.

Los algoritmos no supervisados cuentan con la desventaja de no poder emplear el conocimiento que se podría extraer de la etiqueta de clase para el diseño del modelo. No obstante, en contraste, se trata de algoritmos más robustos, ya que pueden realizar las agrupaciones sin tener que conocer dicha etiqueta (Marsland, 2015, 195-196).

Con la configuración mostrada en la Ilustración 24 y una validación cruzada de 10 iteraciones se han obtenido tres grupos o *clusters* (en el apartado IV. 3.4., p. 438, se observa

cómo los grupos identificados coinciden con las tres clases del grado de especialización textual).

The image shows the 'SimpleKMeans' configuration window in WEKA. The window title is 'weka.clusterers.SimpleKMeans'. It has an 'About' section at the top with a text box containing 'Cluster data using the k means algorithm.' and buttons for 'More' and 'Capabilities'. Below this are various configuration options:

- canopyMaxNumCanopiesToHoldInMemory: 100
- canopyMinimumCanopyDensity: 2.0
- canopyPeriodicPruningRate: 10000
- canopyT1: -1.25
- canopyT2: -1.0
- debug: False
- displayStdDevs: False
- distanceFunction: Choose (EuclideanDistance -R first-l)
- doNotCheckCapabilities: False
- dontReplaceMissingValues: False
- fastDistanceCalc: False
- initializationMethod: Canopy
- maxIterations: 500
- numClusters: 3
- numExecutionSlots: 1
- preserveInstancesOrder: False
- reduceNumberOfDistanceCalcsViaCanopies: False
- seed: 3

At the bottom, there are buttons for 'Open...', 'Save...', 'OK', and 'Cancel'.

Ilustración 24: Configuración de simple k-means en WEKA

Este algoritmo permite conocer cuáles son las características prototípicas o más representativas de una agrupación, en nuestro caso, de cada una de las tres clases según el grado de especialización. La Tabla 64 relaciona los centroides aleatorios de los que parte k-means antes de iniciar los pasos de asignación y actualización.

Atributo	Cluster 0	Cluster 1	Cluster 2
Tema	Derecho	Medicina	Derecho
Tipo texto	Artículo en revista científica	Entrevista	Noticia
TRR	0,65	0,65	0,6384
IDT global	0,257	0,1776	0,1554
IDT a)	0,1006	0,0821	0,1039
IDT b)	0,0852	0,0573	0,0393
IDT c)	0,0714	0,0382	0,0122
IR global	0,0899	0,04827	0,0195
IR a)	0,1046	0,0548	0,0229
IR b)	0,0191	0,0174	0,0055
IR c)	0,1627	0,0545	0,0278
Emisor	Especialista	Instruido	Instruido
Subemisor	Especialista investigador	Instruido semiespecialista	Instruido semiespecialista
Destinatario	Especialista	Instruido	Lego
Subdestinatario	Especialista investigador	Instruido lego formado	Lego no formado
Clase	Texto especializado	Texto semiespecializado	Texto no especializado

Tabla 64: Centroides aleatorios de partida para k-means en WEKA

IV. ANÁLISIS Y RESULTADOS

Este capítulo pretende describir las relaciones entre las diferentes variables que han sido analizadas en el corpus, así como explicar y justificar todas aquellas particularidades que no quedan suficientemente evidenciadas por la mera relación de las variables. Para ello, llevaremos a cabo tres análisis diferentes:

- a) uno dedicado a la totalidad de los datos, es decir, a analizar en conjunto los 1260 textos analizados por los cuatro grupos de informantes;
- b) un segundo análisis dedicado a profundizar en las similitudes y divergencias perceptivas de los cuatro grupos de informantes con respecto a su análisis individual;
- c) un análisis mediante el empleo de aprendizaje automático, que permita establecer las relaciones de probabilidad de clasificación y las convergencias principales entre variables.

Con el objetivo de facilitar la consulta, se han diseñado tres anexos que exponen las relaciones principales entre las variables:

- a) el anexo 3A incluye las tablas y gráficos correspondientes al cruce de datos entre el tema, la clase, el tipo de texto, el índice de densidad terminológica, el índice de reformulación y la ratio *types/tokens* estandarizada para el conjunto de textos de todo el corpus, es decir, representa los datos globales que corresponden con los cuatro grupos de informantes.
- b) el anexo 3B incluye las tablas y gráficos correspondientes a la clase, el emisor, el subemisor, el destinatario, el subdestinatario, las formas léxicas especializadas y los procedimientos de reformulación formal, así como la relación del índice de densidad terminológica y el índice de reformulación con todas las variables antes mencionadas. La particularidad del anexo 3B se halla en que los datos se presentan atendiendo al informante, y no de forma global, por lo que es posible discriminar qué datos ha proporcionado cada grupo de informantes sobre una variable, conjunto de variables o relación de variables. El objetivo del anexo 3B es contrastar los datos de los cuatro grupos de informantes.

Los datos presentes en tablas y gráficos se relacionan de forma ordinal.

1. ANÁLISIS CORRELACIONAL DESCRIPTIVO GLOBAL

La información que se presenta en este apartado es resultado de la comparación estadística de todas las variables de análisis. Dicha información puede consultarse en mayor profundidad en el anexo 3A.

1.1. Análisis del TEMA

1.1.1. Relación con la CLASE

Los tres temas del corpus están representados de forma equilibrada en las tres clases según la especialización (*vid. infra* Gráfico 4). Es decir, ningún tema se inclina hacia una clase textual específica. Las mayores diferencias se hallan en el texto semiespecializado, en los que el derecho y la informática distan entre sí en más de un 10 % de representación en dicha clase (28,17 % y 38,50 % respectivamente). De hecho, el texto semiespecializado de temática informática es el que más destaca dentro de todo el corpus y el texto semiespecializado de temática jurídica el que menos.

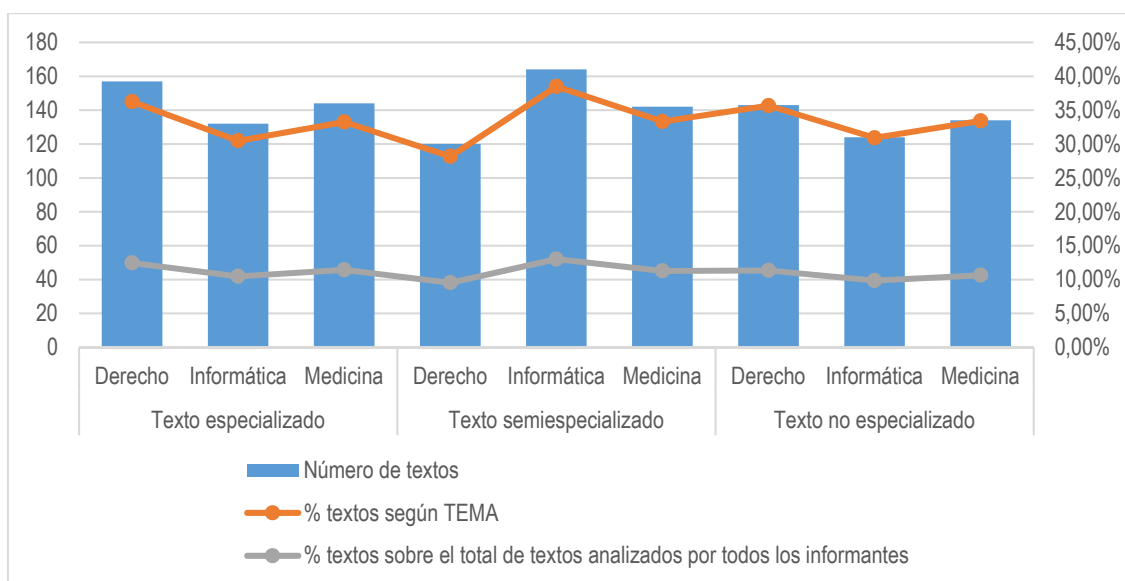


Gráfico 4: Datos globales de clase según tema

1.1.2. Relación con el EMISOR y el SUBEMISOR

Con respecto al emisor de los textos, destaca el hecho de que prácticamente la mitad de los textos (49,76 %) están emitidos por un especialista y la otra mitad (45 %) por un

usuario instruido, correspondiendo el remanente (5,24 %) con la categoría de lego¹⁵¹. Son más numerosos los emisores especialistas en derecho y medicina (36,52 % y 35,41 %), mientras que el peso del emisor instruido lo tiene la informática (41,62 %).

En el caso del emisor lego, donde solo se clasifican 66 textos de un total de 1260, tan solo 8 son de temática informática, siendo el derecho y la medicina los que priman en dicha categoría, con 28 y 30 textos respectivamente.

Los 8 textos de informática con usuario lego son los siguientes:

—Informante estudiante: INF FOR11, INF FOR12, INF FOR13.

—Informante usuario medio: INF FOR12, INF FOR13.

—Informante traductor: INF FOR12, INF FOR13.

—Informante hablante oyente-ideal: INF FOR12.

Todos estos textos son participaciones en foro clasificados como no especializados o semiespecializados. Además, todos los textos tienen un IDT inferior a la media, próximo a 0,1, a excepción de INF FOR11 que posee un IDT por encima de la media (0,275) y una variedad léxica de 58,8 %.

Los 58 textos de derecho y medicina antes mencionados son más variados en la densidad terminológica pero no en la clasificación de la clase textual, puesto que el grueso de ellos se clasifica como texto no especializado, siendo solo 4 de ellos textos semiespecializados. Todos los textos son participaciones en foro, a excepción de un texto publicado en un blog. Que exista menor número de textos emitidos por legos en el campo de la informática hace pensar que los informantes consideran que el emisor del campo de la informática suele mostrar más dominio conceptual (normalmente a través del léxico¹⁵²) y quedan clasificados dentro de la categoría del instruido. En efecto, si consultamos los textos de participación en foro de temática informática, puede distinguirse que 51 de los 60 textos están emitidos por instruidos y, especialmente, por instruidos legos formados (40 textos), que constituye la categoría inmediatamente superior al lego no formado (véase Gráfico 5).

¹⁵¹ Debe tenerse en cuenta que la categoría de lego solo está compuesta por un único subemisor (el lego no formado), mientras que las categorías de especialista y de instruido están formadas por dos y por tres subemisor respectivamente, lo que influye en la categorización.

¹⁵² En el apartado IV. 1.4.1. se muestra que el IDT c) en informática es de 0,067, muy cercano al de la medicina (0,07).

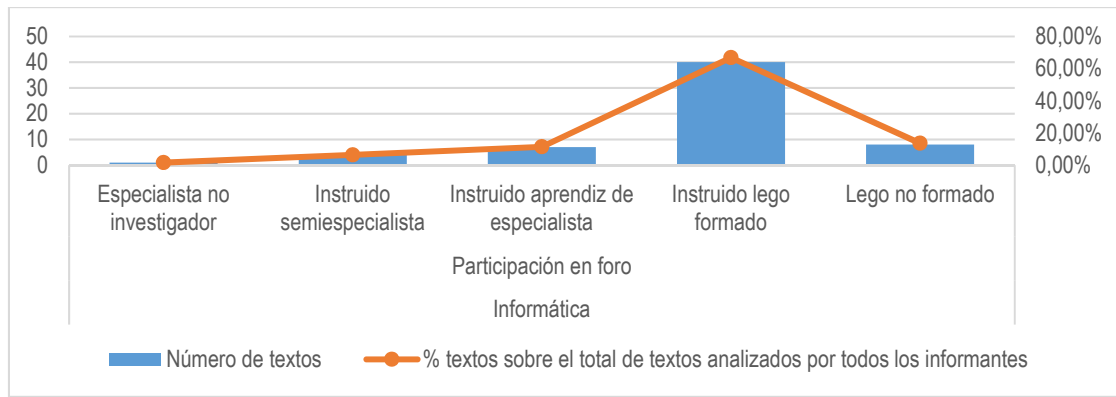


Gráfico 5: Datos de subemisor para participación en foro sobre informática

Por ello, es posible que la causa que disminuya la presencia de textos emitidos por legos en el campo informático se deba a una percepción por parte del informante sobre el control conceptual que tiene el emisor. Como se muestra en la Tabla 65, los textos emitidos por legos no solo cuentan con menor densidad terminológica, sino, además, con menor reformulación, siendo inexistentes los procedimientos de reformulación por reducción o por variación.

	Textos emitidos por legos no formados	Textos emitidos por instruidos legos formados
IDT	0,129	0,161
IR	6,015	15,823
IR por expansión	6,015	6,848
IR por reducción	0	6,375
IR por variación	0	2,600

Tabla 65: Promedio de datos numéricos para textos emitidos por legos e instruidos en participación en foro sobre informática

Profundizando en la categoría de subemisor para todos los temas (*vid. infra* Gráfico 6), las diferencias entre medicina y derecho e informática siguen manteniéndose: el derecho y la medicina muestran mayor porcentaje de textos emitidos por especialista investigador y especialista no investigador, y la informática, mayor número de textos en los tres tipos de instruido.

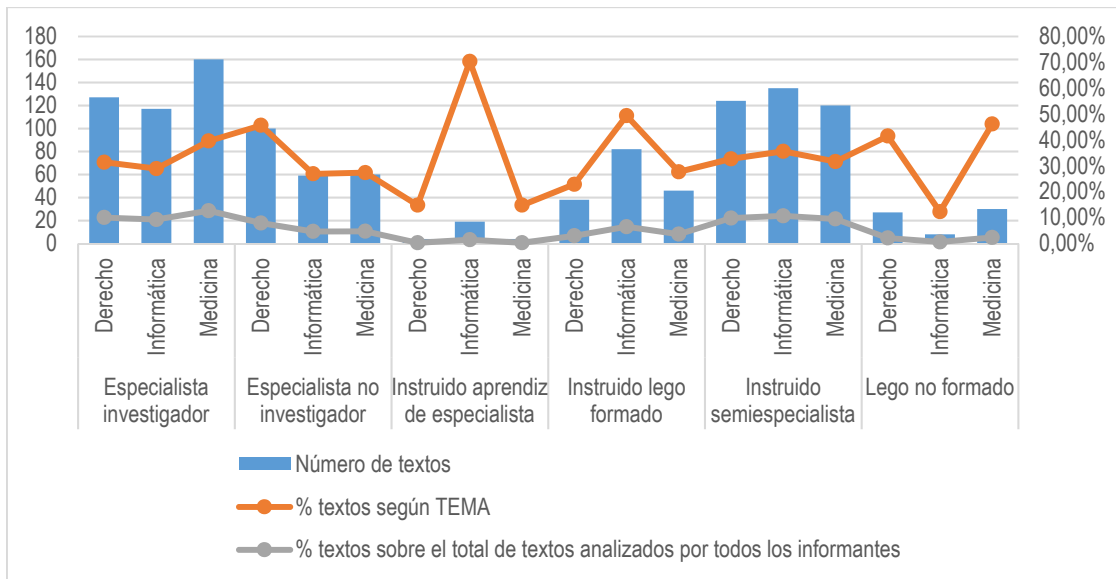


Gráfico 6: Datos globales de subemisor según tema

Centrándonos en la categoría de emisor especialista, parece que los informantes consideran que existe mayor número de subemisores especialistas investigadores en medicina que en cualquier otro campo (39,60 % sobre el total de la subcategoría) y mayor número de especialistas no investigadores en derecho que en cualquier otro campo (45,66 % sobre el total de la subcategoría). Aun así, en relación con el total del corpus, el especialista investigador supone un 32,06 % frente al 17,38 % del especialista no investigador, lo que supone dos tercios de la categoría de emisor especialista.

En cuanto a la categoría de instruido, el instruido semiespecialista es una subcategoría que, dentro de los emisores, muestra diferencias apenas perceptibles con diferentes porcentuales ligeras (apenas un 3 % de diferencia). Las diferencias son más acusadas en el subemisor instruido lego formado, como se mencionó con anterioridad, donde destaca la informática. En cierta forma, la categoría de emisor instruido recoge el número de textos que no han sido clasificados por los informantes como textos emitidos por legos o por especialistas, puesto que constituye la categoría intermedia.

En general, más allá de las diferencias comentadas, los porcentajes de tipo de subemisor por tema son muy similares en todo el corpus.

1.1.3. Relación con el DESTINATARIO y el SUBDESTINATARIO

La descompensación presente en el caso del emisor se mantiene en el destinatario, si bien de forma diferente: la categoría instruido alcanza casi la mitad del corpus (48,57 %), mientras que el destinatario especialista y el destinatario lego muestran valores más equilibrados (27,86 % y 23,57 %; *vid. infra* Gráfico 7).

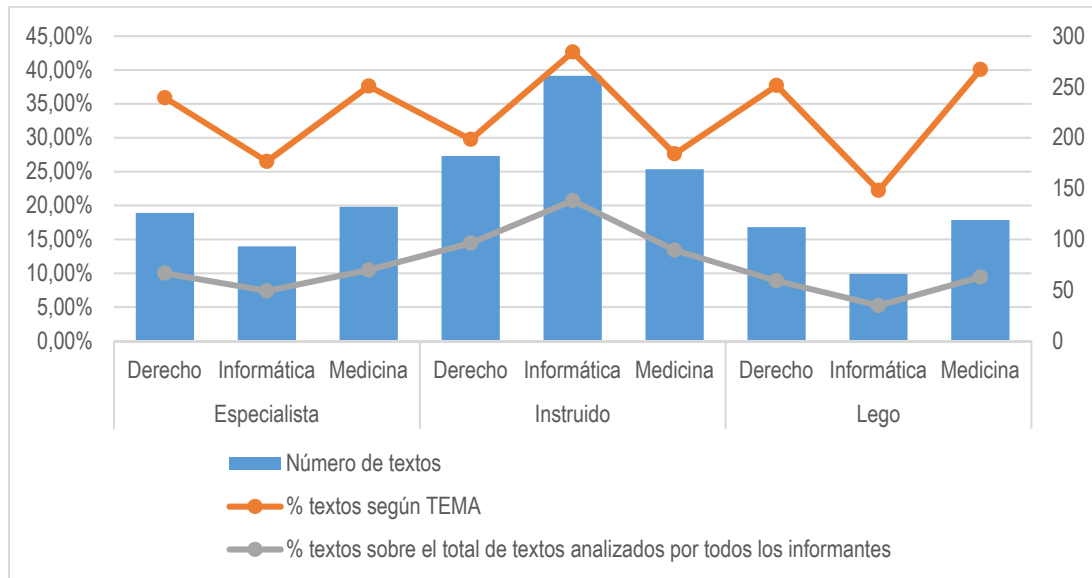


Gráfico 7: Datos globales de destinatario según tema

Un dato que llama la atención es que el patrón observado para el emisor en los textos sobre informática se repite en el destinatario: los textos con destinatario especialista y destinatario lego son los textos menos presentes en el corpus (7,38 % y 5,24 %), mientras que los textos con destinatario instruido son los más numerosos de todo el corpus (20,71 %). En el caso del destinatario instruido, el derecho supera muy ligeramente a la medicina, mientras que esta despunta al derecho, también con diferencias poco notables (menos de un 3 %), en el destinatario lego y en el destinatario especialista (menos de un 1 %).

Para profundizar en la justificación del destinatario instruido en informática, vamos a tener en cuenta la subcategoría de subdestinatario. El principal dato que sobresale al comparar las estadísticas para subemisor y para subdestinatario según el tema es que el patrón observado en el subemisor también se repite en el subdestinatario, como muestra la superposición de líneas del Gráfico 8.

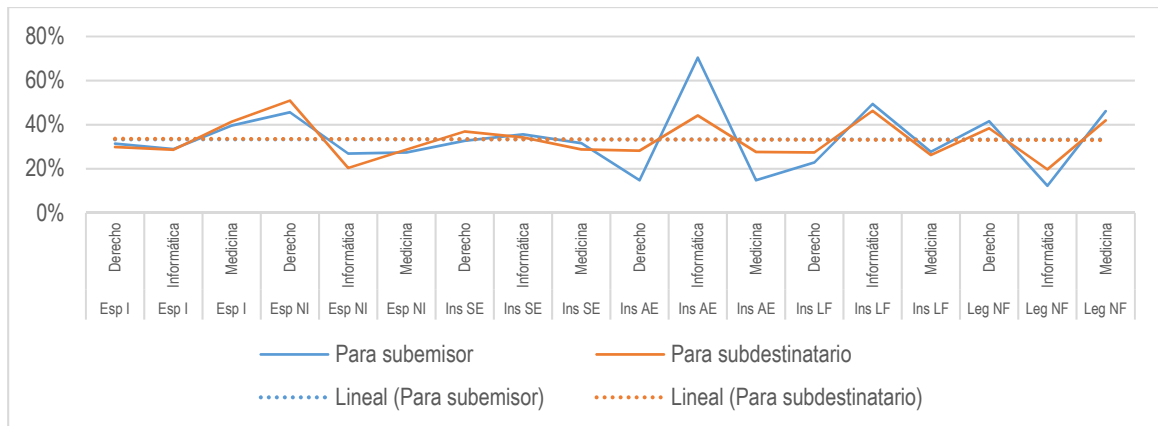


Gráfico 8: Superposición de datos porcentuales para subemisor y subdestinatario según temas

La mayor diferencia se halla en el instruido aprendiz de especialista, que es más acusada en el subemisor que en el subdestinatario. Desde nuestro punto de vista, esta diferencia puede deberse a que los informantes pueden tener dificultades para identificar cuándo un emisor corresponde con un instruido aprendiz de especialista, dados los límites difusos de concreción del conocimiento en el acto discursivo, sobre todo, con las subcategorías de semiespecialista o lego formado.

A pesar de que el patrón subemisor-subdestinatario sea casi idéntico, existe una diferencia que se observa con mayor nitidez al comparar el número de textos, y no los valores porcentuales sobre el tema. Es la que existe entre el lego no formado y el instruido semiespecialista (véase Gráfico 9). Los valores para subemisor y subdestinatario prácticamente se invierten, como muestran la línea de media móvil y la Tabla 66.

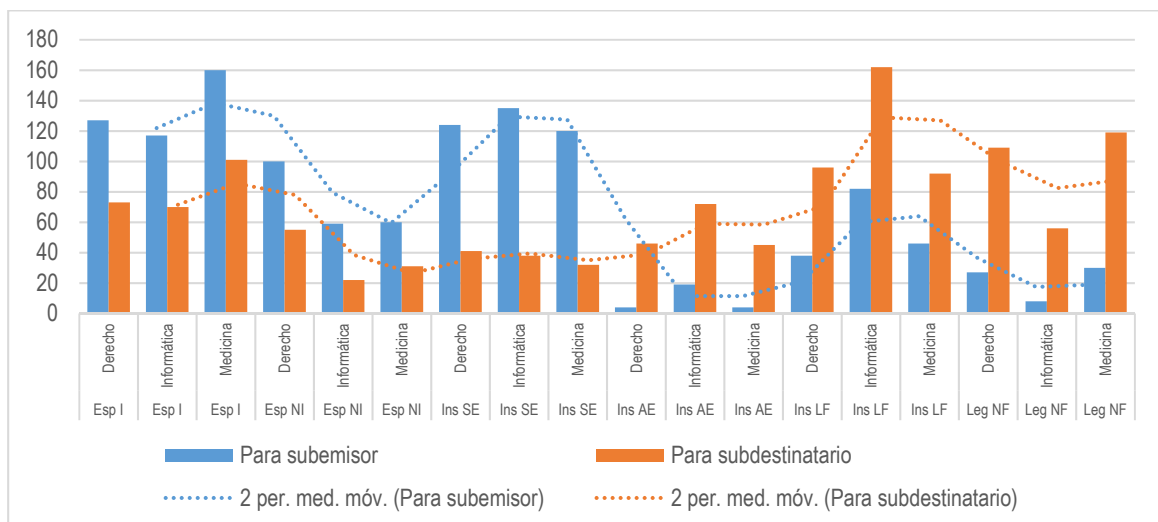


Gráfico 9: Comparación de número de textos por subemisor y subdestinatario según temas

	Instruido semiespecialista	Lego no formado
Subemisor	30,08 %	5,16 %
Subdestinatario	8,81 %	22,54 %

Tabla 66: Comparación de media para todos los temas para instruido semiespecialista y lego no formado según subemisor y subdestinatario

Se puede inferir, a partir de estos resultados, que los grupos de informantes consideran que, para poder emitir un texto, dicho emisor debe mostrar dominio conceptual sobre un tema. No obstante, los destinatarios de los textos solo están sujetos al conocimiento enciclopédico o compartido según los grupos sociales, lo que hace que se identifique un mayor número de legos formados y legos no formados en los textos. Con todo, estos datos deben vincularse obligatoriamente al tipo de corpus y al conjunto de tipos textuales que lo forman.

1.1.4. Relación SUBEMISOR-SUBDESTINATARIO

Si continuamos con las relaciones entre subemisor y subdestinatario, un dato relevante en el corpus es la correspondencia que existe entre los diferentes subemisores y subdestinatario. En este apartado vamos a comparar la correspondencia que existe entre subemisor y subdestinatario en los tres temas del corpus, con el objetivo de identificar las principales similitudes y diferencias.

En los siguientes gráficos el eje de abscisas contiene dos categorías: la inferior representa el tipo de subemisor y la superior representa el tipo de subdestinatario. Así, la primera columna de datos del Gráfico 10 representa la cantidad de textos y el porcentaje correspondiente a los textos emitidos por un especialista investigador y cuyo destinatario también es un especialista investigador.

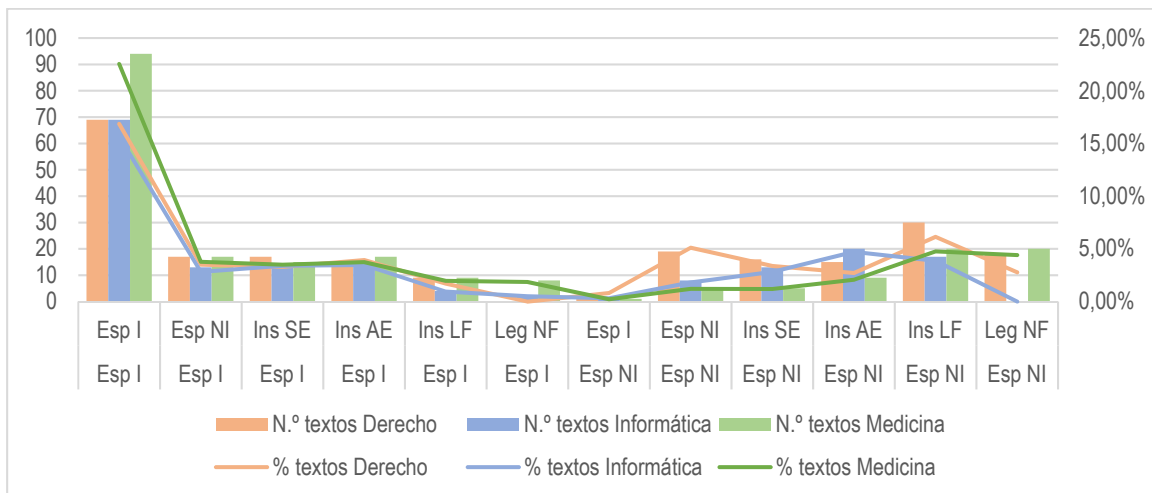


Gráfico 10: Relación de datos globales para la correspondencia emisor especialista-tipo de subdestinatario según tema

En primer lugar, las dos subcategorías para el emisor especialista se combinan con todos los tipos de subdestinatarios planteados en la metodología. El dato esperable y coincidente con los fundamentos teórico-metodológicos de la TCT es que el grueso de textos emitidos por un especialista investigador va dirigido a otro especialista investigador en todos los temas. Solo la medicina destaca sobre el derecho y la informática, con una diferencia de un 7 % (un dato curioso es que tanto en derecho como en informática existen exactamente 69 textos de especialista investigador a especialista investigador). De hecho, la estadística para el emisor especialista se encuentra muy desbalanceada a favor del destinatario especialista investigador, siendo el resto de destinatarios poco relevantes sobre el total del corpus.

En el caso de las relaciones a partir del emisor instruido, debe resaltarse que, al contrario que con el emisor especialista, las subcategorías de emisor instruido no se combinan con todas las subcategorías de destinatario (*vid.* Gráfico 11). Tan solo el instruido semiespecialista (que se trata de la subcategoría más numerosa para el emisor instruido, como ya comprobamos en el apartado IV. 1.1.2., p. 244) se combina con todos los posibles subdestinatarios.

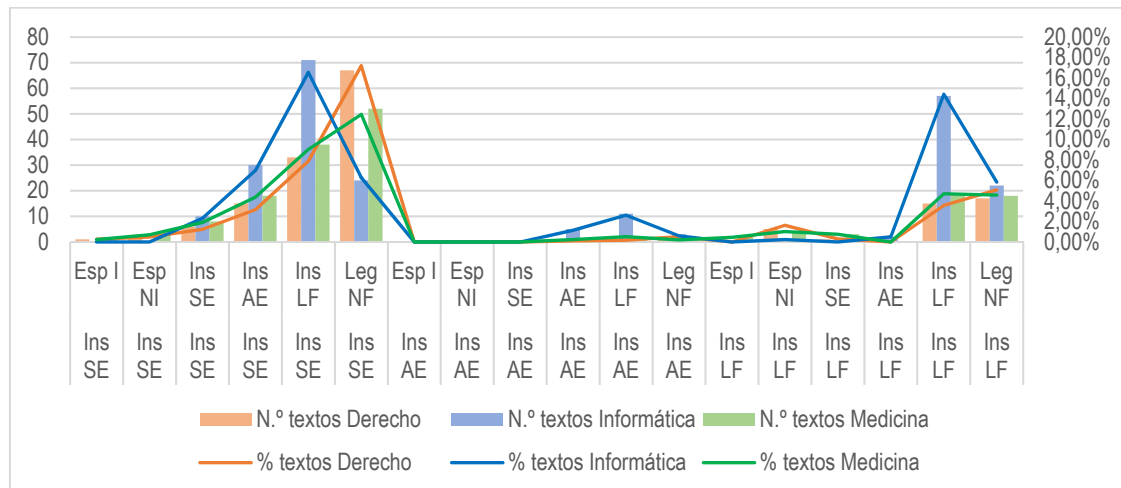


Gráfico 11: Relación de datos globales para la correspondencia emisor instruido-tipo de subdestinatario según tema

Como se observa en el Gráfico 11, la relación instruido semiespecialista-instruido lego formado es la que más destaca para la informática. Si consultamos dichos textos, puede observarse que entre ellos destacan especialmente los textos semiespecializados, aunque también hay no especializados y tres especializados (uno de ellos coincide en el informe del estudiante y del traductor, el INF ENT06). El IDT oscila entre 0,014 y 0,360 y el IR oscila entre 0 y 512,057, lo cual quiere decir que existen textos muy variados en densidad

terminológica y reformulación. Los tipos textuales de estos textos son el blog, la entrevista, la noticia, la participación en foro, la revista o web de divulgación especializada y, contra todo pronóstico, la tesis doctoral. Uno de los informantes del grupo de usuario medio ha clasificado una tesis doctoral sobre informática, con un IDT por debajo de la media (0,158) y un IR por debajo de la media (37,896) como texto semiespecializado, y también es el que ha sido considerado como emitido por un semiespecialista y dirigido a un lego formado. Con todo, se trata de un patrón no relevante en el conjunto del corpus.

La segunda combinación que más destaca es la relación semiespecialista-instruido lego no formado, en la que son más numerosos los textos de derecho y medicina, y no los de informática. Las características de estos textos en derecho y medicina coinciden en gran medida: destacan los textos no especializados, aunque también existen algunos textos clasificados como semiespecializados (más en medicina, en proporción con el derecho), los tipos textuales son el blog, la entrevista, la noticia, la participación en foro y la revista o web de divulgación especializada; y la densidad terminológica y reformulación oscila entre valores muy reducidos o muy altos (sobre todo en medicina, donde la reformulación parece ser mayor). La informática, aunque menos frecuente en esta combinación, también comparte algunas de estas características¹⁵³: destacan los textos no especializados (también hay 4 semiespecializados; el caso de INF NOT12 se repite en tres informantes), los tipos textuales son el blog, la noticia y la revista o web de divulgación especializada; y la densidad terminológica y la reformulación es similar en los otros casos, si bien el IDT máximo es algo más de la mitad que el de los otros dos casos.

La principal diferencia entre el derecho y la medicina y la informática se encuentra precisamente en la delgada línea de separación que existe entre el subdestinatario lego formado, más frecuente en informática, y el subdestinatario lego no formado, más frecuente en derecho y en medicina. Esto permitiría interpretar, de nuevo, que los informantes consideran que, debido a la configuración discursiva de los textos sobre informática, estos necesitan de un destinatario con una competencia cognitivo-comunicativa mayor.

El tercer caso de la informática que destaca muy por encima de los otros dos temas se halla en la relación entre el emisor instruido lego formado y el destinatario instruido lego formado. Entre los textos filtrados no hay textos especializados, los tipos textuales identificados son el blog, la entrevista, la noticia, la participación en foro y la revista o web

¹⁵³ A nuestro juicio, no existen más similitudes entre los tres campos debido a que la presencia de la informática en esta combinación es mucho menor, lo que reduce la cantidad de casos que pueden establecer similitudes.

de divulgación especializada y el IDT e IR son similares al resto de casos. No obstante, los valores máximos son superiores a los presentes en semiespecialista-lego no formado.

En cuanto a la última subcategoría, la de lego no formado (véase Gráfico 12), es posible hallar combinaciones con todas las subcategorías, a excepción de la de aprendiz de especialista. Varios de los datos que más llaman la atención en este gráfico corresponden con las líneas que sobresalen en dicha estadística. No obstante, debe señalarse que las diferencias entre los valores máximos y mínimos no son mayores que un 4 %.

Al identificar los textos que corresponden con dichos elementos más destacables, es decir, a) la relación lego no formado-especialista no investigador en derecho; b) lego no formado-instruido lego formado en derecho; y c) lego no formado-lego no formado en medicina, se puede prestar atención al hecho de que se trata de textos con un alto grado de similitud, puesto que casi todos ellos son participaciones en foro, clasificados como no especializados, con una densidad terminológica y reformulación parecidas.

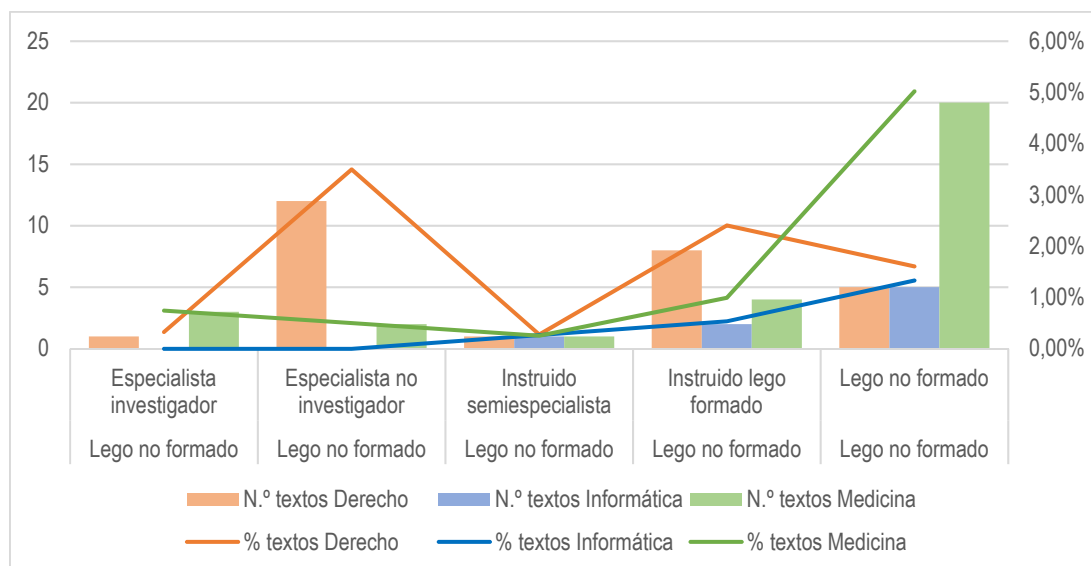


Gráfico 12: Relación de datos globales para la correspondencia emisor lego-tipo de subdestinatario según tema

Con respecto a las excepciones, en la primera relación existen dos textos semiespecializados según los informantes estudiantes y la densidad terminológica es menor que la media; en la tercera relación, por otra parte, la excepción al tipo textual común es un texto procedente de un blog, clasificado, además, como texto semiespecializado.

A nuestro juicio, la diferencia en la distribución de las relaciones, sobre todo entre medicina y derecho, puede deberse a la abstracción con la que las consultas en foro son tratadas en derecho, al tratarse de conceptos no materiales en la realidad, lo que haría que los informantes aumentasen el número de clasificaciones dirigidas a un especialista no investigador, en lugar de a un lego no formado, como en el caso de medicina.

1.1.5. Relación SUBEMISOR-CLASE

En este apartado pretendemos identificar la correspondencia que existe entre el subemisor de cada texto y su clase, y observar si existe alguna diferencia entre los diferentes temas.

En primer lugar, la diferencia entre temas no es muy representativa en ninguna de las combinaciones emisor-clase. De hecho, los tres temas suelen mostrar el mismo tipo de comportamiento en dicha relación.

Entre las similitudes, el emisor especialista investigador emite aproximadamente un 30 % textos especializados sobre el total del corpus, siendo muy poco frecuentes los no especializados y los semiespecializados. Aunque con menor presencia en el corpus, los textos emitidos por especialistas no investigadores se distribuyen de forma más homogénea en las tres clases (véase Gráfico 13).

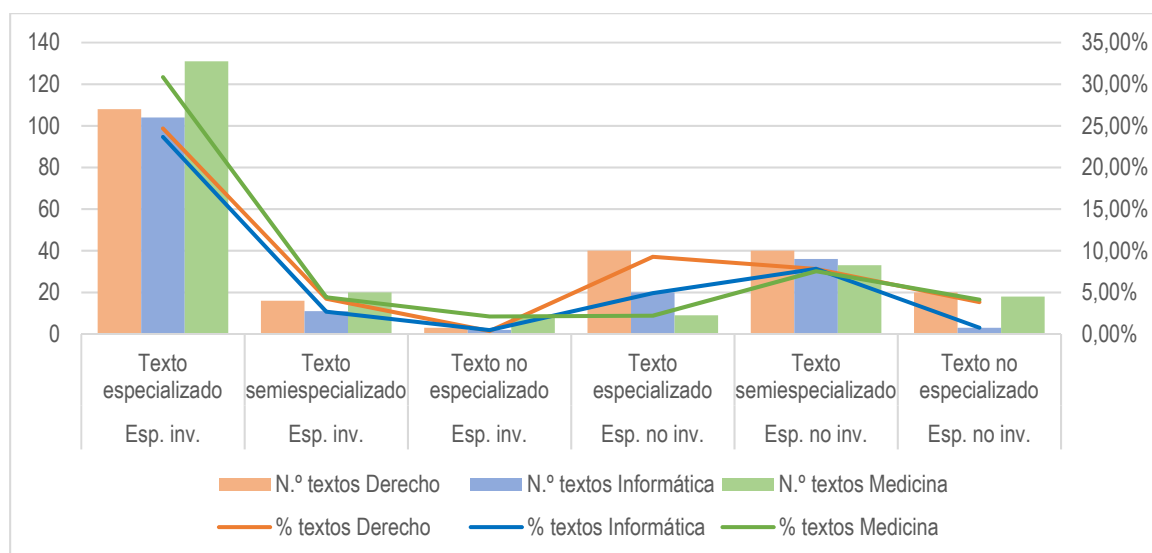


Gráfico 13: Relación de datos globales para la correspondencia emisor especialista-tipo de clase según tema

El patrón de distribución también es similar en la correspondencia emisor instruido-clase, con clara relevancia del texto no especializado y semiespecializado, exceptuando algunas diferencias temáticas que trataremos a continuación (*vid.* Gráfico 14). Por último, la relación lego no formado-texto no especializado parece clara, aunque con un despunte notorio en medicina y derecho en contraposición con la informática (*vid.* Gráfico 15).

En cuanto a las diferencias entre los temas, no suelen superar el 10 %. Empezando por el semiespecialista, la relación más evidente es que la frecuencia de textos según los temas es inversa en el texto no especializado y el semiespecializado (véase Gráfico 14).

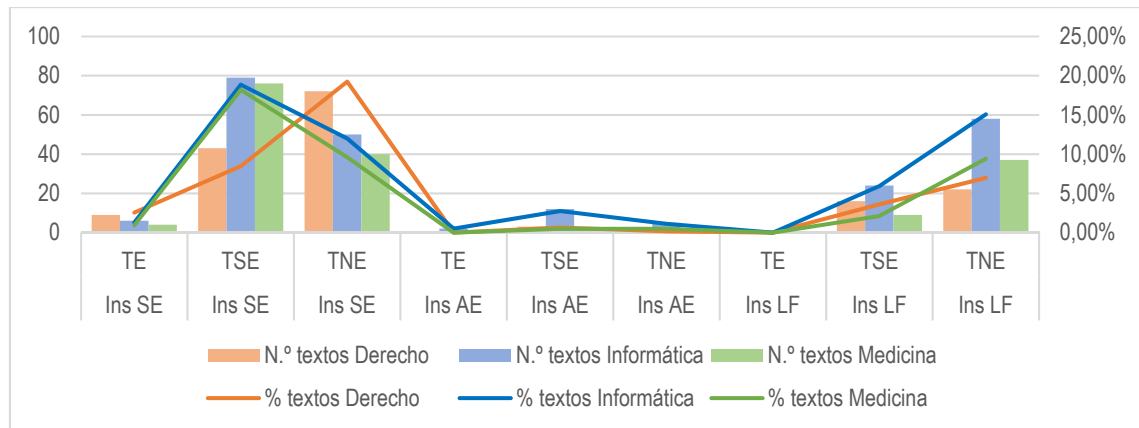


Gráfico 14: Relación de datos globales para la correspondencia emisor instruido-tipo de clase según tema

Si consultamos las características que comparten estos grupos de relación inversa, hallamos que las diferencias principales se encuentran en los valores relativos al índice de reformulación, que corresponden con los valores más reducidos en derecho en el texto no especializado (Tabla 67) y los valores mayores en informática en el texto semiespecializado (Tabla 68).

		Derecho	Informática	Medicina
Destinatario		Principalmente, lego no formado, algunos casos de instruido lego formado y solo uno de especialista no investigador.	Principalmente, lego formado y lego no formado. También algunos casos de las subcategorías de instruido.	Suele ser un lego no formado. También aparece el resto de las subcategorías, a excepción de las de especialista.
Tipo de texto		Blog, entrevista, noticia, participación en foro y web o revista de divulgación especializada.	Blog, entrevista, noticia, y web o revista de divulgación especializada.	Blog, entrevista, noticia, participación en foro y web o revista de divulgación especializada.
Variedad léxica	Mínimo	56,80	59,60	60,87
	Promedio	68,68	69,68	71,66
	Máximo	77,98	82,63	79,83
IDT	Mínimo	0,008	0,028	0,028
	Promedio	0,13	0,15	0,15
	Máximo	0,332	0,386	0,386
IR	Mínimo	0	0	0
	Promedio	8,20	40,48	34,18
	Máximo	40,828	432	281

Tabla 67: Diferencias según temas para textos emitidos por instruido semiespecialista para texto no especializado

		Derecho	Informática	Medicina
Destinatario (por orden de frecuencia)		Lego formado, aprendiz de especialista, lego no formado y semiespecialista.	Lego formado, aprendiz de especialista, semiespecialista y lego no formado.	Lego formado, lego no formado, aprendiz de especialista, semiespecialista y lego no formado.
Tipo de texto		Blog, entrevista, noticia, participación en foro y web o revista	Blog, entrevista, noticia, participación en foro y web o revista	Blog, entrevista, noticia, participación en foro y web o revista

		Derecho	Informática	Medicina
		de divulgación especializada.	de divulgación especializada. Y tesis doctoral ¹⁵⁴ .	de divulgación especializada.
Variedad léxica	Mínimo	61,20	56,32	57,00
	Promedio	69,06	67,53	68,35
	Máximo	78,43	83,28	80,76
IDT	Mínimo	0,047	0,036	0,071
	Promedio	0,22	0,18	0,2
	Máximo	0,404	0,382	0,503
IR	Mínimo	0	0	0
	Promedio	29,38	62,12	36,60
	Máximo	300,926	796,533	140

Tabla 68: Diferencias según temas para textos emitidos por instruido semiespecialista para texto semiespecializado

Siguiendo con la relación de emisor instruido lego formado, las diferencias se hallan principalmente en la informática, que comparte algunas características tanto en la clase de texto semiespecializado como de no especializado. Como se puede comprobar en la Tabla 69 y la Tabla 70, el destinatario se clasifica como lego formado (o también lego no formado en el caso de textos no especializados), el IDT es menor con respecto a los otros temas y los valores máximos de reformulación son mayores. Ahora bien, la media de reformulación se sitúa en las dos clases por debajo de la medicina, aunque siguiendo sus valores de muy cerca.

		Derecho	Informática	Medicina
Destinatario (por orden de frecuencia)		Lego formado, especialista no investigador, lego no formado.	Principalmente lego formado. Solo un especialista no investigador.	Lego formado, lego no formado, especialista investigador.
Tipo de texto		Blog, revista o web de divulgación especializada, participación en foro y noticia.	Principalmente revista o web de divulgación especializada y participación en foro. Solo una entrevista y dos blogs.	Blog, revista o web de divulgación especializada, entrevista, participación en foro y noticia.
Variedad léxica	Mínimo	64,64	52,76	59,30
	Promedio	69,46	61,21	71,76
	Máximo	78,41	80,09	79,83
IDT	Mínimo	0,111	0,092	0,141
	Promedio	0,21	0,18	0,24
	Máximo	0,355	0,375	0,347
IR	Mínimo	0	0	0
	Promedio	7,71	12,39	16,35
	Máximo	17,960	74,933	35,412

Tabla 69: Diferencias según temas para textos emitidos por instruido lego formado para texto semiespecializado

¹⁵⁴ Estas tesis doctorales son dos textos: INF TES05 e INF TES07. Son dos textos repetidos en los informes de 3 informantes. El IDT es menor que la media global del corpus y la reformulación es ligeramente mayor que la media global del corpus.

		Derecho	Informática	Medicina
Destinatario (por orden de frecuencia)		Lego no formado, lego formado, semiespecialista, especialista no investigador.	Lego formado y lego no formado. Dos casos de aprendiz de especialista.	Lego no formado, lego formado, semiespecialista y especialista no investigador.
Tipo de texto		Participación en foro, noticia, revista o web de divulgación especializada y noticia.	Entrevista, noticia, participación en foro y revista o web de divulgación especializada.	Blog, entrevista, noticia, participación en foro.
Variedad léxica	Mínimo	56,80	55,31	63,70
	Promedio	67,32	69,04	74,14
	Máximo	77,51	82	89,54
IDT	Mínimo	0,042	0,013	0,020
	Promedio	0,15	0,14	0,13
	Máximo	0,325	0,321	0,359
IR	Mínimo	0	0	2,843
	Promedio	14,57	35,42	39,19
	Máximo	72	510,955	122,462

Tabla 70: Diferencias según temas para textos emitidos por instruido lego formado para texto no especializado

En el caso del Gráfico 15, se observa una relación con los datos anteriores expuestos en la Tabla 69 y la Tabla 70, ya que la informática es el texto menos presente en los textos no especializados emitidos por legos no formados.

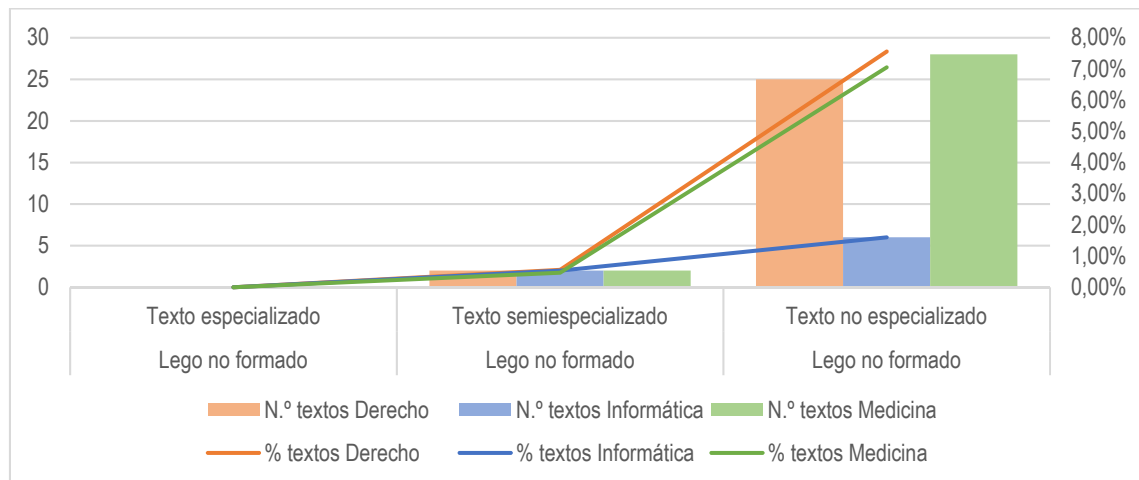


Gráfico 15: Relación de datos globales para la correspondencia emisor lego-tipo de clase según tema

La diferencia principal con respecto a los otros dos temas, como se muestra en la Tabla 71, está en el IR, mucho menor que en las otras categorías, y en la selección de destinatarios, que queda reducida. Llama la atención este hecho, que muestra que los textos que podrían haber sido asignados a esta correspondencia emisor lego no formato-texto no especializado fueron asignados a un emisor instruido lego formado, que constituye la categoría inmediatamente superior al lego no formado. Esto permite defender la existencia de un criterio común por parte de los informantes en considerar que los textos emitidos

sobre informática son emitidos por usuarios con mayor conocimiento en la materia que en otros casos, como el derecho o la medicina, cuyos valores de relación emisor lego formado-tipo de clase son más elevados). Esto quizá esté relacionado con la opacidad terminológica de sus FLE, que suelen ser formas especializadas: acrónimos, formas alfanuméricas, anglicismos, etc.

		Derecho	Informática	Medicina
Destinatario (por orden de frecuencia)		Especialista no investigador, lego formado, lego no formado, instruido semiespecialista.	Lego no formado y lego formado.	Lego no formado, lego formado, especialista investigador. Un caso de semiespecialista y un caso de especialista no investigador.
Tipo de texto		Participación en foro.	Participación en foro.	Participación en foro.
Variedad léxica	Mínimo	53,60	58,80	56,80
	Promedio	63,53	65,89	70,17
	Máximo	67,97	68,53	75,29
IDT	Mínimo	0,017	0,081	0,065
	Promedio	0,11	0,14	0,17
	Máximo	0,358	0,275	0,350
IR	Mínimo	0	0	0
	Promedio	9,59	1,8	29,01
	Máximo	47,818	10,733	91,846

Tabla 71: Diferencias según temas para textos emitidos por lego no formado para texto no especializado

1.1.6. Relación SUBDESTINATARIO -CLASE

Como en el apartado anterior, las diferencias entre temas no son muy representativas. Sin embargo, se pueden encontrar algunos contrastes más destacables con la informática en los casos del destinatario instruido o lego.

La relación del destinatario especialista con las clases de texto (*vid. infra* Gráfico 16) es muy similar a la que existe entre emisor especialista y clase de texto (*vid. supra* Gráfico 13): la clase textual más frecuente es el texto especializado tanto para el subdestinatario especialista investigador (hasta un 22,80 %) como para el especialista no investigador (no alcanza el 10 %), si bien existe una notoria diferencia de frecuencias entre ellos. Debe subrayarse que la distribución de frecuencias por temática es, asimismo, igual en los casos tanto emitidos como recibidos por especialista investigador y especialista no investigador (*vid. supra* Gráfico 8 y Gráfico 9, p. 249), siendo más frecuentes los textos con subdestinatario especialista investigador en medicina (derecho e informática cuentan con una frecuencia más pareja) y, en el caso del destinatario especialista no investigador, son más frecuentes los textos especializados en derecho (siendo, por el contrario, los casos de informática y medicina más homogéneos).

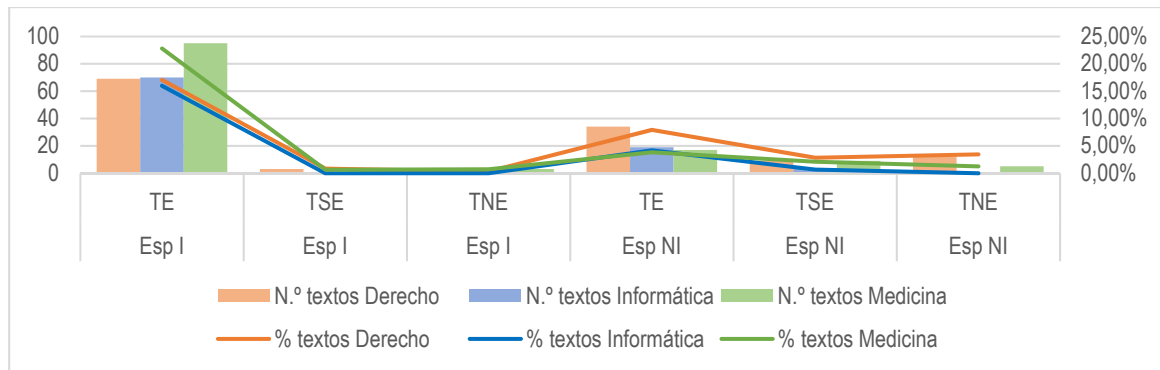


Gráfico 16: Relación de datos globales para la correspondencia destinatario especialista-tipo de clase según tema

Respecto a la diferencia entre derecho e informática y medicina en el texto especializado con destinatario especialista investigador, como puede observarse en la Tabla 72, las diferencias cualitativas son menos de las esperadas, existiendo una clara similitud con el derecho (véanse especialmente los valores para tipo de emisor y tipo de texto). Parece ser que las diferencias cuantitativas son mayores en los otros dos temas, ya que la medicina se sitúa ligeramente por encima del derecho y la informática en la densidad terminológica, y muy por debajo en los valores que corresponden a la reformulación. Es decir, es posible afirmar que la medicina se diferencia del derecho y la informática en los textos especializados recibidos por especialistas en que estos textos médicos cuentan con mayor terminología y menor reformulación.

		Derecho	Informática	Medicina
Emisor		Principalmente especialista investigador. Hay un caso de no investigador y de semiespecialista.	Principalmente especialista investigador. Hay solo un caso de no investigador.	Principalmente especialista investigador. Hay un caso de no investigador y de semiespecialista.
	Tipo de texto	Especialmente tesis doctoral y artículo en revista científica. También existen casos residuales del resto de tipos, a excepción de la participación en foro.	Principalmente tesis doctoral y artículo en revista científica. Solo dos casos de blogs.	Principalmente tesis doctoral y artículo en revista científica. También hay casos residuales del resto de tipos, a excepción de la participación en foro y la noticia.
Variedad léxica	Mínimo	54,60	42,40	48,00
	Promedio	63,14	59,12	56,93
	Máximo	72,20	69,47	73,07
IDT	Mínimo	0,047	0,120	0,093
	Promedio	0,27	0,29	0,32
	Máximo	0,540	0,430	0,576
IR	Mínimo	0	3,548	0
	Promedio	57,66	103,62	47,92
	Máximo	317,463	219,784	116,269

Tabla 72: Diferencias según temas para textos recibidos por destinatario especialista investigador para texto especializado

El caso de las diferencias entre clases para el destinatario especialista investigador (*vid. supra* Gráfico 16) es mucho más notorio, puesto que es posible afirmar que los informantes rechazan considerar otras clases textuales según el grado de especialización si el destinatario del texto es un especialista investigador. En el caso del destinatario especialista no investigador, las frecuencias se difuminan algo más. No obstante, la clase textual de texto especializado sigue siendo la favorita de los informantes cuando el destinatario es un especialista no investigador.

La diferencia principal del Gráfico 16 con respecto al Gráfico 13 se encuentra en los textos semiespecializados (tanto emitidos como recibidos por un especialista no investigador), que son mucho menos frecuentes en el caso de los textos cuyo destinatario es un especialista no investigador. Así, puede identificarse una clara relación entre el destinatario especialista (tanto investigador como no investigador) con el texto especializado en cualesquiera de los temas del corpus.

Por otro lado, en cuanto al destinatario instruido (*vid. infra* Gráfico 17), destaca su vinculación con el texto semiespecializado, aunque es menos acentuada que en el caso de la relación del destinatario especialista con el texto especializado. La clase de texto no especializado queda reservada prácticamente para el destinatario instruido lego formado, siendo escasa en los otros destinatarios en derecho y medicina y muy poco frecuente en informática.

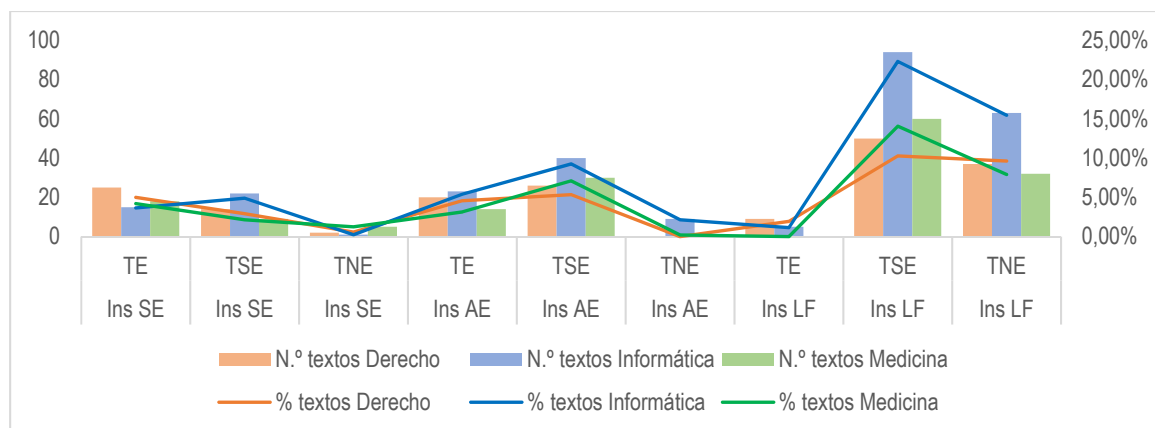


Gráfico 17: Relación de datos globales para la correspondencia destinatario instruido-tipo de clase según tema

Aunque las diferencias entre temas no son muy significativas en la mayoría de casos, sí debe mencionarse que es la informática la que destaca en todas las relaciones por encima del derecho o la medicina (a excepción del destinatario semiespecialista, con una diferencia de menos de un 1 %), sobre todo en los textos con destinatario lego formado, en

los que la informática puede llegar a superar en casi un 10 % a los otros temas, tanto en texto semiespecializado como texto no especializado.

Si centramos nuestra atención en intentar justificar las diferencias de frecuencia entre informática y medicina y derecho para el texto semiespecializado con destinatario lego formado, puede afirmarse que las diferencias cualitativas son muy leves en el caso del emisor o el tipo de texto, y prácticamente inexistentes en el recuento de la densidad terminológica y el índice de reformulación (*vid. infra* Tabla 73).

		Derecho	Informática	Medicina
Emisor		Principalmente especialista no investigador y semiespecialista. También hay lego formado y especialista investigador.	Principalmente especialista no investigador y semiespecialista. También hay casos de lego formado y aprendiz de especialista. Existen casos residuales de especialista investigador y uno de lego no formado.	Principalmente especialista no investigador y semiespecialista. También hay casos de lego formado, aprendiz de especialista y especialista investigador.
Tipo de texto		Todos los tipos a excepción de la tesis doctoral y de dos casos de artículo en revista científica.	Todos los tipos a excepción: solo dos casos de tesis doctoral y uno de artículo en revista científica.	Todos los tipos a excepción de tesis doctoral y de artículo en revista científica.
Variedad léxica	Mínimo	60,62	52,76	54,92
	Promedio	68,36	65,85	67,86
	Máximo	78,43	83,28	82,93
IDT	Mínimo	0,050	0,057	0,073
	Promedio	0,20	0,19	0,21
	Máximo	0,404	0,502	0,506
IR	Mínimo	0	0	0
	Promedio	40,79	40,93	31,27
	Máximo	339,209	512,057	132,632

Tabla 73: Diferencias según temas para textos recibidos por destinatario instruido lego formado para texto semiespecializado

Por ello, se ha decidido profundizar en los datos cuantitativos, especialmente en el recuento de valores mínimos, promedios y máximos para los tipos de términos en el IDT y para los tipos de reformulación en el IR (*vid. infra* Tabla 74). La única diferencia notable se puede hallar en el promedio y en el máximo de la reformulación en informática, que es la menor de todas en el caso de la reformulación por expansión, y la mayor de todas en el caso de la reformulación por reducción.

		Mínimo			Promedio			Máximo		
		DER	INF	MED	DER	INF	MED	DER	INF	MED
IDT	a)	0	0,06	0	0,10	0,09	0,09	0,387	0,336	0,391
	b)	0	0	0,05	0,06	0,04	0,06	0,139	0,156	0,139
	c)	0	0	0	0,04	0,05	0,05	0,135	0,220	0,274
IR	a)	0	0	0	21,42	16,85	19,46	220,216	76,519	124,342
	b)	0	0	0	11,29	17,92	3,88	169,604	388,457	22,712
	c)	0	0	0	8,08	6,16	7,94	84,802	70,629	41,019

Tabla 74: Diferencias según temas para tipos de IDT y de IR en textos con destinatario instruido lego formado para texto semiespecializado

A pesar de estas diferencias, no es posible hallar una relación de causalidad entre la reformulación por expansión y por reducción con el caso de que los textos semiespecializados recibidos por legos formados sean mucho más frecuentes que sus análogos en derecho y medicina.

En otro orden de cosas, atendiendo al otro grupo en el que destaca la informática, los textos no especializados con destinatario lego formado tampoco son distinguibles por sus características cualitativas o cuantitativas, ya que la informática suele hallarse en los valores intermedios (*vid. infra* Tabla 75).

		Derecho	Informática	Medicina
Emisor		Principalmente semiespecialista y lego no formado. Hay casos de lego formado y especialista no investigador.	Principalmente lego formado y semiespecialista. Hay casos residuales del resto de subcategorías.	Principalmente lego formado y semiespecialista. También se incluyen otras categorías, como lego no formado o especialista no investigador.
Tipo de texto		Existen casos para todos los tipos de texto. Los residuales son artículo en revista científica, tesis doctoral y revista o web de divulgación.	Existen casos para todos los tipos de texto a excepción del artículo en revista científica y la tesis doctoral.	Principalmente, blog, noticia y participación en foro, así como tres casos de entrevista.
Variedad léxica	Mínimo	53,60	55,31	63,70
	Promedio	64,81	68,24	72,11
	Máximo	74,05	80,09	80,76
IDT	Mínimo	0,023	0,014	0,032
	Promedio	0,19	0,15	0,15
	Máximo	0,476	0,321	0,336
IR	Mínimo	0	0	9,515
	Promedio	14,35	37,38	52,38
	Máximo	109,154	432,00	281,195

Tabla 75: Diferencias según temas para textos recibidos por destinatario instruido lego formado para texto no especializado

Por ello, como en el caso anterior, se ha recurrido a estudiar los valores mínimos, promedios y máximos de la densidad terminológica según los tipos de término y de los

tipos de reformulación de los textos recibidos por instruido lego formado para texto no especializado (*vid. infra* Tabla 76). Tan solo el valor promedio y el valor máximo del IR tipo b) (es decir, los procedimientos de reformulación por reducción) destacan muy por encima del resto de casos. A pesar de esta diferencia, no puede defenderse que sea una causa que justifique el aumento de frecuencia en el corpus.

		Mínimo			Promedio			Máximo		
		DER	INF	MED	DER	INF	MED	DER	INF	MED
IDT	a)	0,009	0	0	0,13	0,08	0,08	0,452	0,302	0,293
	b)	0	0,02	0	0,05	0,04	0,03	0,151	0,154	0,119
	c)	0	0	0	0,01	0,03	0,03	0,091	0,188	0,102
IR	a)	0	0	4,814	8,86	15,97	37,77	55,883	135,750	218,707
	b)	0	0	0	2,10	16,71	5,73	43,662	360,00	29,487
	c)	0	0	0	3,39	4,71	8,89	19,602	33,635	62,488

Tabla 76: Diferencias según temas para tipos de IDT y de IR en textos recibidos por destinatario instruido lego formado para texto no especializado

Se puede comprobar una diferencia con la vinculación emisor-clase (*vid. supra* Gráfico 14 y Gráfico 17, pp. 255 y 260), que se halla en el contraste entre el semiespecialista y el instruido lego formado, siendo el primero mucho más frecuente en el caso del emisor y el segundo mucho más frecuente en el caso del destinatario. En este último caso, puede afirmarse que la dificultad de establecer un claro límite entre las categorías parece que lleva a los informantes a optar por la categoría del destinatario lego formado, es decir aquel usuario que, sin formación académica o profesional en el tema, cuenta con conocimientos fruto de la autonomía y la documentación por interés personal.

Por último, al estudiar la relación de clases textuales según tema en el destinatario lego (*vid.* Gráfico 18), puede observarse una clara relación entre el destinatario lego no formado y la clasificación de texto no especializado, siendo mucho menos frecuente los textos semiespecializados e inexistentes los textos especializados.

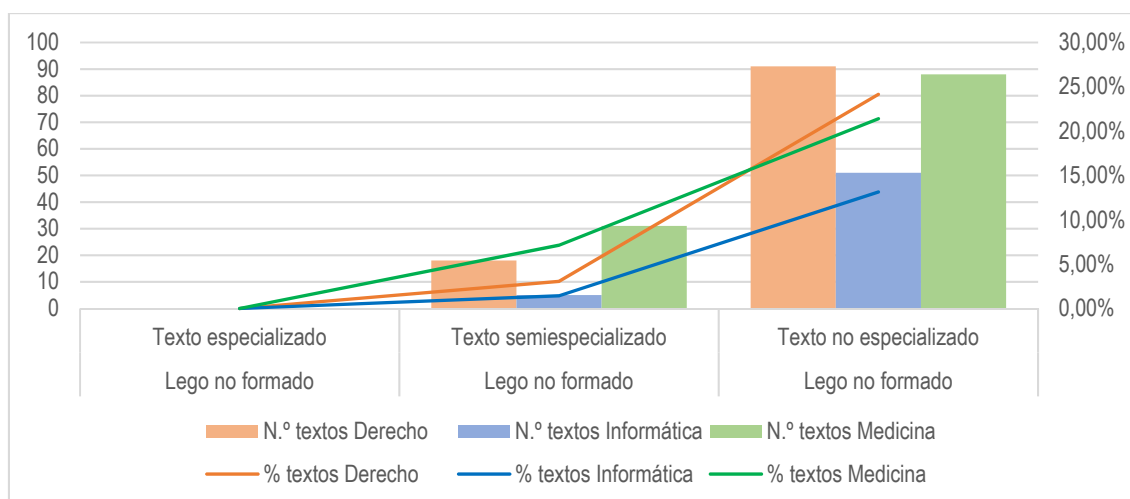


Gráfico 18: Relación de datos globales para la correspondencia destinatario lego-tipo de clase según tema

Las principales características de los textos no especializados con destinatario lego no formado se relacionan en la Tabla 77. La comparación de dichas características atestigua que la densidad terminológica en los textos sobre informática es menor, así como la reformulación terminológica, que resulta mayor que en los otros temas. Además, existe una selección algo más restrictiva de los tipos textuales, aunque puede deberse a la menor frecuencia de los textos informáticos clasificados como no especializados. De forma contraria a lo que previamente podría pensarse, los textos no especializados sobre informática con destinatario lego no formado cuentan con menor densidad terminológica y mayor reformulación terminológica que los textos en derecho o medicina, si bien son menos frecuentes que estos últimos, con una diferencia de aproximadamente un 10 %.

		Derecho	Informática	Medicina
Emisor		Principalmente semiespecialista y lego formado. Existen otros casos de especialista y de lego no formado.	Principalmente semiespecialista y lego formado. Existen otros muy residuales de especialista investigador, aprendiz de especialista o lego no formado.	Principalmente semiespecialista, lego formado y lego no formado. Hay otros casos de especialista.
	Tipo de texto	Son muy frecuentes el blog, la entrevista, la noticia, la participación en foro y la revista o web de divulgación especializada.	Los tipos más frecuentes son el blog, la entrevista o la noticia. La participación en foro o la revista o web de divulgación especializada son menos frecuentes.	Son muy frecuentes el blog, la entrevista, la noticia o la participación en foro. Son escasos los casos de revista o web de divulgación especializada.
Variedad léxica	Mínimo	56,80	58,80	56,80
	Promedio	68,67	70,42	70,79
	Máximo	77,98	82,63	80,76
IDT	Mínimo	0,017	0,088	0,020
	Promedio	0,18	0,11	0,16

	Derecho	Informática	Medicina	
IR	Máximo	0,457	0,386	
	Mínimo	0	0	
	Promedio	10,29	40,13	28,74
	Máximo	72	510,955	122,462

Tabla 77: Diferencias según temas para textos recibidos por destinatario lego no formado para texto no especializado

1.1.7. Relación con las FORMAS LÉXICAS ESPECIALIZADAS

Este apartado relaciona el número de formas léxicas especializadas que identifica el conjunto de grupos de informantes, pero no la densidad terminológica resultado de la operación de dividir términos entre formas léxicas totales, que se estudia en profundidad en relación con todas las variables en IV. 1.4. (p. 307). Se recuerda que las formas léxicas especializadas que pueden identificar los informantes se dividen en tres grupos (*vid. supra* III. 6.7.1., p. 223):

Tipo de FLE	Descripción
FLE a) <i>Términos délficos</i>	Conocimiento de la denominación y del concepto, no necesariamente con precisión, sino posiblemente de forma laxa. El informante lo reconoce como frecuente en el discurso común.
FLE b) <i>Términos semicrípticos</i>	Conocimiento de la denominación, pero no del concepto o conocimiento muy pobre del concepto. El informante lo reconoce como poco frecuente en el discurso común.
FLE c) <i>Términos crípticos</i>	Conocimiento nulo de la denominación o del concepto. El informante lo reconoce como poco frecuente en el discurso común.

Tabla 78: Descripción de los tipos de FLE que distinguen los grupos de informantes

En el Gráfico 19 puede observarse la diferencia entre los tres tipos de términos que distinguen los informantes: el tipo a) es el más frecuente, seguido del tipo b) y muy estrechamente del tipo c).

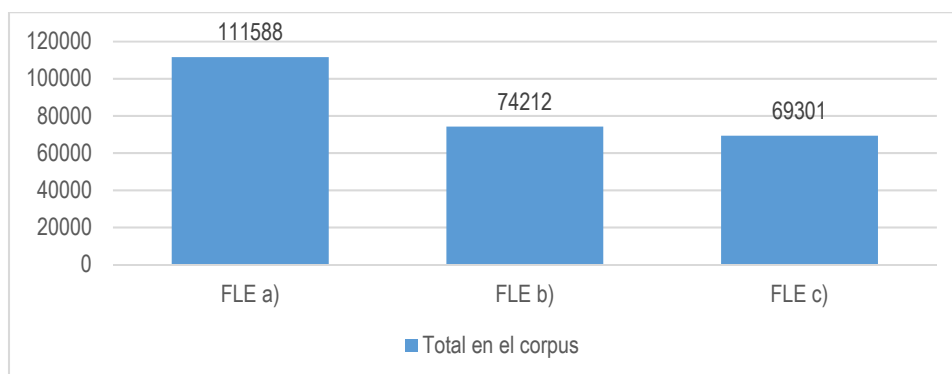


Gráfico 19: Datos globales de tipos de FLE en el corpus

Si distinguimos los tipos de términos según los temas, a pesar de las similitudes en la distribución de los tipos de términos —por ejemplo, que en los tres temas los términos más frecuentes son los del tipo a)—, se aprecian diferencias entre los tres temas (*vid. infra* Gráfico 20). Si comparamos los temas en cada tipo de FLE, las diferencias más importantes se encuentran en los términos délficos y en los términos crípticos, en los que destaca el derecho y la medicina respectivamente. La cantidad de términos semicrípticos son muy similares en los tres temas.

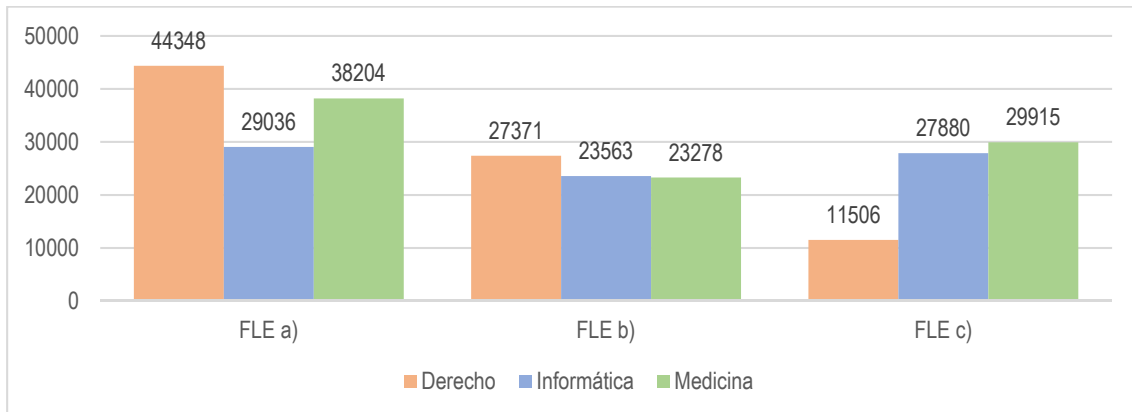


Gráfico 20: Datos globales de tipos de FLE según tema

Atendiendo a la distinción temática, la medicina es el tema con mayor cantidad de terminología (91 397 FLE), seguida del derecho (83 225 FLE) y de la informática (80 479 FLE).

Si centramos la atención en los temas y en la distribución de los tipos de FLE debe mencionarse, en primer lugar, que el derecho es el que más se inclina hacia las formas délficas y el que menos presencia de formas crípticas tiene. En segundo lugar, la informática cuenta con una cantidad de FLE similar en todos los tipos, si bien, como el derecho, cuenta con más términos délficos y menos términos semicrípticos. Por último, la medicina despunta algo en cuanto a su similitud con la informática, ya que cuenta con multitud de formas délficas y, a pesar de que los términos semicrípticos y términos crípticos son similares, cuenta con mayor cantidad de los segundos, con una diferencia mayor con respecto a la informática. En la Tabla 79, la Tabla 80 y la Tabla 81 se relacionan varios ejemplos por tipo de FLE y tema, en los que se ilustran los términos que han sido marcados por los informantes (en este caso el grupo de estudiantes).

Derecho

FLE a) A juicio del Tribunal los progenitores no poseen un permiso de residencia permanente sino una mera autorización de estancia por razones humanitarias que puede ser anulada y que, según el ordenamiento interno Suizo, no les da derecho a la reunificación familiar.

DER ART04 (Estudiante)

Derecho

FLE b) Los elementos del **enunciado legal** de la **estafa** española son cinco: a) Utilización de un «**engaño bastante**» por parte del **autor**; b) Inducción a la víctima a la realización de un acto de **disposición patrimonial perjudicial**; c) **Acto** de **disposición patrimonial**; d) **Perjuicio patrimonial**; e) **Dolo y ánimo de lucro**.

DER TES05 (Estudiante)

FLE c) Su art. 4 señala que «son componentes fundamentales del ejercicio de la **democracia** la transparencia de las actividades gubernamentales, la **probidad**, la **responsabilidad** de los gobiernos en la **gestión pública**, el respeto por los **derechos sociales** y la **libertad** de expresión y de prensa (...)».

Seguindo a la anterior Carta, ha de destacarse el texto al que España adhirió y aprobó recientemente, la **Carta Iberoamericana de Participación Ciudadana en la Gestión Pública**.

DER ART02 (Estudiante)

Tabla 79: Ejemplos de tipos de términos en derecho

Informática

FLE a) El **sistema** trabaja sobre un **canal de comunicaciones ranurado**, cada **ranura de tiempo** representa el tiempo necesario para la **transmisión** de un **paquete de datos**. En este caso, consideramos que la **estación base**, que en principio no tiene restricciones energéticas ya que puede estar conectada directamente a la **red eléctrica**, envía periódicamente un **indicador de inicio de ranura** para sincronizar a todos los **nodos de la red**.

INF ART05 (Estudiante)

FLE b) La tercera idea consiste en aprovechar la **ilusión somatográfica** (MacNeilage, Banks et al. 2007), para hacer percibir al cerebro las **aceleraciones sostenidas** (eliminadas por la **aplicación de filtros pasa-alta**) mediante pequeñas **inclinaciones** que modifican la **dirección relativa del vector gravedad** con respecto al sistema de referencia del **aparato vestibular**. Esta idea se implementa aplicando **filtros pasa-baja** a las **aceleraciones lineales** para inclinar ligera y lentamente la **plataforma** en los **ejes lateral y longitudinal**, en función del valor de estas **aceleraciones filtradas**.

INF ART13 (Estudiante)

FLE c) WISPER (Wireless sensor network alternating cntm/cr Slots protocol for Preemptive Event Reporting) es un protocolo híbrido para WSN basado en TDMA que contempla el uso de radio cognoscitivo para el **reporte de eventos**. Como se puede observar en la Fig. 1, la **ronda** inicia con una etapa de formación de cluster (CF), en la cual los **nodos** se organizan en grupos para mejorar la **eficiencia energética** de las **transmisiones**. En este caso, antes de enviar su **joint packet** (**mensaje de control** que le proporciona al CH correspondiente el **id del nodo** en cuestión para ser incluido en el **cluster**), cada uno de los **CMs** realiza un proceso de selección de **red**, lo cual determinará sus características de funcionamiento hasta la siguiente etapa de CF.

INF ART05 (Estudiante)

Tabla 80: Ejemplos de tipos de términos en informática

Medicina

FLE a) Los **factores** que influyen sobre el **pronóstico** también son muy heterogéneos, abarcando desde **factores demográficos, clínicos, estructurales, biológicos** (63) y otros **factores** especialmente relevantes en ancianos como la **comorbilidad**, la **dependencia funcional** o los **trastornos cognitivos** (8) (63) (64). A continuación se detallan los **factores pronósticos** más relevantes.

MED TES15 (Estudiante)

FLE b) Medialmente la **clavícula** forma parte de la **articulación esterno-clavicular** con **superficies articulares** están cubiertas de **fibrocárilago** que están separadas por un **menisco** o **disco articular** que divide la **articulación** en dos **recesos** separados, aumentando la **estabilidad** ⁽³¹⁾. Es el **elemento de sujeción** de la **extremidad superior** al **esqueleto axial**, aunque los movimientos que realiza son los de **elevación** y **descenso** además de **antepulsión** y **retropulsión**. Sus **refuerzos capsulares** son

Medicina

los **ligamentos esterno-claviculares anterior, posterior y el interclavicular**, aunque el que proporciona más **estabilidad articular** es el **ligamento costo-clavicular**.

MED TES07 (Estudiante)

FLE c)

Los **microtúbulos** son cilindros huecos con un diámetro aproximado de 25 nm formados por la **asociación longitudinal** de protofilamentos de α, β -tubulina.⁵⁶ Los **microtúbulos** se pueden clasificar en **microtúbulos estables**, que son los que constituyen los **cilios y flagelos**, y los **cambiantes o dinámicos**, que se localizan en el **citoplasma**. Los **microtúbulos citoplasmáticos** forman el huso mitótico y son los encargados de la separación de los **cromosomas** en la fase de **mitosis** durante la **división celular**. También participan en el movimiento de **orgánulos**, como las **mitocondrias, lisosomas, pigmentos, gotas de lípidos**, etcétera y en la organización del **tráfico vesicular**.

MED TES01 (Estudiante)

Tabla 81: Ejemplos de tipos de términos en medicina

1.1.8. Relación con los PROCEDIMIENTOS DE REFORMULACIÓN FORMAL

Como en el caso anterior, este apartado enumera los procedimientos de reformulación formal que se emplean puestos en marcha para variar, describir o explicar la terminología de los textos. Este apartado no se dedica a la densidad reformulativa en un texto, que se estudia en profundidad en IV. 1.5. (p. 324).

En el Gráfico 21 se indica el número total procedimientos de reformulación formal que existen en el corpus. Los procedimientos por expansión despuntan muy por encima de los de reducción o variación.

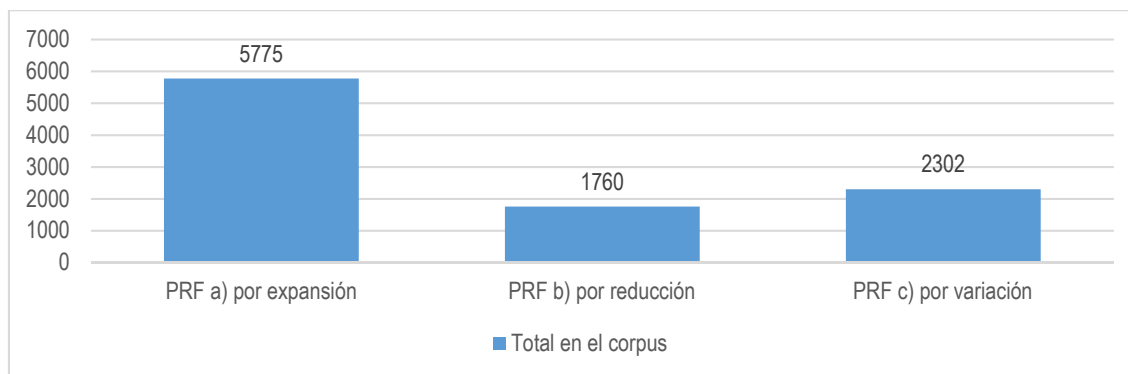


Gráfico 21: Datos globales de tipos de PRF en el corpus

Atendiendo a la distinción temática, como indica el Gráfico 22, el derecho es el tema menos reformulado, seguido de la informática y de la medicina, las cuales se diferencian tan solo por 300 PRF. En el caso de la medicina, no solo se trata del tema con mayor número de términos, sino del tema con mayor número de procedimientos de reformulación. En el caso del derecho y de la informática, la relación es inversa.

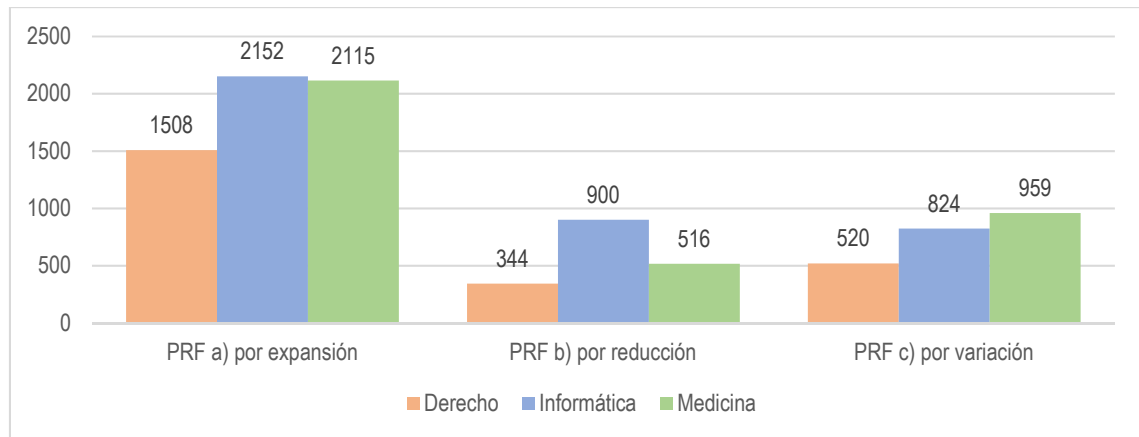


Gráfico 22: Datos globales de tipos de PRF según tema

El hecho de que el derecho sea el tema menos reformulado puede relacionarse con los datos obtenidos en el apartado anterior (*vid. supra* IV. 1.1.7., p. 265) que defienden que los términos más frecuentes son aquellos reconocidos como frecuentes en el discurso común, pero con conocimiento del concepto o aquellos poco frecuentes en el discurso común con conocimiento muy pobre del concepto, i. e. términos délficos y semicrípticos. Desde el punto de vista opuesto, la mayor reformulación de la medicina podría relacionarse quizá con el mayor número de términos crípticos. En estos dos ejemplos, se observa la relación entre PRF y terminología basada en el principio de relevancia de Sperber y Wilson (1995: 260).

Si centramos la atención en los temas y en la distribución de los tipos de PRF debe resaltarse que la informática y la medicina son los que más se inclinan hacia la reformulación por expansión y por variación (con un número de procedimientos muy similar). La informática destaca por encima del derecho y de la medicina en los procedimientos de reducción, posiblemente por el número de siglas, acrónimos y denominaciones que se usan en los textos. En la Tabla 82, la Tabla 83 y la Tabla 84 se relacionan varios ejemplos por tipo de PRF y tema, en los que se ilustran los procedimientos, así como los términos, que han sido marcados por los informantes (en este caso el grupo de estudiantes).

Derecho	
PRF a) por expansión	Si bien a la precisión de dichos conceptos dedicaremos el primer capítulo de nuestro trabajo, <u>con la utilización de la palabra «obra» o «creación» pretendemos identificar el resultado de la actividad de carácter creativo de un sujeto o varios.</u> Con «obra culinaria» o «creación culinaria» nos referiremos, pues, al resultado de la <u>actividad creativa culinaria</u> de un «chef». DER TES02 (Estudiante)
FLE b) por reducción	Desde el primer capítulo de este trabajo se dijo que podría distinguirse entre el sujeto que brinda la información (<u>experto</u>) y la información brindada por éste (<u>expertise</u>). DER TES13 (Estudiante)

Derecho

Así, reconoce **efectos civiles** a los **matrimonios celebrados** ante el **juez, alcalde o funcionario** señalado por el **Código —matrimonios en forma civil—**, y a los celebrados en la **forma religiosa** legalmente prevista **—matrimonios en forma religiosa—**

DER ART03 (Estudiante)

El primer problema que nos encontramos es la variedad de **términos** que se usan en la **literatura científica** para definir lo que una parte de la **comunidad científica** denomina **ciberbullying** (incluso se usa “**cyberbullying**”). Ejemplo de ellos son **bullying electrónico**, **bullying a través de Internet**, **acoso a través de Internet** y **acoso online**.

FLE c)
por
variación

DER TES14 (Estudiante)

La figura del experto como **consejero** del **juez o «technical advisor»** coexiste en Estados Unidos junto a otras figuras distintas como los «**court-appointed experts**» (**expertos nombrados por las partes**) y de «**specials master**».

DER TES08 (Estudiante)

Tabla 82: Ejemplos de tipos de PRF en derecho

Informática

PRF a)
por
expansión

donde $E T_1(h)$ es la **probabilidad de clasificar erróneamente los píxeles** en la clase $(h+1)$, cuando realmente pertenecen a la clase h , mientras que $E T_2(h)$ **representa la probabilidad de clasificar inadecuadamente los píxeles** en la clase h ,

INF ART11 (Estudiante)

FLE b)
por
reducción

Para el **control** de los **movimientos** de la plataforma se diseñan unos **algoritmos de control específicos** conocidos como **algoritmos de generación de claves gravitoinerciales**.

(INF ART13) (Usuario medio)

FLE c)
por
variación

Algunas de las **técnicas** más simples proponen transformar el **problema multiobjetivo** en un problema escalar, ponderando las **funciones objetivo**, como es el caso del **método combinación lineal de pesos** (o **método de factores ponderados**).

INF ART08 (Estudiante)

Tabla 83: Ejemplos de tipos de PRF en informática

Medicina

PRF a)
por
expansión

Los **microtúbulos** son **cilindros huecos con un diámetro aproximado de 25 nm** formados por la **asociación longitudinal de protofilamentos de α,β -tubulina**.⁵⁶ Los **microtúbulos** se pueden clasificar en **microtúbulos estables**, que son los que constituyen los **cilios y flagelos**, y los **cambiantes o dinámicos**, que se localizan en el **citoplasma**.

MED TES01 (Estudiante)

FLE b)
por
reducción

La descarga que efectúa la excitación del músculo cardíaco proviene en estos casos de un punto cualquiera de las **aurículas** o de los **ventrículos** (el corazón está formado por dos **aurículas** y dos **ventrículos**), pero una fracción de tiempo muy pequeña antes de la llegada del impulso sinusal normal (que es el originado en una parte especial del corazón llamado “**nodo sinusal**”, que es el **lugar de inicio normal de los impulsos eléctricos que hacen que el corazón se contraiga e impulse la sangre**).

(MED FOR02) (Usuario medio)

FLE c)
por
variación

Así se ha podido establecer que los **tumores** tienen en general un **coeficiente α/β** alto que se sitúa entre 7 y 25, que los **tejidos** normales con mayor **destrucción celular directa o también llamados de respuesta aguda** sus valores α/β están alrededor de 10 y que los **tejidos de respuesta lenta** donde predomina la **lesión subletal** y por tanto más **radiorresistentes**, tienen unos valores α/β entre 1,7 y 4,5.

MED TES14 (Estudiante)

Tabla 84: Ejemplos de tipos de PRF en medicina

1.1.9. Conclusiones parciales sobre el TEMA

A continuación, se relaciona una recapitulación de las interpretaciones más relevantes sobre los datos relacionados con los temas del corpus.

- a) Los tres temas están representados de forma equilibrada en las tres clases según la especialización.
- b) El emisor especialista es más numeroso en derecho y medicina, mientras que el emisor instruido es más numeroso en informática. En cuanto a los subemisores, la medicina se caracteriza por contar con mayor número de textos emitidos por un especialista investigador, y el derecho por contar con mayor número de textos emitidos por un especialista no investigador. La informática, por otro lado, se caracteriza por contar con mayor número de textos emitidos por instruidos legos formados.
- c) El destinatario más frecuente coincide en los tres temas: el instruido.
- d) El patrón de distribución de número de textos emitidos por los diferentes subemisores y el del número de textos recibidos por los diferentes subdestinatarios coincide.
- e) Los grupos de informantes consideran que, para poder emitir un texto, dicho emisor debe mostrar dominio conceptual sobre un tema. No obstante, los destinatarios de los textos solo están sujetos al conocimiento enciclopédico o compartido según los grupos sociales, lo que hace que se identifique un mayor número de legos formados y legos no formados en los textos.
- f) Existe una clara relación entre emisor especialista investigador y destinatario especialista investigador en todos los temas.
- g) En la relación entre emisor semiespecialista y el resto de subdestinatarios se hallan las principales diferencias entre temas, así como en la relación entre emisor lego formado y destinatario lego formado o lego no formado.
- h) No existen diferencias representativas entre los temas en la combinación de subemisor-clase. La más frecuente es la relación emisor especialista investigador-clase texto especializado, seguida de la relación emisor instruido semiespecialista-texto semiespecializado y texto no especializado, y por último de la relación emisor lego no formado-texto no especializado (a excepción de la informática, que opta por el emisor instruido lego formado para los textos semiespecializados y no especializados).

- i) Tampoco se encuentran diferencias notables entre temas en la combinación de subdestinatario-clase. El destinatario especialista queda vinculado al texto especializado, el destinatario lego no formado queda relacionado con el texto no especializado (a excepción, de nuevo, de la informática, que se inclina hacia la relación instruido lego formado-texto semiespecializado y texto no especializado, con una posible vinculación a la mayor reformulación por reducción de estos textos).
- j) Existen diferencias en la cantidad de términos con las que cuenta cada tema: la medicina es el tema con mayor cantidad de terminología, seguida del derecho y, este último, muy próximo a la informática. El derecho es el que más se inclina hacia los términos délficos y el que menos presencia de términos crípticos tiene. En segundo lugar, la informática cuenta con una cantidad de FLE similar en todos los tipos, si bien, como el derecho, cuenta con más formas délficas y menos formas semicrípticas. Por último, la medicina despunta algo en cuanto a su similitud con la informática, ya que cuenta con multitud de términos délficos y, a pesar de que los términos crípticos y términos semicrípticos son similares, cuenta con mayor cantidad de los segundos, con una diferencia mayor con respecto a la informática.
- k) Los procedimientos de reformulación más numerosos son los procedimientos por expansión, seguidos de los procedimientos por variación y por reducción. Existen, asimismo, diferencias entre los temas, siendo el derecho el tema con menos procedimientos de reformulación. La informática y la medicina cuentan con algunas diferencias muy leves; la mayor de ellas es que la informática supera a la medicina en los procedimientos por reducción.

1.2. Análisis de la CLASE

Las tres clases de texto están muy compensadas en el corpus, ya que las tres rondan el 30 %: el texto no especializado es el que menor representación tiene en el corpus (31,83 %); el texto semiespecializado está representado en el 33,81 % del corpus y el texto especializado en el 34,37 %.

1.2.1. Relación con el TIPO DE TEXTO

En la relación entre tipos de texto y clase puede comprobarse que existen dos bloques claramente diferenciados: el de texto especializado y el conjunto de texto semiespecializado y texto no especializado (*vid. infra* Gráfico 23).

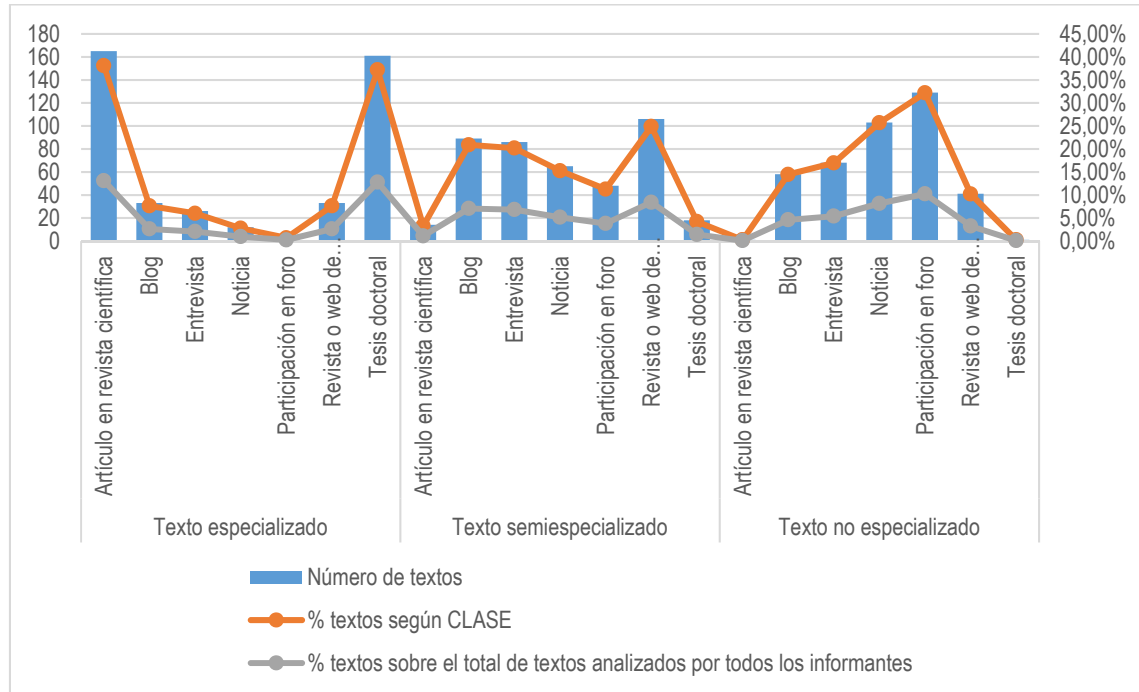


Gráfico 23: Datos globales de tipo de texto según clase

Para representar visualmente la tendencia de distribución de los tipos textuales en las tres clases nos hemos servido de los cálculos probabilísticos que se establecen en la Tabla 146 del apartado 3.1., de forma que la Tabla 85 muestra la proximidad de cada tipo textual a cada uno de los límites que existen entre las clases según el grado de especialización.

TE	TSE	TNE
ART TES	DIV BLOG ENT	NOT FOR

Tabla 85: Distribución de tipos textuales según clases

En el primero, los textos con más presencia son el artículo en revista científica y la tesis doctoral (aunque también aparecen algunos casos en el texto semiespecializado), como podría ser previsible *a priori*. En el segundo grupo se incluyen los otros cinco tipos de texto con una frecuencia similar en ambas clases. Cabe mencionar que, no obstante, se aprecia una relación inversamente proporcional, ya que cuanto más frecuentes son los tipos

en el texto semiespecializado, menos lo son en el no especializado, y viceversa (lo cual es lógico desde el punto de vista de que estos tipos de texto se reparten casi íntegramente entre las dos clases con menor grado de especialización).

En el texto especializado, existen tres casos que, a pesar de ser poco frecuentes, cuentan con una frecuencia que ronda en cada uno el 7 %. Se trata del blog, la entrevista y la revista o web de divulgación especializada. A continuación, se presentan tres tablas que pretenden justificar la presencia de estos textos en la clase reservada a textos especializados. En la Tabla 86 se muestran las características atribuidas al artículo en revista científica y tesis doctoral en los textos especializados. Entre ellas destaca el emisor y el destinatario investigador y no investigador, una densidad terminológica elevada, entre 0,27, y un índice de reformulación también elevado, en torno a 80.

		Artículo en revista científica	Tesis doctoral
Emisor		Especialista investigador y no investigador. Solo un caso de semiespecialista.	Especialista investigador y no investigador.
Destinatario		Especialista investigador y no investigador. También existen casos de todas las subcategorías de instruido.	Especialista investigador y no investigador. También existen casos de todas las subcategorías de instruido.
Variedad léxica	Mínimo	48,00	42,40
	Promedio	59,02	60,07
	Máximo	73,05	69,47
IDT	Mínimo	0,047	0,060
	Promedio	0,25	0,30
	Máximo	0,525	0,576
IR	Mínimo	0	9,509
	Promedio	82,51	75,23
	Máximo	378,646	418,403

Tabla 86: Similitudes entre artículo en revista científica y tesis doctoral especializados

En la Tabla 87 se relacionan las características para el blog, la entrevista o la revista o web de divulgación especializada clasificados como textos especializados. Como se puede comprobar, las categorías de emisor y destinatario coinciden con las de la Tabla 86, así como el valor promedio del IDT, aunque no la variedad léxica (que es mayor en los casos de blog, entrevista y revista o web de divulgación especializada) o el índice de reformulación, que es mayor en los casos de artículo en revista científica y tesis doctoral). A continuación, se emplea el color verde para mostrar los valores numéricos próximos.

	Blog	Entrevista	Revista o web de divulgación especializada
Emisor	Principalmente especialista investigador y no investigador. Existe un caso de semiespecialista.	Principalmente especialista investigador y no investigador, así como semiespecialista.	Principalmente especialista investigador y no investigador. Existen dos casos de semiespecialista.

		Blog	Entrevista	Revista o web de divulgación especializada
Destinatario		Especialista investigador y no investigador y todas las subcategorías de instruido.	Especialista investigador y no investigador y todas las subcategorías de instruido.	Especialista investigador, no investigador, semiespecialista y aprendiz de especialista.
Variedad léxica	Mínimo	54,60	57,44	50,60
	Promedio	65,16	64,78	64,25
	Máximo	82,93	72,20	80,27
IDT	Mínimo	0,130	0,095	0,114
	Promedio	0,32	0,24	0,30
	Máximo	0,558	0,361	0,487
IR	Mínimo	0	0	0
	Promedio	17,11	29,98	21,54
	Máximo	66,0	87,500	52,448

Tabla 87: Similitudes entre blog, entrevista y revista o web de divulgación especializada especializados

Para tratar de justificar que estos textos hayan sido clasificados como especializados, recurrimos a las características que se les atribuye a blog, entrevista y revista o web de divulgación especializada en el bloque de textos semiespecializado y no especializado, como se relaciona en la Tabla 88.

		Blog	Entrevista	Revista o web de divulgación especializada
Emisor		Semiespecialista y especialista no investigador. Otros casos son especialista investigador, aprendiz de especialista y lego formado.	Semiespecialista y especialista no investigador. Otros casos son investigador y lego formado.	Semiespecialista y especialista no investigador. Otros casos son lego formado.
Destinatario		Lego no formado y lego formado. Otros casos son semiespecialista y aprendiz de especialista.	Lego no formado y lego formado. Otros casos son semiespecialista, aprendiz de especialista y especialista no investigador.	Lego no formado y lego formado. Otros casos son semiespecialista y aprendiz de especialista.
Variedad léxica	Mínimo	57	57,44	54,92
	Promedio	67,31	67,97	67,45
	Máximo	82,93	83,28	80,27
IDT	Mínimo	0,044	0,013	0,023
	Promedio	0,24	0,14	0,22
	Máximo	0,503	0,361	0,506
IR	Mínimo	0	0	0
	Promedio	21,87	68,77	26,57
	Máximo	181,000	796,533	146,571

Tabla 88: Similitudes entre blog, entrevista y revista o web de divulgación especializada semiespecializados y no especializados

En los casos relacionados en la tabla anterior, se observa que la principal diferencia con los textos especializados se halla en la relación que se establece entre emisor y

destinatario, que deja de ser casi exclusiva del especialista y que la densidad terminológica es menor en todos los casos. Se observa, por tanto, que la vinculación con el texto especializado está relacionada con la densidad terminológica y las relaciones discursivas según la competencia cognitivo-comunicativa de los usuarios.

En otro orden de cosas, en el Gráfico 23 también se observan algunos casos muy reducidos de artículo en revista científica y tesis doctoral clasificados como textos semiespecializados (no superan el 1 % de la clase). Para dar respuesta a esta clasificación, se indican las características para estos textos en la Tabla 89 (marcamos en rojo las diferencias principales que se observan).

		Artículo en revista científica	Tesis doctoral
Emisor		Especialista investigador y no investigador. Solo un caso de aprendiz de especialista.	Principalmente especialista investigador y semiespecialista. Solo dos casos de especialista no investigador.
Destinatario		Principalmente aprendiz de especialista, semiespecialista y lego formado.	Principalmente aprendiz de especialista y semiespecialista. Otros casos de lego formado y especialista investigador.
Variedad léxica	Mínimo	54,27	59,07
	Promedio	59,51	64,30
	Máximo	65,12	68,53
IDT	Mínimo	0,071	0,064
	Promedio	0,20	0,17
	Máximo	0,389	0,410
IR	Mínimo	13,657	11,299
	Promedio	164,55	42,79
	Máximo	339,209	77,708

Tabla 89: Similitudes entre artículo en revista científica y tesis doctoral semiespecializados

Las principales diferencias con las características relacionadas en la Tabla 86 se marcan en rojo: el destinatario deja de ser especialista, para formar parte de la categoría de instruido. A su vez, existe una reducción en el valor del IDT, que resulta más próximo a la media de IDT para texto semiespecializado (0,21) que a la media de IDT para texto especializado (0,28).

Por último, en el contraste de textos semiespecializados y textos no especializados del Gráfico 23 también se puede hallar dos tipos de texto que cuentan con diferencias más notables que sus análogos: se trata de la participación en foro y de la revista o web de divulgación especializada. En la Tabla 90, se especifican las características de estos textos en las dos clases, de forma que puedan contrastarse.

		Texto semiespecializado		Texto no especializado	
		Participación en foro	Revista o web de divulgación especializada	Participación en foro	Revista o web de divulgación especializada
Emisor		Principalmente lego formado y especialista no investigador. Existen casos de	Semiespecialista y especialista no investigador. Hay casos de investigador y lego formado.	Principalmente lego no formado y lego formado. Existen casos de	Semiespecialista y lego formado. Existen otros casos de especialista no investigador.
Destinatario		semiespecialista y lego no formado.	Lego formado, semiespecialista, lego no formado y aprendiz de especialista.	semiespecialista y especialista no investigador.	Principalmente lego no formado y lego formado. Solo dos casos de aprendiz y uno de semiespecialista.
Variedad léxica	Mínimo	52,76	54,92	53,60	56,32
	Promedio	63,56	66,51	67,66	69,90
	Máximo	80,09	80,27	89,54	77,51
IDT	Mínimo	0,088	0,052	0,017	0,023
	Promedio	0,18	0,24	0,14	0,18
	Máximo	0,355	0,506	0,358	0,439
IR	Mínimo	0	0	0	0
	Promedio	20,38	29,30	22,12	19,52
	Máximo	140,000	146,571	122,462	84,160

Tabla 90: Similitudes y diferencias entre participación en foro y revista o web de divulgación especializada en las clases de texto semiespecializado y de texto no especializado

El dato que más resalta en los dos casos es la gran diversidad de emisores y destinatarios con los que cuentan los tipos de texto. De hecho, los tipos de emisor y destinatario se suelen repetir en todos los casos y, a nuestro juicio, no resultan pertinentes en la discriminación de las clases. Tampoco resulta relevante la variedad léxica o la reformulación, si bien la terminología sí es menor en los textos no especializados (se aproximan a la media de IDT para textos no especializados, que es de 0,16). Así, como en el caso anterior, parece que existe vinculación entre la terminología y la clasificación textual según el grado de especialización.

1.2.2. Relación con el EMISOR y el SUBEMISOR

Las relaciones entre clase y emisor siguen algunos de los presupuestos de la TCT: los textos especializados son emitidos principalmente por especialistas (95 % sobre el total de la clase), seguidos de un porcentaje muy reducido de instruidos (principalmente semiespecialistas) y ausencia total de emisores legos (*vid. infra* Gráfico 24). Los textos semiespecializados y no especializados son más variados, aunque predomina principalmente la categoría de emisor instruido. En el caso concreto de texto

semiespecializado, puede decirse que guarda algunas similitudes con el texto especializado, en tanto que el lego supone menos de un 2 % de los emisores y el especialista cuenta con representación de un 37 % sobre la clase. El instruido es el emisor más relacionado con el texto semiespecializado, lo que supone que los dos entes de transición guardan cierta vinculación. La principal diferencia con el texto especializado se halla en que el texto semiespecializado cuenta con mayor presencia de emisores especialistas no investigadores (25,59 % en el texto semiespecializado y 15,94 % en el texto especializado). Ambas clases coinciden en el subemisor más frecuente, el semiespecialista (aunque debe resaltarse la estrecha relación entre texto semiespecializado y subemisor semiespecialista, mucho más notoria que en el texto especializado).

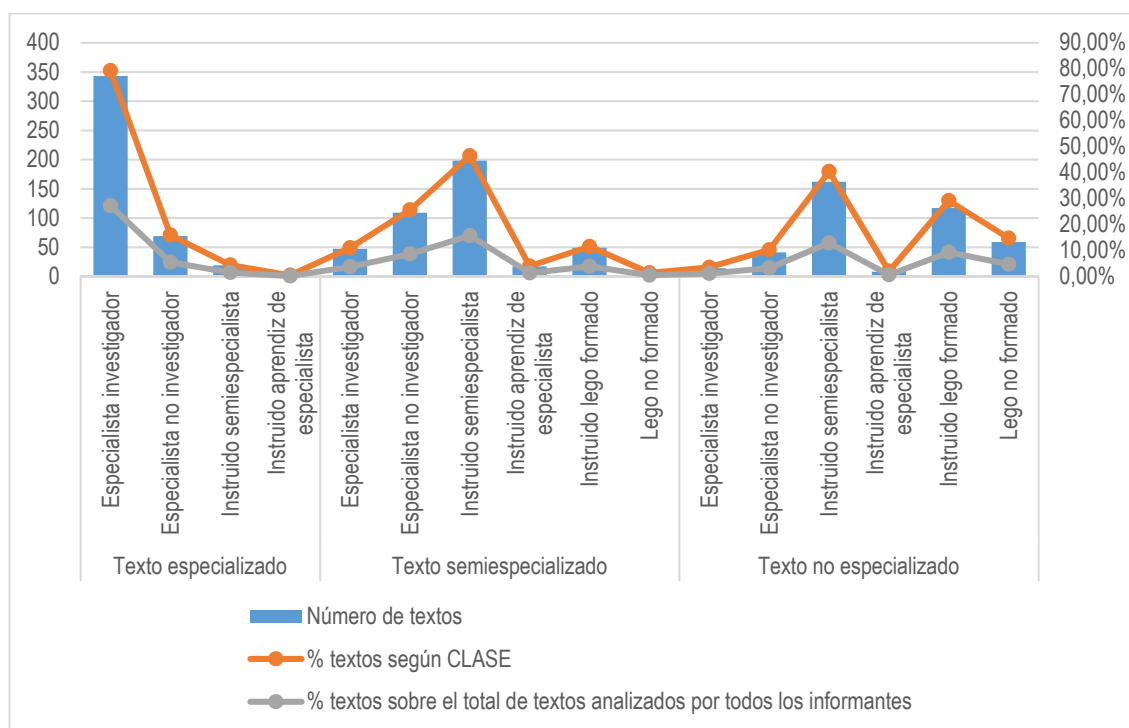


Gráfico 24: Datos globales de subemisor según clase

Por último, el texto no especializado también queda vinculado al emisor instruido (principalmente instruido semiespecialista e instruido lego formado, que se trata de la subcategoría más cercana al lego) así como al lego no formado, que cuenta con un 14,71 % de representación entre los emisores para el texto no especializado. De hecho, es en el caso de texto no especializado donde más frecuente es el lego no formado de todo el corpus, lo cual se relaciona estrechamente con su opuesto especialista-texto especializado y los fundamentos teórico-metodológicos de la TCT (*vid. infra* Tabla 93: Correspondencias más frecuentes entre subemisor y subdestinatario según clases).

1.2.3. Relación con el DESTINATARIO y el SUBDESTINATARIO

El destinatario de los textos según la clase conserva similitudes, pero también presenta divergencias con el emisor (*vid. supra* Gráfico 24 y *vid. infra* Gráfico 25). En primer lugar, entre algunas analogías, en el texto especializado prima la categoría de especialista (69,75 %), a la que le sigue el instruido (29,79 %). Cuenta con tan solo dos textos con destinatario lego (0,46 %). Existen, además, otras semejanzas con el emisor en el texto semiespecializado, como, por ejemplo, la alta ocurrencia de la categoría instruido frente al especialista o al lego. Sin embargo, son el aprendiz de especialista (22,54 %) y el lego formado (47,89 %) las categorías más frecuentes para el destinatario en el texto semiespecializado. Además, el lego no formado también cuenta con representación entre los destinatarios para esta clase de texto, la cual no tenía entre los emisores.

En el caso del texto no especializado, las categorías con mayor competencia cognitivo-comunicativa quedan descartadas, siendo el lego formado y el lego no formado las que, combinadas, alcanzan el 94,76 % de representación del texto no especializado.

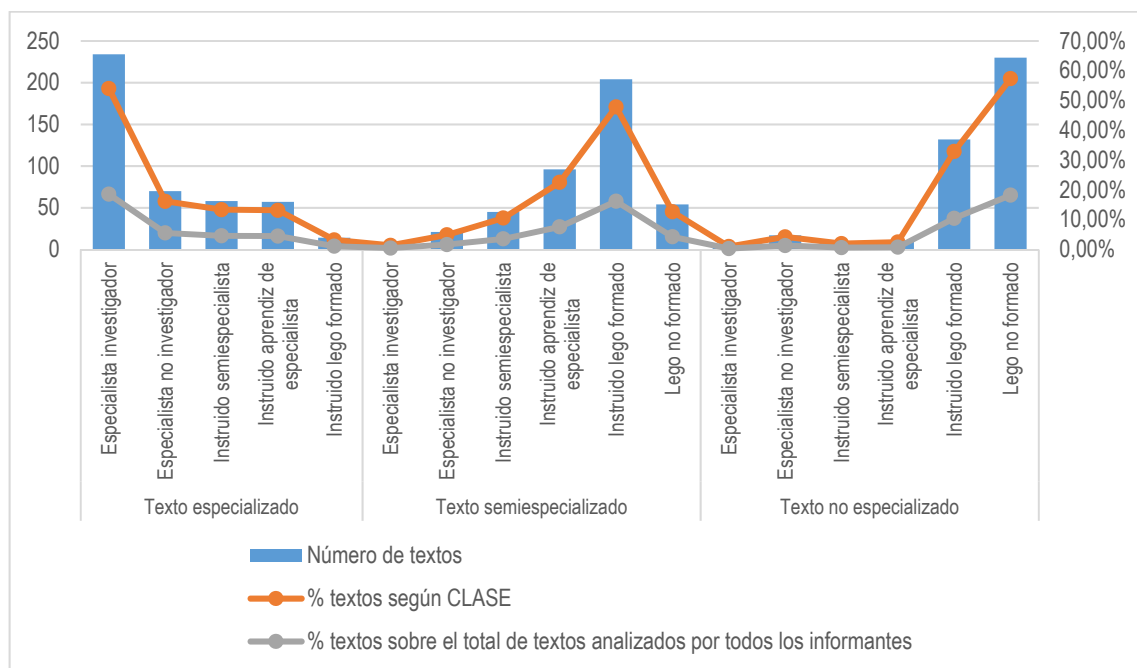


Gráfico 25: Datos globales de subdestinatario según clase

En el Gráfico 26 se comprueban las diferencias principales entre la relación emisor-clase y destinatario-clase. Las zonas marcadas en gris reflejan las grandes diferencias entre subemisor y subdestinatario en la clase. A partir del instruido aprendiz de especialista, se observa una relación inversa entre dos grandes bloques, los de mayor y los de menor competencia cognitivo-comunicativa, que puede observarse en que las zonas marcadas en gris se sitúan por encima y por debajo del eje de abscisas respectivamente. Las líneas de

tendencia en el Gráfico 8 se superponen y permiten identificar una coincidencia de la distribución de los emisores y los destinatarios en los temas del corpus. Sin embargo, en el Gráfico 26 se observa una tendencia inversa generalizada, con solo algunos casos de coincidencia.

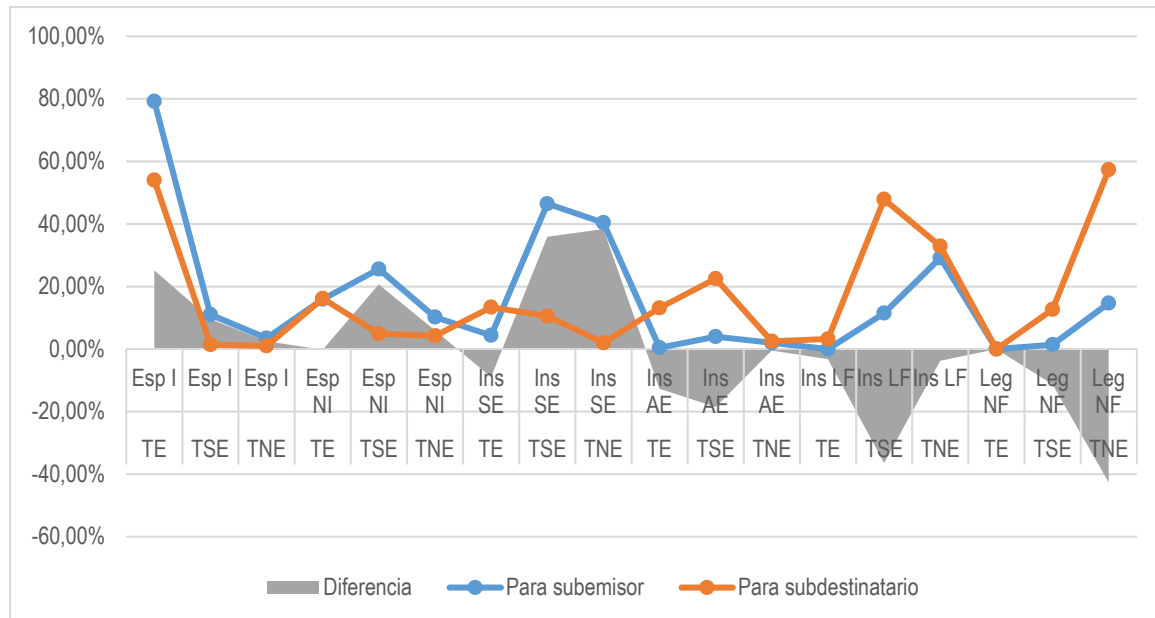


Gráfico 26: Superposición de datos porcentuales para subemisor y subdestinatario según clases

Las diferencias más destacables se relacionan en la Tabla 91, donde se muestran los valores absolutos de las diferencias porcentuales que existen entre subemisor y subdestinatario. Los porcentajes más cercanos al subemisor se marcan en azul y los que se aproximan más al subdestinatario se marcan en rojo.

Usuario	Clase	Porcentaje para subemisor	Porcentaje para subdestinatario	Diferencia
Especialista no investigador	TSE	25,59 %	4,93 %	20,66 %
Instruido semiespecialista	TSE	46,48 %	10,56 %	35,92 %
Instruido semiespecialista	TNE	40,40 %	2,00 %	38,40 %
Instruido aprendiz de especialista	TE	0,46 %	13,16 %	12,70 %
Instruido aprendiz de especialista	TSE	3,99 %	22,54 %	8,55 %
Instruido lego formado	TSE	11,50 %	47,89 %	36,39 %
Lego no formado	TSE	1,41 %	12,68 %	11,27 %
Lego no formado	TNE	14,71 %	57,36 %	42,65 %

Tabla 91: Diferencias porcentuales para subemisor y subdestinatario según tipo de usuario y clase

A partir de estos datos, puede inferirse que los informantes presuponen cierto conocimiento sobre el tema que exponen los emisores de los textos, mientras que dicho conocimiento queda alejado del destinatario de los textos. Esto se vincula estrechamente con la relación entre procesos de producción y de recepción propuestos por Bernárdez

(1995: 139-141) (*vid. supra* II. 2.1., p. 37) cuando menciona las expectativas del productor acerca del conocimiento de las estrategias comunicativas por el destinatario.

1.2.4. Relación SUBEMISOR-SUBDESTINATARIO

La vinculación del subemisor con cada subdestinatario según la clase permite caracterizar la situación discursiva de los textos, de forma que sea posible aproximarse a la predicción de las clases según el grado de especialización a partir del conocimiento del perfil cognitivo-comunicativo del emisor y del destinatario.

En el caso del especialista (*vid. infra* Gráfico 27), es clara la relación especialista investigador-especialista investigador para el texto especializado. En efecto, ninguna otra relación es tan frecuente en dicho gráfico. Los textos vinculados a esta relación son el artículo en revista científica (107 textos) y la tesis doctoral (106 textos), entre casos muy poco frecuentes de blog (7), entrevista (1), noticia (1) o revista o web de divulgación especializada (7).

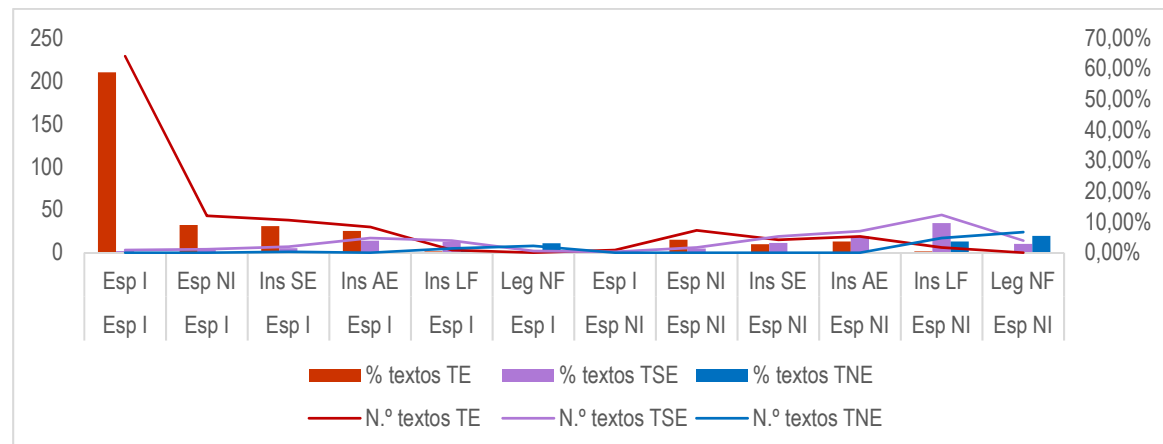


Gráfico 27: Relación de datos globales para la correspondencia emisor especialista-tipo de subdestinatario según clase

El caso del instruido, como en otras ocasiones, es más diverso, principalmente a causa de su estatus de categoría difusa y que cuenta con tres subcategorías. En el Gráfico 28 destacan, para el texto semiespecializado, cinco relaciones y, para el texto no especializado, cuatro relaciones.

En el texto semiespecializado cuentan con elevada ocurrencia las relaciones semiespecialista-aprendiz de especialista¹⁵⁵ y semiespecialista-lego formado^{156, 157}. En el texto no especializado, la relación semiespecialista-lego no formado¹⁵⁸ y lego formado-lego formado son las más frecuentes¹⁵⁹.

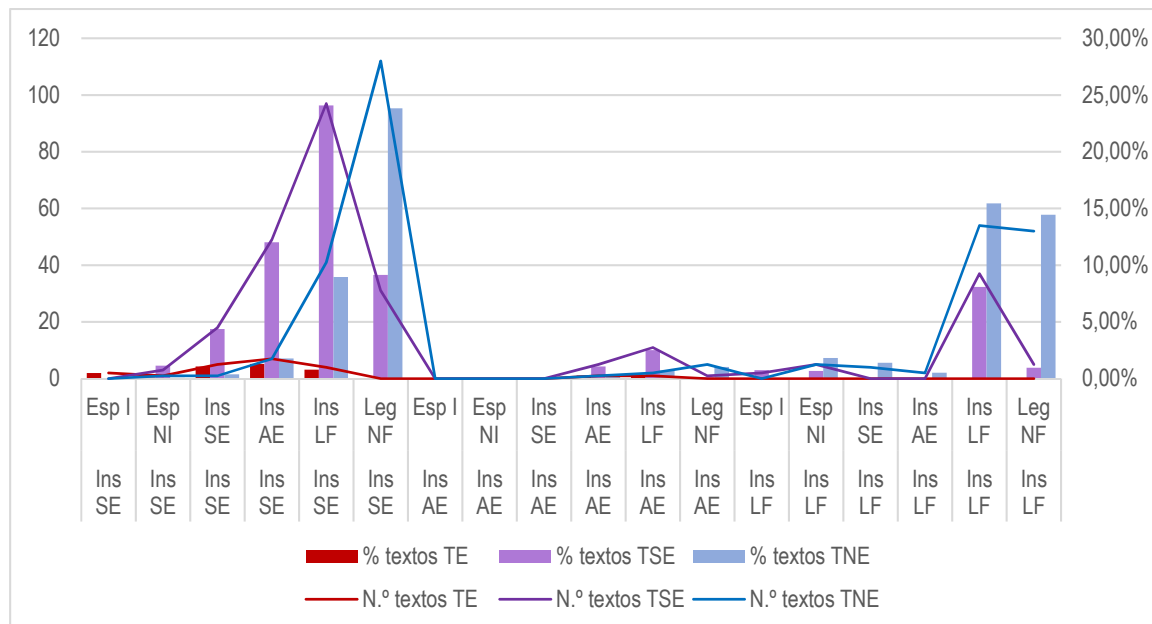


Gráfico 28: Relación de datos globales para la correspondencia emisor instruido-tipo de subdestinatario según clase

Por último, el emisor lego no formado se vincula estrechamente con el destinatario lego no formado en el texto no especializado (*vid. infra* Gráfico 29). En este caso, el único tipo textual para esta combinación es la participación en foro, con 29 textos.

¹⁵⁵ Los tipos textuales más frecuentes son la entrevista y la revista o web de divulgación especializada, con 14 y 13 textos respectivamente. Otros casos para la relación semiespecialista-aprendiz de especialista en el texto semiespecializado son el blog (8), la noticia (9), la participación en foro (1) y la tesis doctoral (4).

¹⁵⁶ Los tipos textuales más frecuentes son la entrevista (26), la noticia (26) y la revista o web de divulgación especializada (26). Otros casos para la relación semiespecialista-lego formado en el texto semiespecializado son el blog (15), la participación en foro (3) o la tesis doctoral (1).

¹⁵⁷ En el caso del emisor semiespecialista en el texto semiespecializado, se vincula estrictamente con aquellos destinatarios que no sean especialistas.

¹⁵⁸ Los tipos textuales más frecuentes en dicha relación son la noticia (47), la entrevista (22), la revista o web de divulgación especializada (19), el blog (18) y la participación en foro (6).

¹⁵⁹ El tipo textual más frecuente para el lego formado-lego formado en el texto no especializado es la participación en foro (38). Otros tipos frecuentes son la noticia (7), la revista o web de divulgación especializada (5) o la entrevista (4).

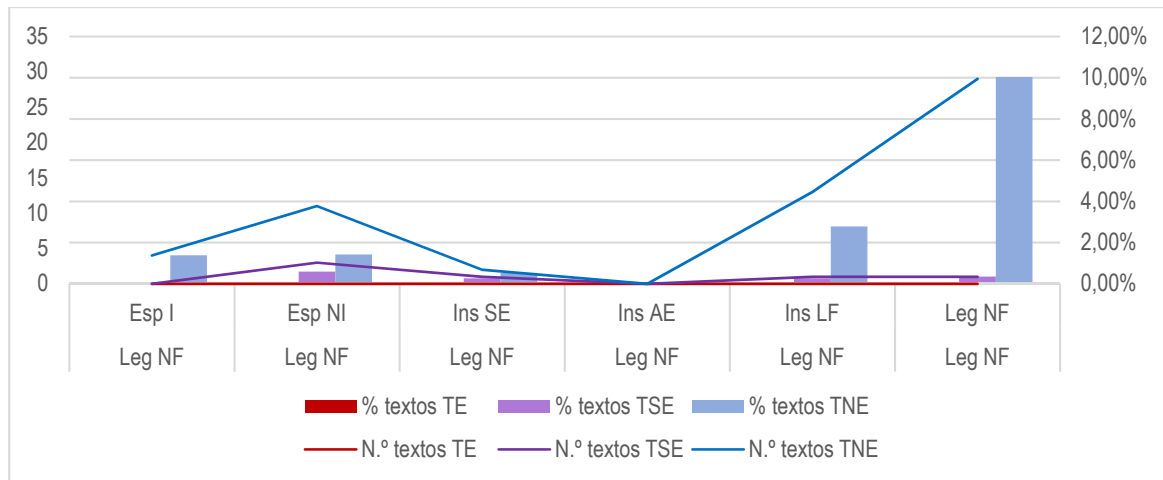


Gráfico 29: Relación de datos globales para la correspondencia emisor lego-tipo de subdestinatario según clase

En resumen, la Tabla 92 sintetiza las relaciones subemisor-subdestinatario características de cada clase según el grado de especialización. Las relaciones más frecuentes quedan acompañadas del tipo textual más recurrente para dicha relación.

Clase	Relación subemisor-subdestinatario	Tipo textual más frecuente
Texto especializado	Investigador-investigador	artículo en revista científica (107) tesis doctoral (106)
	Semiespecialista-semiespecialista	
Texto semiespecializado	Semiespecialista-aprendiz de especialista	entrevista (14) revista o web de divulgación especializada (13)
	Semiespecialista-lego formado	entrevista (26) noticia (26) revista o web de divulgación especializada (26)
	Semiespecialista-lego no formado	
	Semiespecialista-lego formado	
Texto no especializado	Semiespecialista-lego no formado	
	Lego formado-lego formado	
	Lego formado-lego no formado	
	Lego no formado-lego no formado	participación en foro (29)

Tabla 92: Correspondencias más frecuentes entre subemisor, subdestinatario y tipos textuales según clases

De forma esquemática y algo más gráfica, la Tabla 93 muestra las relaciones de la Tabla 92. En ella, se puede comprobar cómo el texto especializado queda restringido a la relación según la competencia cognitivo-comunicativa más elevada, mientras que el texto no especializado contrasta con esta relación y se sitúa en el último punto. Además, destaca la importancia del subemisor semiespecialista tanto para el texto semiespecializado como para el texto no especializado, así como la relevancia del subdestinatario lego formado y lego no formado para el texto no especializado. Por último, cabe señalar la similitud que es

posible identificar entre el texto semiespecializado y el texto no especializado en alguna de las relaciones, así como en ciertos tipos textuales que seleccionan.

Subemisor	Subdestinatario					
	Investigador	No investigador	Semiespecialista	Aprendiz de especialista	Lego formado	Lego no formado
Investigador	TE					
No investigador						
Semiespecialista			TSE	TSE	TSE / TNE	TSE / TNE
Aprendiz de especialista						
Lego formado					TSE / TNE	TNE
Lego no formado						TNE

Tabla 93: Correspondencias más frecuentes entre subemisor y subdestinatario según clases

1.2.5. Relación con las FORMAS LÉXICAS ESPECIALIZADAS

La terminología queda estrechamente vinculada al texto especializado (*vid. infra* Gráfico 30), como era previsible, ya que se trata del objeto comunicativo diseñado para la transmisión de conocimiento especializado. El texto especializado cuenta con el 70,28 % de la terminología del corpus, seguido del texto semiespecializado (19,03 %) y del texto no especializado (10,70 %). En este sentido, el grado de especialización textual coincide con el grado de frecuencia de formas léxicas especializadas en el corpus. La distribución de los tipos de términos, además, es idéntica en los tres casos.

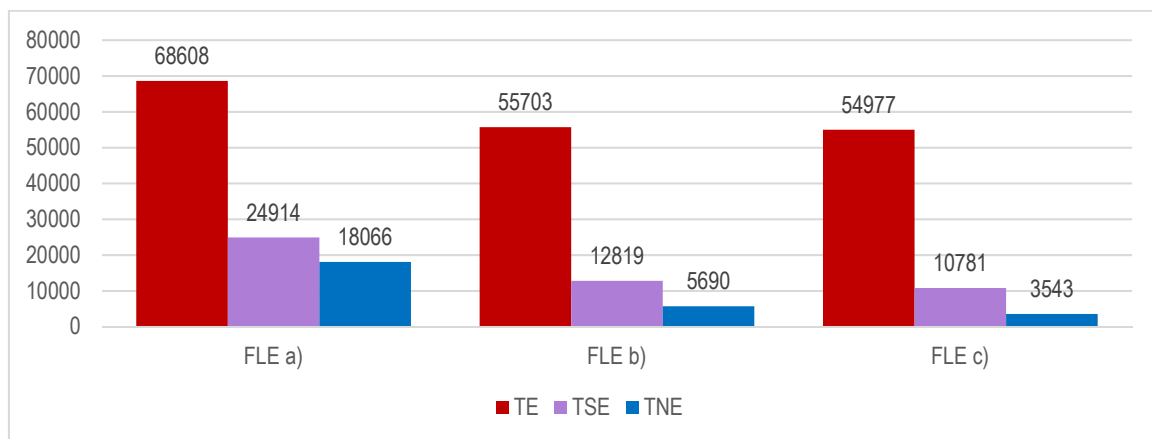


Gráfico 30: Datos globales de tipos de FLE según clase

A continuación, el Gráfico 31, el Gráfico 32 y el Gráfico 33 presentan los datos que permiten describir los tipos de FLE en las clases textuales según los temas. El dato que más destaca en estos tres gráficos es la estrecha relación que existe entre los distintos tipos de

unidades terminológicas y la clase de texto especializado¹⁶⁰. Además, las diferencias entre temas no son muy relevantes, a excepción de algunos casos que se comentarán a continuación. La presencia de terminología en el texto semiespecializado y no especializado es similar tanto cuantitativamente como en los diferentes temas aunque existen algunas leves diferencias, que son las que se comentarán a continuación.

En el caso de las formas délficas (*vid. infra* Gráfico 31), el derecho también cuenta con mayor cantidad de unidades tanto en el texto semiespecializado como en el no especializado, si bien las diferencias con los otros temas son muy estrechas. En comparación con los otros temas, el derecho duplica la cantidad de términos délficos de la informática y supera en un 60 % los términos délficos de la medicina. Esto puede sugerir que la disciplina jurídica del derecho se construye empleando unidades léxicas que parten del discurso común y que adquieren un significado actualizado al campo científico.

Además, en el caso de la informática y la medicina, existe menos cantidad de términos délficos en el texto no especializado que en el semiespecializado aunque es posible hallar una frecuencia similar de FLE a) tanto en el texto semiespecializado como en el texto no especializado, lo que no ocurre con los otros tipos de FLE.

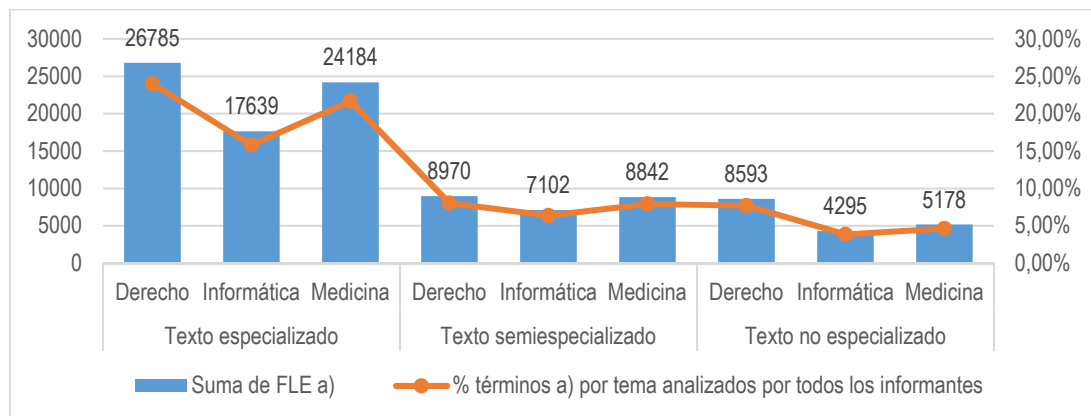


Gráfico 31: Suma de FLE a) según tema y clase de texto

En relación con los términos semicrípticos (*vid. infra* Gráfico 32), como puede comprobarse, la presencia de estas formas es similar en los tres temas en las tres clases de texto, aunque, diferenciando entre clases se observa un crecimiento lineal de texto no especializado a texto semiespecializado y un crecimiento exponencial de texto semiespecializado a texto especializado. En este caso, se observa que todas las disciplinas

¹⁶⁰ Tanto es así que la curva de distribución numérica del total de FLE en el corpus (*vid. supra* Gráfico 20) coincide con la curva de distribución numérica en el texto especializado en el Gráfico 31, el Gráfico 32 y el Gráfico 33).

cuentan con textos que poseen terminología reconocible como poco frecuente en el discurso común.

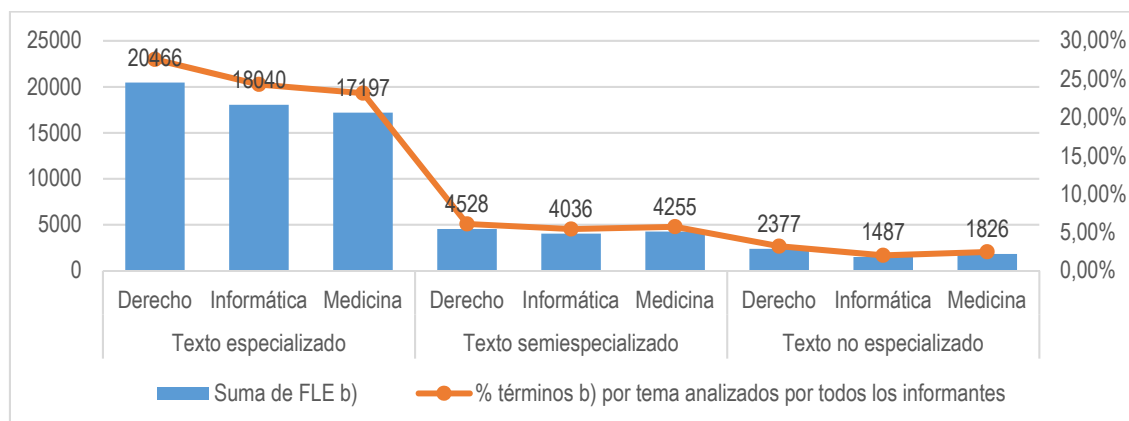


Gráfico 32: Suma de FLE b) según tema y clase de texto

Por último, los términos crípticos, los menos frecuentes del corpus, se distribuyen de forma diferenciada entre texto semiespecializado y texto no especializado. En cuanto al primero, la informática destaca por encima del derecho y la medicina, casi duplicando los términos crípticos en cada una de ellas. En relación con el segundo, las frecuencias son muy similares, si bien la medicina duplica el número de términos crípticos del derecho en el texto no especializado.

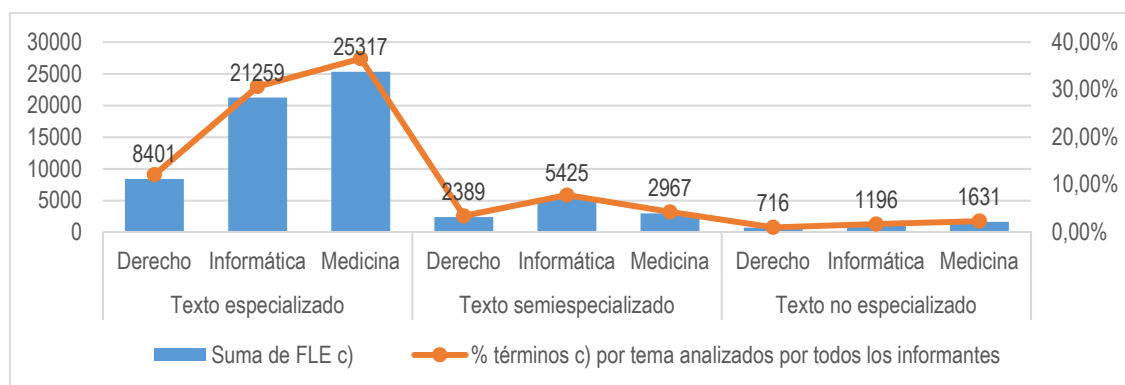


Gráfico 33: Suma de FLE c) según tema y clase de texto

1.2.6. Relación con los PROCEDIMIENTOS DE REFORMULACIÓN FORMAL

De forma similar a la terminología, los procedimientos de reformulación priman en el texto especializado (65,61 % del total del corpus), al que le sigue el texto semiespecializado (22,88 %) y el texto no especializado (11,51 %). En este caso, los procedimientos de reformulación aumentan cuanto mayor es el grado de especialización (al igual que el número de formas léxicas especializadas, *vid. supra* IV.1.2.5., p. 284). La

distribución también coincide, si bien los procedimientos de expansión son los más numerosos con diferencia (59 % del total del corpus).

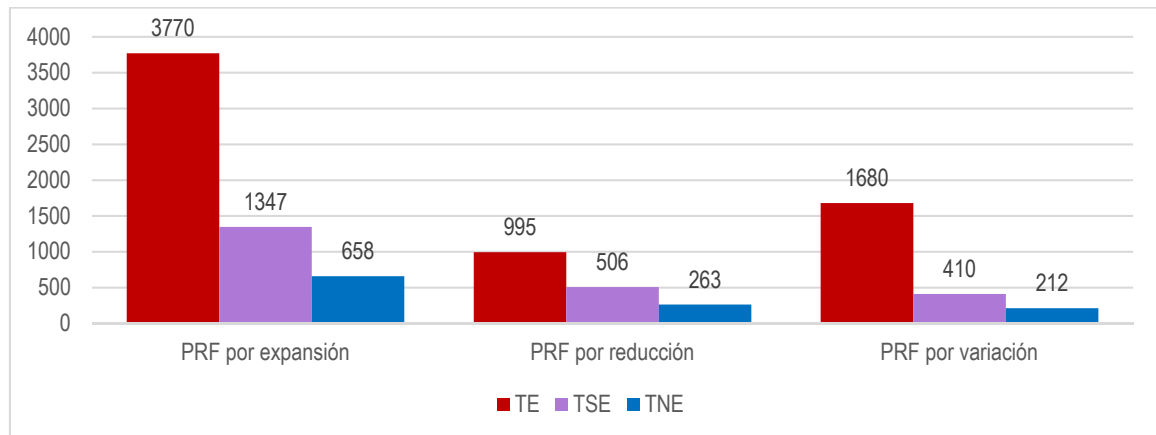


Gráfico 34: Datos globales de tipos de PRF según clase

A continuación, el Gráfico 35, el Gráfico 36 y el Gráfico 37 presentan los datos que permiten describir los tipos de PRF en las clases textuales según los temas. A partir de estos gráficos, es posible hallar una relación entre el texto especializado y los procedimientos de reformulación, especialmente en los procedimientos por expansión y por variación (los procedimientos por reducción, si bien siguen siendo más numerosos en el texto especializado, se reparten de forma más homogénea entre el resto de las clases).

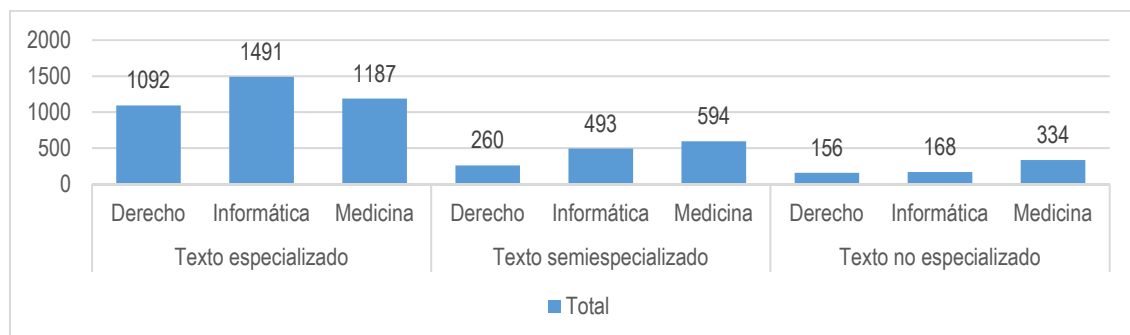


Gráfico 35: Suma de PRF a) según tema y clase de texto

En el caso de los procedimientos por reducción (*vid. infra* Gráfico 36), destaca la informática muy por encima de la medicina y el derecho, lo que puede justificarse mediante la gran cantidad de siglas, acrónimos y denominaciones que se registran en esta clase textual (y en general en este tema), además de que, como se mencionaba en IV. 1.1.1. (p. 244) el texto semiespecializado de temática informática es el que más destaca dentro de todo el corpus.

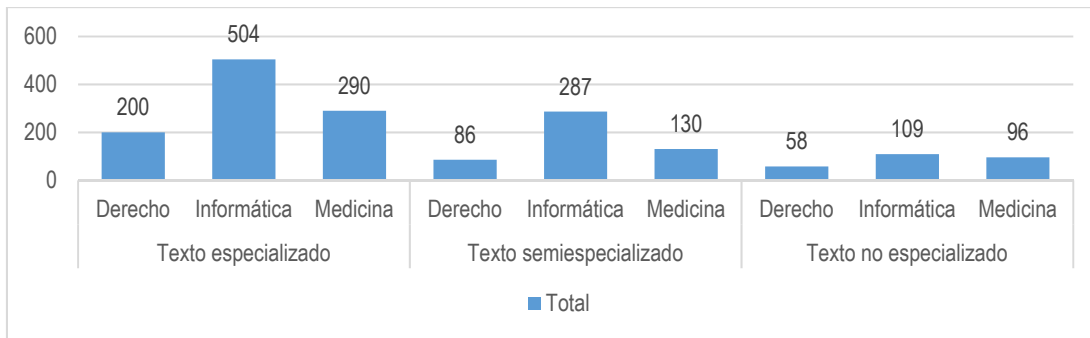


Gráfico 36: Suma de PRF b) según tema y clase de texto

En el Gráfico 37, el texto semiespecializado y no especializado cuenta con un número muy reducido de procedimientos de variación (en el corpus, principalmente se hallan denominaciones alternativas, préstamos o acrónimos).

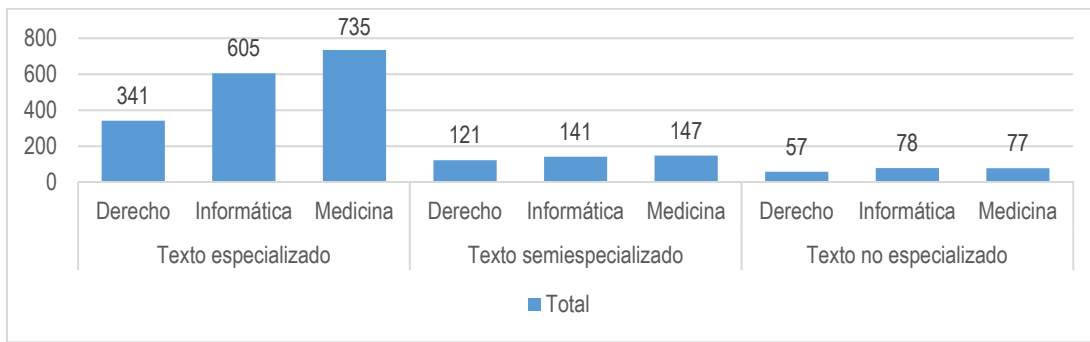


Gráfico 37: Suma de PRF c) según tema y clase de texto

1.2.7. Conclusiones parciales sobre la CLASE

A continuación, incluimos los puntos más importantes que sintetizan la información que relaciona las clases del corpus con el resto de variables.

- a) Sobre los tipos de texto, se identifican dos grandes bloques: el de texto especializado y el conjunto texto semiespecializado-texto no especializado. En el primer bloque los textos más frecuentes son el artículo en revista científica y la tesis doctoral y en el segundo bloque el resto de los tipos textuales, que se reparten de forma inversamente proporcional entre el texto semiespecializado y el texto no especializado.
- b) En cuanto a las relaciones discursivas, es clara la relación especialista investigador-especialista investigador en el texto especializado. El caso del instruido se distribuye de forma más dispar, si bien el subemisor semiespecialista es predominante para el texto semiespecializado y no especializado. En concreto, para el texto semiespecializado, el emisor semiespecialista se combina con todos

los usuarios que no son especialista, especialmente con el aprendiz de especialista y el lego formado, con los cuales la entrevista o la revista o web de divulgación especializada está muy relacionada. El texto no especializado comparte algunas relaciones con el texto semiespecializado, como las de semiespecialista-lego formado y semiespecialista lego no formado. No obstante, la relación más característica para esta clase es la de lego no formado-lego no formado, cuyo tipo textual más usual es la participación en foro.

- c) Respecto a las formas léxicas especializadas, el texto especializado cuenta con el 70 % de la terminología de todo el corpus (el semiespecializado cuenta con el 19 % y el no especializado con el 10 %). El número de términos aumenta conforme aumenta el grado de especialización textual. Las diferencias de los tipos de términos entre los temas y las clases son prácticamente inexistentes: tan solo se observa una diferencia notable entre la medicina y el derecho los términos crípticos del texto especializado, ya que la primera cuenta con tres veces más términos que la segunda.
- d) En lo que concierne a los procedimientos de reformulación formal, se asemejan al comportamiento de la terminología: el texto especializado cuenta con la mayoría de los procedimientos de reformulación (66 %) frente al semiespecializado (23 %) o el no especializado (11 %). El número de procedimientos de reformulación formal aumenta conforme aumenta el grado de especialización textual. Por otro lado, los procedimientos por expansión y por variación se vinculan más estrechamente con el texto especializado, mientras que los procedimientos por reducción tienen una ocurrencia más homogénea en todos los textos, si bien el texto especializado es el que prima.

1.3. Análisis del TIPO DE TEXTO

Como punto de partida, debe recordarse que los 7 tipos textuales cuentan con una representación de muestras equilibrada, ya que cada uno está compuesto por 15 textos.

1.3.1. Relación con el EMISOR y el SUBEMISOR

El Gráfico 38 se asemeja a los datos proporcionados por la relación entre clase y tipo de texto (*vid. supra* Gráfico 23, apartado IV. 1.2.1., p. 273). Mientras que el artículo en revista científica y la tesis doctoral (los dos textos más vinculados al texto especializado) están estrechamente relacionados con el emisor especialista investigador, el resto de los tipos textuales se distribuye de forma más o menos homogénea entre el resto de emisores (a excepción del instruido aprendiz de especialista, que es el menos frecuente). El hecho de que el blog y la revista o web de divulgación especializada también esté dentro de la categoría de especialista (en concreto, dentro del especialista no investigador) concuerda con los datos analizados en la Tabla 87 (*vid. supra* apartado IV. 1.2.1., p. 275), en los que se defendía que la diferencia con los tipos textuales especializados característicos se hallaba en su menor densidad terminológica y su relación con el destinatario.

Como en análisis previos, los datos globales de subemisor semiespecialista se vinculan al blog, la entrevista, la noticia o la revista o web de divulgación especializada y el instruido lego formado y el lego no formado se relacionan con la participación en foro (*vid. Gráfico 38*).

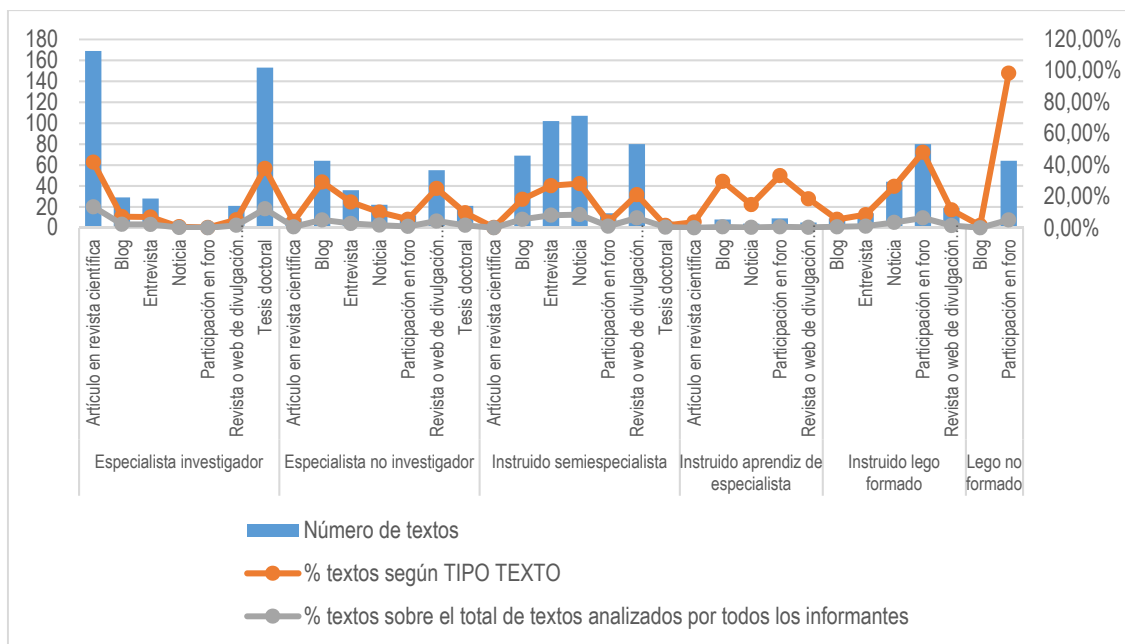


Gráfico 38: Datos globales de subemisor según tipo de texto

1.3.2. Relación con el DESTINATARIO y el SUBDESTINATARIO

En el caso del destinatario se observan tres bloques (*vid. infra* Gráfico 39): (i) el primero, dedicado al artículo en revista científica y tesis doctoral con destinatario

especialista investigador; (ii) el segundo, vinculado al instruido lego formado y al lego no formado, con frecuencias para todos los tipos textuales de en torno al 20 %; y (iii) el tercero y último, en el que se incluyen el resto de destinatarios, con frecuencias muy escasas en el corpus. De acuerdo con estos datos, en relación con el destinatario de un texto, es posible afirmar que los informantes optan por vincular al texto aquel destinatario que está más alejado del centro del grado según la competencia cognitivo-comunicativa del usuario. Como se observa, el destinatario especialista investigador es exclusivo de dos tipos textuales y los destinatarios que menos competencia cognitiva y comunicativa tienen sobre un sistema conceptual (lego formado y lego no formado) se vinculan al resto de tipos textuales. De hecho, en el caso del lego formado se observan dos porcentajes muy reducidos, correspondientes con el artículo en revista científica y con la tesis doctoral. Estos dos tipos textuales ya no aparecen como destinatarios en el lego no formado.

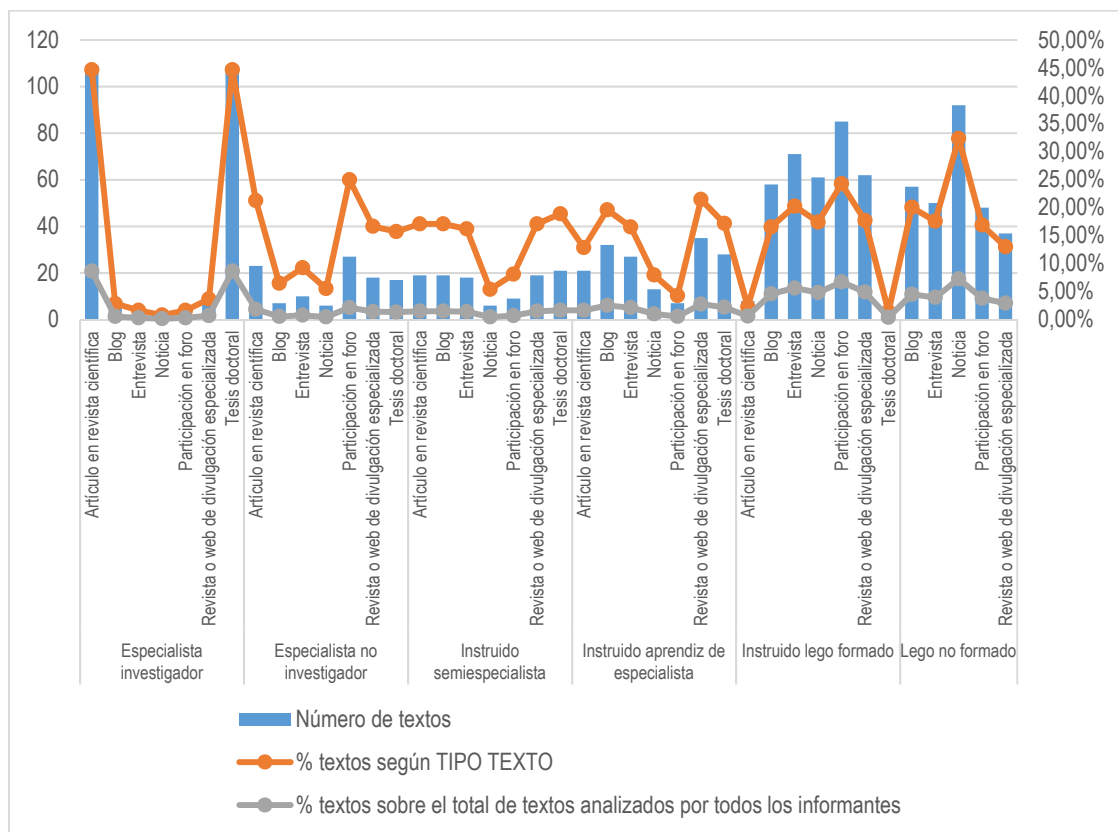


Gráfico 39: Datos globales de subdestinatario según tipo de texto

La Tabla 94 relaciona las características para destinatario instruido lego formado y lego no formado para los tipos textuales comunes a dichos destinatarios en el Gráfico 39. Se ha usado el color rojo para marcar las diferencias entre pares y el símbolo (↓) o (↑) para indicar que se trata de un valor menor o mayor que su análogo.

DESTINATARIO INSTRUIDO LEGO FORMADO				DESTINATARIO LEGO NO FORMADO		
BLOG						
Semiespecialista y especialista no investigador.			Subemisor	Semiespecialista y especialista no investigador.		
Semiespecializado y no especializado.			Clase	No especializado y semiespecializado.		
Mínimo	Promedio	Máximo		Mínimo	Promedio	Máximo
57,00	67,163	82,93	Variedad léxica	57,00	67,736	82,93
0,044	0,208	0,502	IDT	0,096	0,233 (↑)	0,490
0	18,824	181	IR	25,170	18,789	78,545
ENTREVISTA						
Semiespecialista y especialista no investigador. Existen casos de especialista investigador y lego formado.			Subemisor	Semiespecialista y especialista no investigador. Existen casos de especialista investigador y lego formado.		
Semiespecializado y no especializado.			Clase	No especializado y semiespecializado.		
Mínimo	Promedio	Máximo		Mínimo	Promedio	Máximo
57,44	68,159	83,28	Variedad léxica	63,60	69,056	78,57
0,014	0,138	0,361	IDT	0,013	0,125 (↓)	0,289
0	82,502	512,057	IR	0	47,551 (↓)	510,955
NOTICIA						
Semiespecialista, lego formado y especialista no investigador.			Subemisor	Lego formado y semiespecialista.		
Semiespecializado y no especializado.			Clase	No especializado y semiespecializado.		
Mínimo	Promedio	Máximo		Mínimo	Promedio	Máximo
64,60	71,161	80,76	Variedad léxica	64,57	72,578	82,63
0,048	0,200	0,406	IDT	0,008	0,152 (↓)	0,392
0	19,293	83,500	IR	0	16,542	81,594
PARTICIPACIÓN EN FORO						
Lego formado y lego no formado. Hay casos de semiespecialista, especialista no investigador y aprendiz de especialista.			Subemisor	Lego no formado, lego formado y semiespecialista. Hay casos de especialista no investigador.		
Semiespecializado y no especializado.			Clase	No especializado.		
Mínimo	Promedio	Máximo		Mínimo	Promedio	Máximo
52,76	65,417	80,09	Variedad léxica	56,80	68,094	74,60
0,023	0,162	0,358	IDT	0,017	0,127 (↓)	0,350
0	23,570	132,632	IR	0	32,706 (↑)	140
REVISTA O WEB DE DIVULGACIÓN ESPECIALIZADA						
Semiespecialista, lego formado y especialista investigador.			Subemisor	Semiespecialista y lego formado.		
Semiespecializado y no especializado.			Clase	No especializado y semiespecializado.		
Mínimo	Promedio	Máximo		Mínimo	Promedio	Máximo
54,92	67,343	78,43	Variedad léxica	60,87	70,344	78,43
0,052	0,215	0,506	IDT	0,023	0,150 (↓)	0,442
0	18,836	135,625	IR	0	17,641	44,289

Tabla 94: Diferencias según subdestinatario instruido lego formado y lego no formado para blog, entrevista, noticia, participación en foro y revista o web de divulgación especializada

Al contrastar los dos grupos se observa que los subemisores casi siempre coinciden, aunque a veces en menor medida. En cuanto a la clase, todos comparten la clasificación de

texto semiespecializado y no especializado, aunque la distribución es inversa en el destinatario lego no formado con respecto al destinatario lego formado. La variedad léxica y la reformulación tampoco parecen ser relevantes en la discriminación de los dos grupos. Por otro lado, debe añadirse que es un factor común a todos los tipos textuales que la densidad terminológica sea menor en los textos asociados al destinatario lego no formado (a excepción del blog).

Dado que la diferencia principal parece hallarse en la densidad terminológica, en la Tabla 95 se relacionan el tipo de IDT y el tipo de IR según instruido lego formado y lego no formado para el blog, la entrevista, la noticia, la participación en foro y la revista o web de divulgación especializada.

		BLOG	ENT	NOT	FOR	DIV
IDT a)	LF	0,087	0,078	0,092	0,092	0,092
	LNF	0,127	0,105	0,096	0,080	0,104
IDT b)	LF	0,061	0,037	0,059	0,036	0,066
	LNF	0,063	0,023	0,051	0,024	0,044
IDT c)	LF	0,058	0,026	0,047	0,034	0,052
	LNF	0,037	0,008	0,018	0,023	0,032
IR a)	LF	19,723	27,221	11,238	15,823	16,418
	LNF	12,608	36,857	8,774	26,104	10,663
IR b)	LF	2,955	37,825	2,102	4,743	5,494
	LNF	6,459	2,383	4,087	1,783	3,811
IR c)	LF	6,309	7,434	7,529	3,004	4,900
	LNF	2,797	10,891	2,980	4,819	4,736

Tabla 95: Diferencias de tipo de IDT y tipo de IR según subdestinatario instruido lego formado (LF) y lego no formado (LNF) para blog, entrevista, noticia, participación en foro y revista o web de divulgación especializada

En dicha tabla se encuentran valores diversos y diferentes para cada tipo, a excepción de los valores marcados en rojo para el IDT c), es decir, las unidades poco frecuentes en el discurso común y cuyo conocimiento de forma o contenido no es compartido por los informantes (términos crípticos). Los valores para el IDT de términos crípticos son comunes, es decir, en el lego no formado existe menor densidad terminológica de estas unidades, lo cual concuerda con el hecho de que este usuario no cuenta con la competencia cognitivo-comunicativa necesaria para comprender las formas léxicas especializadas más opacas.

1.3.3. Relación SUBEMISOR-SUBDESTINATARIO

Las relaciones analizadas previamente en varios apartados también se comprueban al cruzar los datos de subemisor, subdestinatario y tipo de texto (*vid. infra* Gráfico 40 a

Gráfico 46). En el caso de artículo en revista científica (Gráfico 40) y tesis doctoral (Gráfico 46), las semejanzas son claras y favorables hacia el especialista investigador, siendo el resto de casos poco diferenciados entre emisor y destinatario si bien existen casos más frecuentes para subdestinatario que para subemisor (debe recordarse la prominencia del subemisor especialista investigador en estos dos tipos textuales).

En otro orden de cosas, la participación en foro (Gráfico 44) es la que presenta menos diferencias entre emisor y destinatario. Las que existen divergen del resto de los tipos textuales, puesto que la prominencia del subdestinatario se encuentra en el especialista no investigador y la prominencia del subemisor se encuentra en el lego no formado. A nuestro juicio, esto se justifica por el tipo de texto, ya que en la participación en foro los participantes son muy diversos y las consultas suelen realizarlas usuarios que desconocen un tema y reclaman el conocimiento de un usuario que pueda suplir sus carencias conceptuales.

Las similitudes en el patrón de distribución con mayor frecuencia en emisores especialista investigador a semiespecialista y con mayor frecuencia en destinatarios aprendiz de especialista a lego no formado se repite en el blog (Gráfico 41), la entrevista (Gráfico 42), la noticia (Gráfico 43) o la revista o web de divulgación especializada (Gráfico 45). De hecho, todos ellos comparten el punto en los que se establecen las grandes diferencias entre emisor y destinatario: a partir del instruido aprendiz de especialista las diferencias entre subemisor y subdestinatario se inclinan hacia el subdestinatario, lo que coincide con lo descrito en el Gráfico 26 y la Tabla 91 (*vid. supra* IV. 1.2.3., p. 279). En ese caso, la distribución inversa para la clase también se iniciaba en el instruido aprendiz de especialista. Además, en el Gráfico 23 y la Tabla 88 (*vid. supra*, p. 273) se vinculaban los tipos textuales del Gráfico 41 a Gráfico 45 al grupo de textos semiespecializados y no especializados. A través de la comparación de estas similitudes se puede encontrar un patrón de correlación entre el tipo textual y la clase textual, que se establece a través de las frecuencias de asignación de subemisor y subdestinatario a cada texto del corpus.

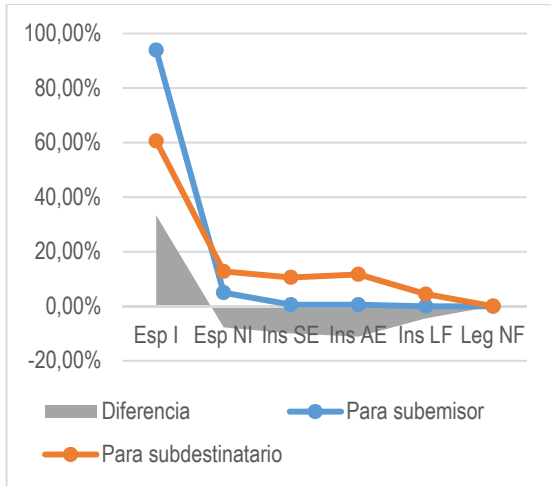


Gráfico 40: Superposición de datos porcentuales para subemisor y subdestinatario (artículo en revista científica)

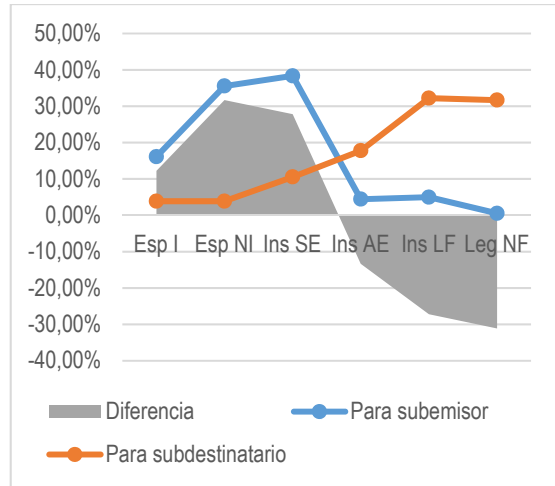


Gráfico 41: Superposición de datos porcentuales para subemisor y subdestinatario (blog)

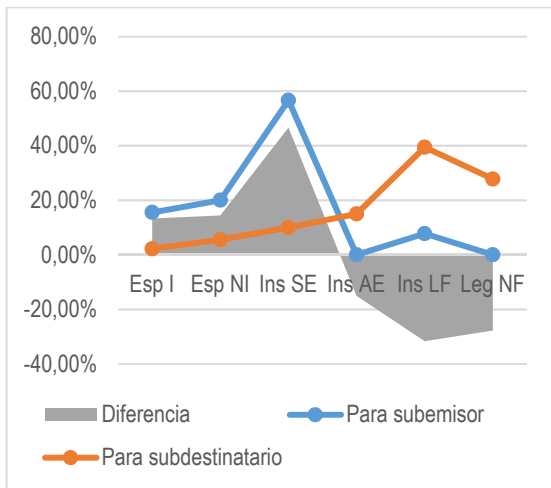


Gráfico 42: Superposición de datos porcentuales para subemisor y subdestinatario (entrevista)

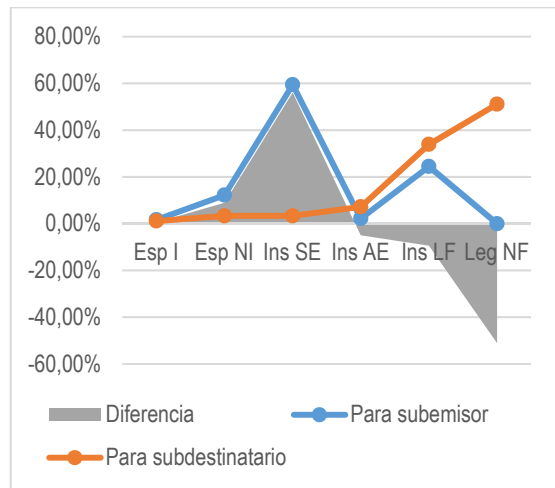


Gráfico 43: Superposición de datos porcentuales para subemisor y subdestinatario (noticia)

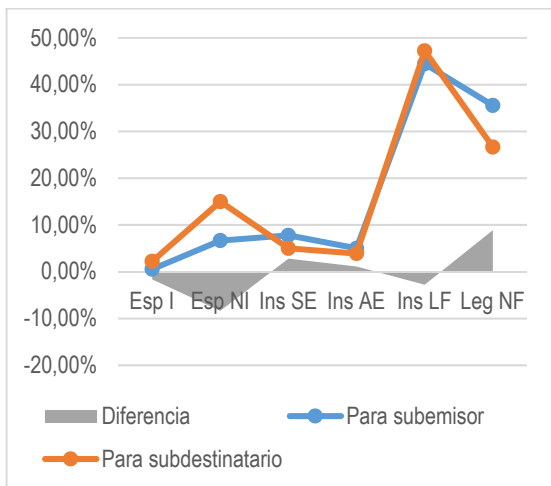


Gráfico 44: Superposición de datos porcentuales para subemisor y subdestinatario (participación en foro)

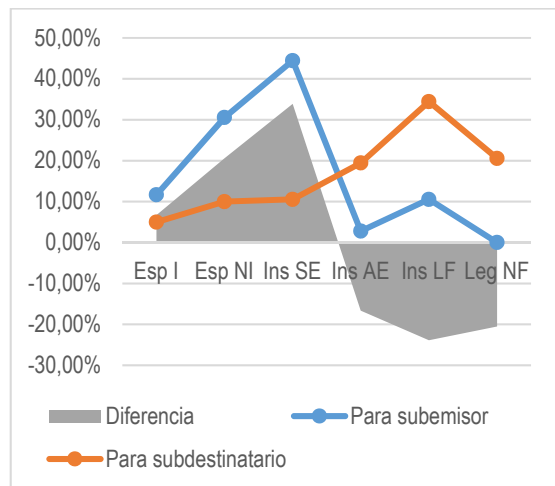


Gráfico 45: Superposición de datos porcentuales para subemisor y subdestinatario (revista o web de divulgación especializada)

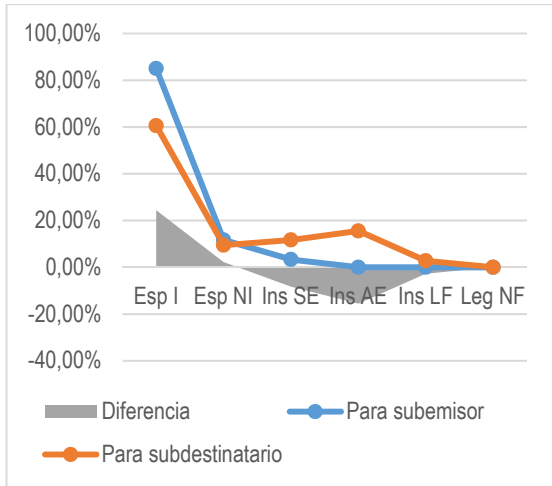


Gráfico 46: Superposición de datos porcentuales para subemisor y subdestinatario (tesis doctoral)

Si relacionamos las posibles combinaciones entre subemisor y subdestinatario por cada tipo de texto según los tres temas del corpus, se identifican algunas diferencias. Los porcentajes que aparecen en cada una de las siguientes tablas corresponden con cada subcorpus según cada tema sobre un valor de 100.

En el caso del Gráfico 47 se observa una clara diferencia entre el derecho y la informática y la medicina para el artículo científico. Mientras que las dos últimas disciplinas alcanzan aproximadamente el 7 % del subcorpus, el caso del derecho no llega a superar el 2 % en la relación especialista investigador-especialista investigador, lo cual puede explicarse a través de la relevancia que tienen otras relaciones para el derecho (*vid. infra* Gráfico 49 y Gráfico 51).

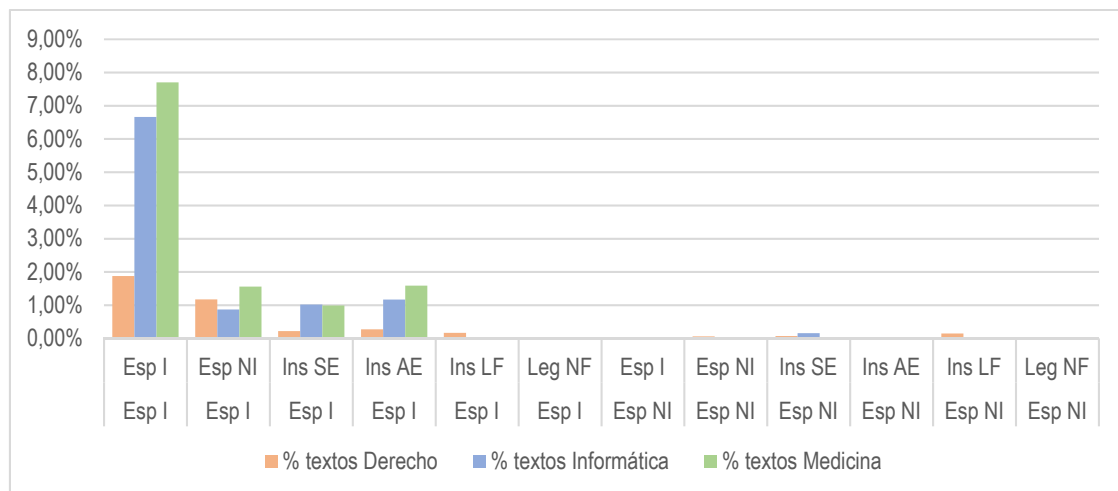


Gráfico 47: Datos comparativos sobre relación subemisor especialista-subdestinatario según temas para el artículo en revista científica

Otros datos que permiten contrastar la relación entre instruido semiespecialista y las dos categorías de destinatario con menor competencia cognitivo-comunicativa

corresponden con la entrevista (*vid. infra* Gráfico 48). Aunque ninguno de los casos alcanza el 5 % del total de cada subcorpus, llama la atención la relación inversa entre el destinatario lego formado y el lego no formado, especialmente en el caso del derecho¹⁶¹.

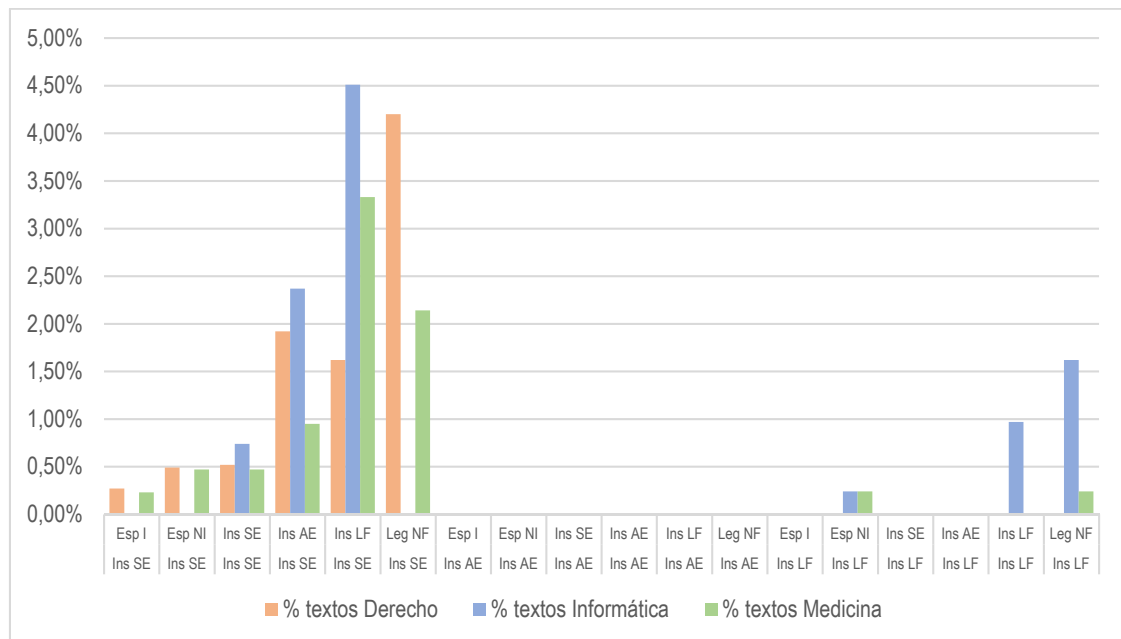


Gráfico 48: Datos comparativos sobre relación subemisor instruido-subdestinatario según temas para la entrevista

Para ilustrar este fenómeno, nos servimos de dos ejemplos (*vid. infra* Tabla 96). En DER ENT07 se observan algunas formas semicrípticas, es decir, el informante tiene un conocimiento laxo del concepto o de la denominación de la forma léxica, junto con una reformulación y una forma délfica. En el segundo caso, el DER ENT06 presenta una densidad terminológica menor, así como una selección de términos délficos mayor que en el caso anterior.

No creo en el **derecho de autodeterminación**

Entonces no me sirves de ejemplo. Mira: aunque votaran el 100% de los catalanes y el 100% de los españoles, eso no serviría de nada. El **derecho de autodeterminación** no existe. Lenin, Stalin, Otto Bauer... Ya Marx, en un artículo en Nueva York, escribió que ninguna de las naciones que descubrieron la libertad política de los pueblos con la Revolución Francesa y que tenían ya su unidad política no tenían **derecho de autodeterminación**; y Marx puso como ejemplo a España. Todo el nacionalismo sin excepción es burgués y de derechas. ¡Y la CUP dice que es de izquierdas! La **autodeterminación es un derecho derivado del derecho de libre determinación, creado para liberar a Grecia del Imperio Otomano**. Luego se transformó en **derecho de autodeterminación** con Otto Bauer, que incorporó la **doctrina** de Marx, de Lenin y de Stalin, que escribió una obra notable: *Nacionalidades*.

(DER ENT07) (Investigador)

Francisco Iniesto: Hay otro problema, sobre la **inseguridad jurídica**, y es que nosotros también tenemos clientes extranjeros, que perciben que en España ocurren cosas que no suceden en el resto de la Unión

¹⁶¹ En la entrevista de derecho con subdestinatario lego formado encontramos 7 textos en el corpus, mientras que en la de subdestinatario lego no formado es posible hallar 18.

Europea. **Detectan una serie de complejidades no razonadas y no fundamentadas.** De cara a los **inversores** extranjeros estamos fallando. No estamos haciendo España atractiva.

Nosotros tenemos que hacer trajes a medida, porque la propia **legislación fiscal** nos obliga a hacerlos. No hay respuestas ni soluciones globales.

José Luis Pacheco: Otra cuestión fundamental es que se regule nuestra actividad. Hoy en día, **asesor se puede llamar a cualquiera**. Por ejemplo, el ex trabajador de un banco que monta su despacho en su casa a bajos precios, que le alcanzan ni para tomar los cursos de formación.

Asesorar sobre la **vida económica** de una persona es tan importante como hacerlo sobre la salud.

(DER ENT06) (Traductor)

Tabla 96: Comparación de entrevista de derecho entre semiespecialista-lego formado (DER ENT07) y semiespecialista-lego no formado (DER ENT06)

Por otra parte, en la noticia (Gráfico 49) destaca, sobre todo, la relación semiespecialista-lego formado en derecho, que constituye casi el 9 % de todas las posibles relaciones del subcorpus de derecho (lo que contrasta con la ausencia del derecho en la comparación anterior del Gráfico 47 para el artículo en revista científica).

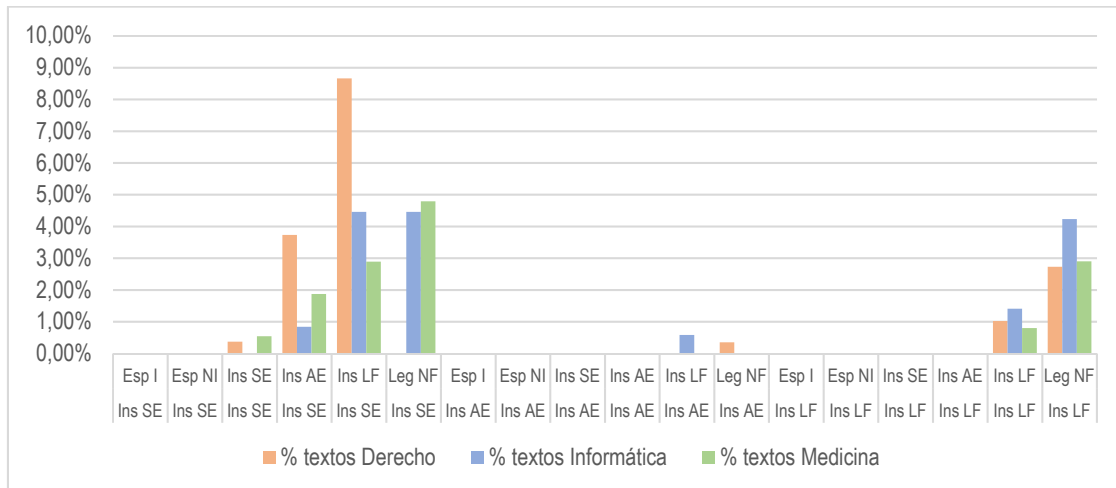


Gráfico 49: Datos comparativos sobre relación subemisor instruido-subdestinatario según temas para la noticia

En cuanto a la informática, la participación en foro alcanza el 10 % del subcorpus de informática en la relación lego formado-lego formado. Además, la informática es el tema que presenta mayor frecuencia para dicha relación (*vid. supra* Gráfico 11) y supone el 14,42 % de todo el corpus. De hecho, la participación en foro es el tipo textual más numeroso en informática para la relación lego formado-lego formado (38 textos), casi quintuplicando el valor del segundo valor más frecuente (la revista o web de divulgación especializada, con 8 textos).

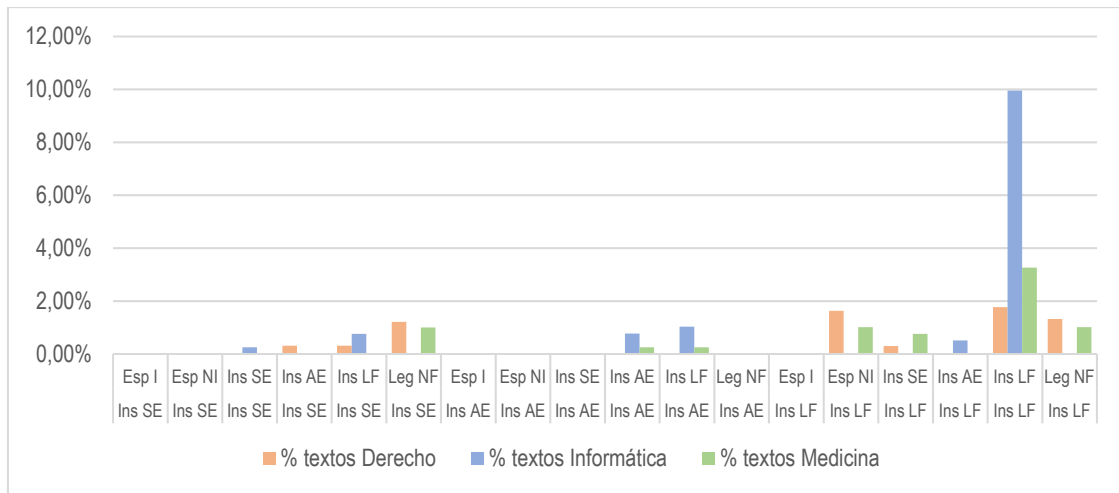


Gráfico 50: Datos comparativos sobre relación subemisor instruido-subdestinatario según temas para la participación en foro

En la Tabla 97 se ilustra una participación en foro sobre informática, en la que se observa la terminología empleada en el texto. Destaca la presencia de terminología poco frecuente en el discurso común.

Encore enhwi-g3, problemas de **microcortes**
 Estimados, experimentos este problema. Alguien que me pueda brindar algo de su sabiduría.
 Las cosas que ya probe:
 1. **Deshabilitar UPnP**
 2. **Actualizar la firmware**, le puse la **version de fabrica** (v2), dado que la version 3 de la **Firmware** no funca (**LINK**). Se puede bajar normalmente pero el aparato la rechaza, lei en varios foros que experimetan el mismo problema, cuando quieren pasar a la Firmware 3.
 3. Por si alguno sugiere que lo **reseteo**, ya lo intente ciento de veces.
 Alguna sugerencia?
 Desde ya, estoy agradecido.
 Sal2
 [...]
 Por suerte creo que encontré una SOLUCION, la cual consiste en **DESHABILITAR EL SERVIDOR DHCP del router** y asignar manualmente las **direcciones IP, MASCARA DE SUBRED, PUERTA DE ENLACE Y SERVIDOR DNS** a cada PC en la red.
 Sí, bastante tedioso si son varias maquinas, y tengan en cuenta de que hay que hacerlo tambien para cada nueva PC que se agregue a la **red**. Por ahora me está dando buenos resultados. Ni una **desconexión** en varias horas de descargas y **msn**.
 ¿Alguien logró actualizar el **firmware**?

(INF FOR10) (Investigador)

Tabla 97: Ejemplo de participación en foro lego formado-lego formado sobre informática

Por último, el caso de especialista investigador-especialista investigador para la tesis doctoral, similar al del artículo en revista científica, supone el 14,62 % del subcorpus en derecho y el 12,86 % del subcorpus en medicina. Aunque la informática se diferencia de los otros dos temas en un porcentaje medio de 5 %, se trata de otra de las correspondencias subemisor-subdestinatario más frecuentes.

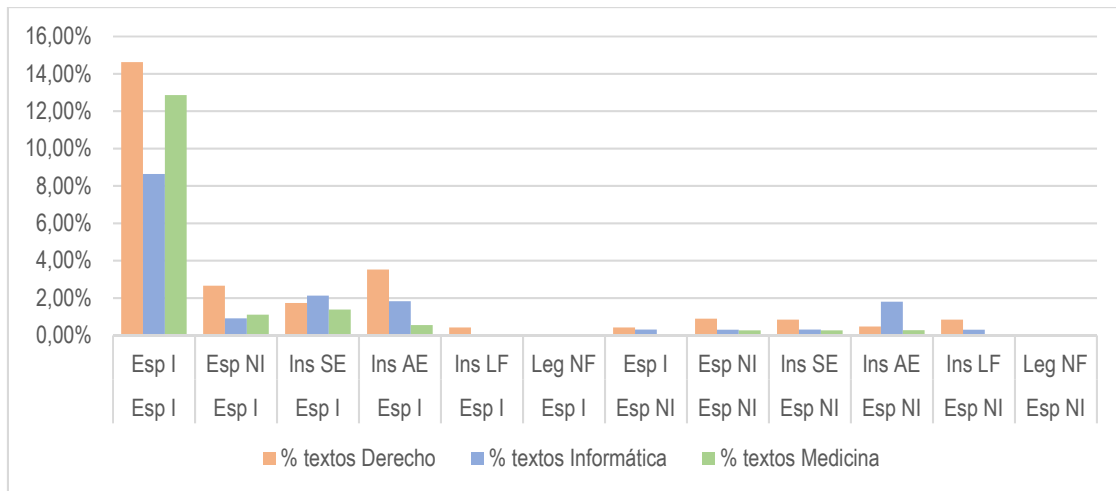


Gráfico 51: Datos comparativos sobre relación subemisor especialista-subdestinatario según temas para la tesis doctoral

1.3.4. Relación con las FORMAS LÉXICAS ESPECIALIZADAS

Este apartado compara las formas léxicas especializadas de cada subcorpus según el tema para los siete tipos textuales del corpus (*vid. infra* Gráfico 52 a Gráfico 58).

En primer lugar, desde el punto de vista del derecho, la distribución de tipos de FLE siempre es la misma en los siete tipos textuales. Las FLE c) son las menos frecuentes (también comparadas con el resto de temas), especialmente en la participación en foro y la entrevista. Esto se relaciona con la selección de las denominaciones del derecho, que parten del discurso común (*dictarse*) o que han sufrido el proceso de banalización (*recurso, normativa o inconstitucionalidad*) (*vid. infra* Tabla 98).

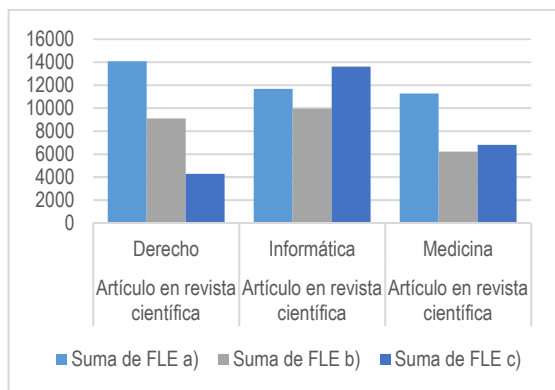


Gráfico 52: Tipos de FLE según temas para el artículo en revista científica

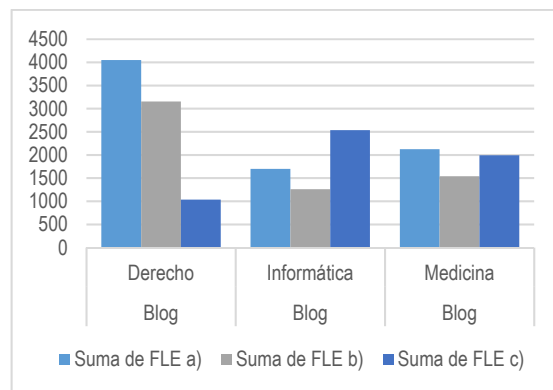


Gráfico 53: Tipos de FLE según temas para el blog

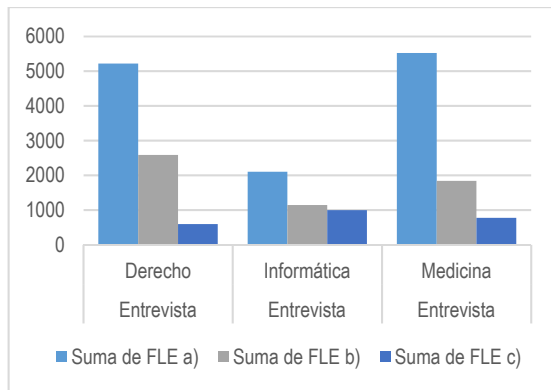


Gráfico 54: Tipos de FLE según temas para la entrevista

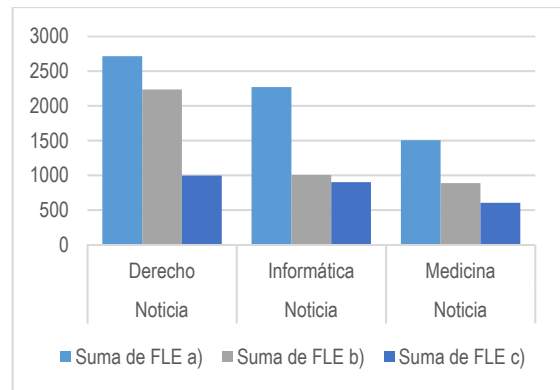


Gráfico 55: Tipos de FLE según temas para la noticia

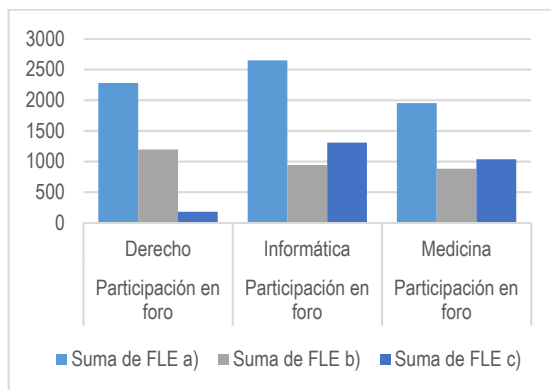


Gráfico 56: Tipos de FLE según temas para la participación en foro

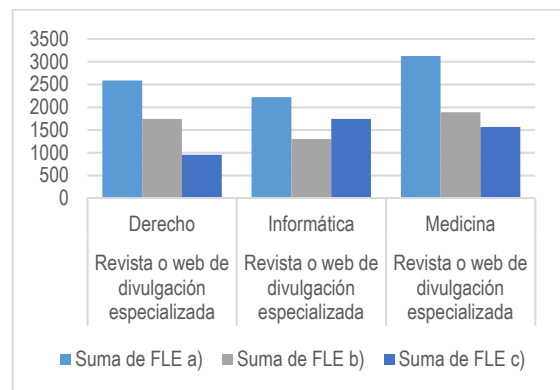


Gráfico 57: Tipos de FLE según temas para la revista o web de divulgación especializada

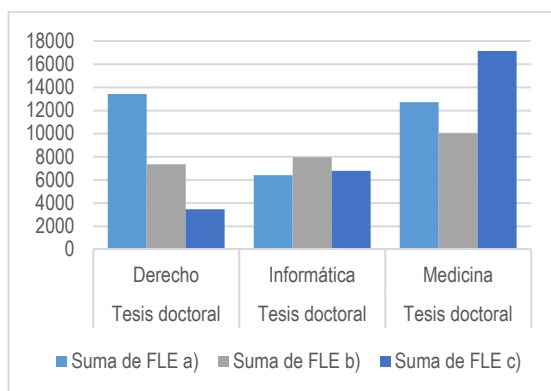


Gráfico 58: Tipos de FLE según temas para la tesis doctoral

El recurso interpuesto se fundamenta en la **disconformidad a derecho del acto administrativo impugnado**, al dictarse en aplicación del **Real Decreto-Ley 8/2014**, de 4 de julio, por el que se crea el **sistema nacional de obligaciones de eficiencia energética**, por considerar que esta norma legal **infringe el Derecho de la Unión Europea y la Constitución Española**. La entidad recurrente solicita la **declaración de nulidad del acto recurrido** como consecuencia de la **inaplicación por el Tribunal de varios preceptos del citado Real Decreto-Ley**, por **infringir preceptos constitucionales y de la normativa de la Unión Europea**. Subsidiariamente se solicita el planteamiento de una **cuestión prejudicial ante el TJUE** y una **cuestión de inconstitucionalidad ante el TC**.

(DER BLOG08) (Traductor)

Tabla 98: Ejemplo de FLE en blog sobre derecho

En el caso de la informática, la distribución es muy desigual. El artículo en revista científica y el blog presentan mayor frecuencia de formas crípticas que de ningún otro tipo de término. Por otra parte, las formas crípticas de la participación en foro y la revista o web de divulgación especializada superan a las formas semicrípticas pero no a los términos más frecuentes en el discurso común según los informantes (términos délficos). En tercer lugar, las formas semicrípticas superan a las formas crípticas en la tesis doctoral y, por último, el caso de la entrevista corresponde con un decrecimiento en el número de términos desde las formas délficas hasta las formas crípticas. Estos datos permiten afirmar que la informática regula la cantidad de terminología (y sus tipos) según los tipos textuales (y, al fin y al cabo, el tipo de destinatario). Uno de los datos que llama la atención es el desequilibrio de FLE entre el artículo en revista científica y la tesis doctoral, no solo ya en la cantidad, sino también en los tipos de términos.

En relación con la medicina, al igual que el derecho, se observa predominancia de formas délficas en todos los tipos textuales, a excepción de la tesis doctoral, donde las formas crípticas son las más numerosas (de hecho, son las más numerosas de todo el subcorpus de tesis doctoral en los tres temas). Por otra parte, las formas semicrípticas son menos frecuentes que las formas crípticas en el artículo en revista científica, el blog, la participación en foro o la tesis doctoral. En contraste, las formas semicrípticas son más frecuentes que las formas crípticas en la entrevista, la noticia o la revista o web de divulgación especializada. En estos casos, no se observa una relación directa entre el tipo de terminología empleada y la clase según el grado de especialización más frecuente, el subemisor o el subdestinatario. Una justificación plausible, aunque no demostrada, podría ser la relación que existe entre el interés personal o profesional de un destinatario y el tipo de configuración discursiva del blog o la revista o web de divulgación especializada, que tienen un público más restringido (y, por tanto, con mayor conocimiento sobre los temas; con ello, sobre la terminología) que la entrevista o la noticia, que pueden ser textos con un abanico más amplio de posibles destinatarios.

Si en lugar de centrarnos en la homogeneidad de la distribución de las FLE en un tema optamos por contrastar los diferentes temas según cada tipo textual, también es posible hallar algunas diferencias.

En el artículo en revista científica (*vid. supra* Gráfico 52) existe un número más o menos equilibrado de FLE a) y b) en los tres temas. Aun así, las FLE c) son mucho más numerosas en la informática (más del doble que en medicina, de hecho). Esto se justifica por el gran número de términos préstamo del inglés y la gran cantidad de términos en forma

acronímica. Si comparamos este tipo textual con el más similar, la tesis doctoral (*vid. supra* Gráfico 58), las diferencias son mucho más notables. En primer lugar, la medicina duplica en número de FLE a la informática (39 868 frente a 21 159 respectivamente). No solo eso, sino que, además, casi triplica el número de formas crípticas de la informática (17 133 frente a 6787), a pesar de que el subcorpus de tesis doctoral de tema médico cuenta con 44 754 palabras y el subcorpus de tesis doctoral de tema informático cuenta con 51 679 palabras. En este caso, puede considerarse que la medicina hace uso de un número mayor de FLE para transmitir el conocimiento que codifica en sus textos.

El blog, la participación en foro y la revista o web de divulgación especializada también son virtualmente similares a la distribución del artículo en revista científica, si bien cuentan con mucha menos terminología. Esta distribución similar puede justificarse, de nuevo, con el argumento esgrimido anteriormente: la mayor restricción de potenciales lectores de dichos textos. Además, debe recordarse que desde el punto de vista del subemisor, el blog y la revista o web de divulgación especializada se incluyen dentro de la categoría de especialista no investigador (*vid. supra* apartado IV. 1.3.1., p. 290). Por último, en el caso de la noticia y la entrevista, no existen diferencias ni entre los tipos textuales ni entre los temas, ya que la distribución es similar (si bien la entrevista cuenta con muchas más FLE que la noticia).

1.3.5. Relación con los PROCEDIMIENTOS DE REFORMULACIÓN FORMAL

Este apartado compara los procedimientos de reformulación formal de cada subcorpus según el tema para los siete tipos textuales del corpus (*vid. infra* Gráfico 59 a Gráfico 65).

Como punto de partida puede afirmarse que la distribución de frecuencia de los tres tipos de PRF son similares en los tres temas en todos los textos, a excepción de algunos casos que se detallarán a continuación. Como norma general, los procedimientos de expansión son los más numerosos, a los que le siguen los de variación y, por último, los de reducción (como se indicaba en el Gráfico 21, *vid. supra* p. 268) y, además, los tipos textuales con mayor número de PRF son el artículo en revista científica y la tesis doctoral, lo que resulta lógico si tenemos en cuenta que son, a su vez, los textos con mayor terminología.

Destaca la informática en el caso de la entrevista en los PRF por expansión, sobre todo en aquellos casos en los que se observan estrategias denominativas (*vid. infra* Tabla 99).

Los «hackers» son investigadores de seguridad, los cibercriminales son los que cometen delitos y los «hacktivistas» son los que indican ideología por medio de las técnicas de «hacking».

(INF ENT05) (Traductor)

Tabla 99: Ejemplo de PRF por expansión en blog sobre informática

Los procedimientos de variación, por otra parte, tienen un protagonismo relevante en la tesis doctoral, especialmente en medicina, donde existen numerosos casos de acrónimos o variantes denominativas (Gráfico 65). El caso de los procedimientos por variación en informática es mucho más numeroso en el artículo en revista científica (Gráfico 59), lo cual puede vincularse al hecho de que también sea el que cuenta con mayor número de FLE c) (que son las más crípticas y poco frecuentes en el discurso común). Es normal encontrar en el corpus una forma críptica a la que acompaña un procedimiento de variación, como una variante denominativa o un acrónimo (*vid. infra* Tabla 100).

Para el control de los movimientos de la plataforma se diseñan unos algoritmos de control específicos conocidos como algoritmos de generación de claves gravitoinerciales (en inglés Motion Cueing/Drive Algorithms – MCA/MDA) (Schmidt and Conrad 1969), también conocidos como algoritmos de washout. Estos algoritmos toman como entrada el estado físico del vehículo simulado y generan como salida la pose deseada para la plataforma robótica en forma de grados de libertad (GdL) traslacionales y rotacionales, que es transformada en consignas para el control de los actuadores (Garrett and Best 2010).

(INF ART13) (Usuario medio)

Tabla 100: Ejemplo de PRF por variación en artículo en revista científica sobre informática

Otro caso que también despunta si se compara con los otros temas es la participación en foro en medicina, que cuenta con 232 PRF por expansión (Gráfico 63), ya que en dicho tipo textual se suelen explicitar o describir diferentes fenómenos, enfermedades, medicamentos, síntomas o signos (*vid. infra* Tabla 101).

La descarga que efectúa la excitación del músculo cardíaco proviene en estos casos de un punto cualquiera de las aurículas o de los ventrículos (el corazón está formado por dos aurículas y dos ventrículos), pero una fracción de tiempo muy pequeña antes de la llegada del impulso sinusal normal (que es el originado en una parte especial del corazón llamado “nodo sinusal”, que es el lugar de inicio normal de los impulsos eléctricos que hacen que el corazón se contraiga e impulse la sangre).

(MED FOR02) (Usuario medio)

Tabla 101: Ejemplo de PRF por variación en artículo en revista científica sobre informática

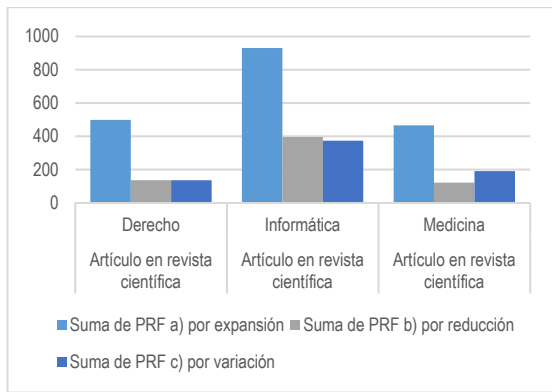


Gráfico 59: Tipos de PRF según temas para el artículo en revista científica

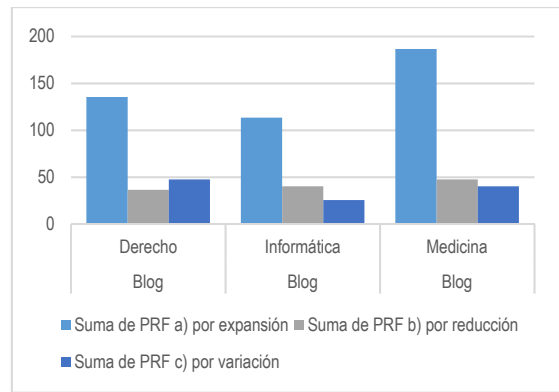


Gráfico 60: Tipos de PRF según temas para el blog

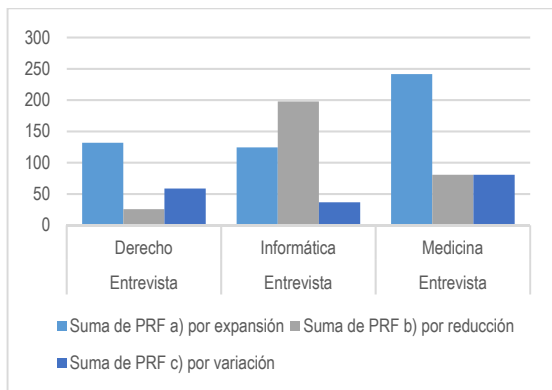


Gráfico 61: Tipos de PRF según temas para la entrevista

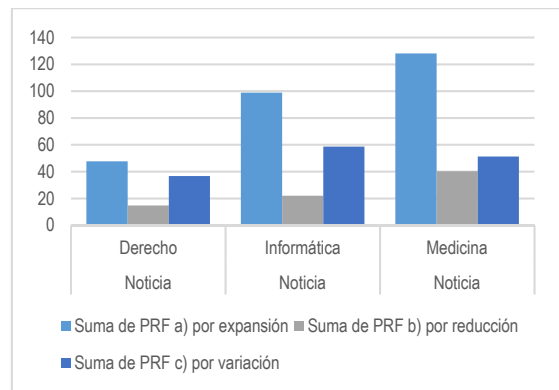


Gráfico 62: Tipos de PRF según temas para la noticia

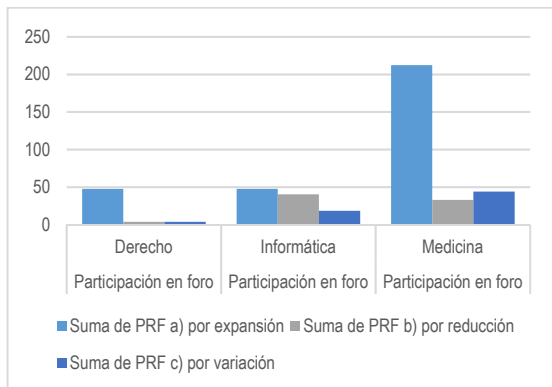


Gráfico 63: Tipos de PRF según temas para la participación en foro

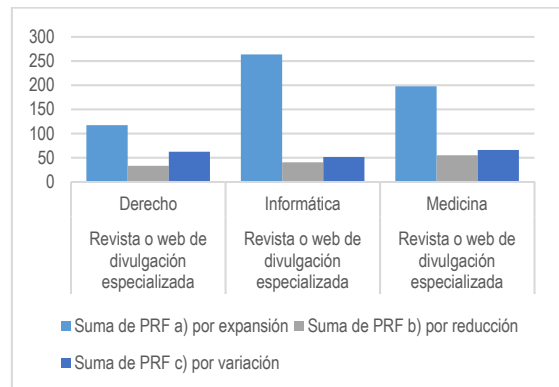


Gráfico 64: Tipos de PRF según temas para la revista o web de divulgación especializada

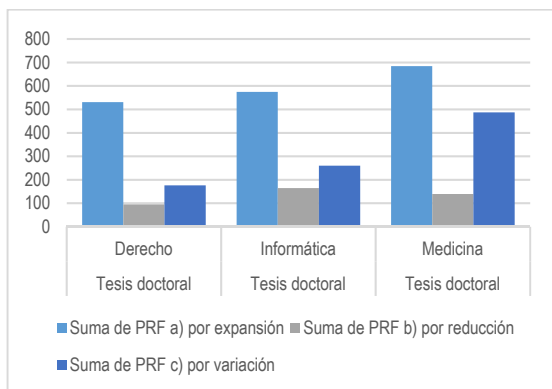


Gráfico 65: Tipos de PRF según temas para la tesis doctoral

1.3.6. Conclusiones parciales sobre el TIPO DE TEXTO

Esta última sección sirve de recapitulación de los datos más representativos del cruce de la variable de tipo de texto con los datos del resto de variables.

- a) La vinculación entre el tipo de texto y el subemisor es similar a la que existe con la clase según el grado de especialización.
- b) La relación con el subdestinatario se divide en tres grupos: (i) el primero, dedicado al artículo en revista científica y tesis doctoral con destinatario especialista investigador; (ii) el segundo, vinculado al instruido lego formado y al lego no formado, con frecuencias para todos los tipos textuales de en torno al 20 %; y (iii) el tercero y último, en el que se incluyen el resto de destinatarios, con frecuencias muy escasas en el corpus. El grupo que se encuentra a caballo entre el instruido lego formado y el lego no formado parece diferenciarse a través de la densidad terminológica, que es menor en los casos con subdestinatario lego no formado, especialmente la relacionada con los términos crípticos.
- c) El patrón de distribución de los perfiles según la competencia cognitivo-comunicativa del subemisor y el subdestinatario es similar en varios tipos textuales. En primer lugar, coincide en el artículo en revista científica y la tesis doctoral, que tienen el punto de inversión de la distribución en el especialista no investigador. Por otro lado, coinciden el blog, la entrevista, la noticia y la revista o web de divulgación especializada, que cuentan con el punto de inversión en el instruido aprendiz de especialista. Por último, la participación en foro es la que más difiere en la distribución de subemisor-subdestinatario del resto de tipos textuales, ya que el subemisor más frecuente es aquel que cuenta con menor competencia cognitivo-comunicativa (el lego no formado), al contrario que el resto de tipos textuales.
- d) En cuanto a las FLE, existen diferencias entre temas y entre tipos textuales. El derecho presenta una distribución homogénea (decreciente desde términos délficos hasta términos crípticos) en todos los tipos textuales. La informática presenta una distribución desigual en sus tipos de FLE y en los tipos textuales del corpus. Por último, la medicina también se vincula con una distribución desigual de tipos de FLE y tipos textuales, con predominancia de los términos délficos en todos los tipos textuales, a excepción de la tesis doctoral, en la que predominan los términos crípticos.

e) La distribución de frecuencia de los tres tipos de PRF son similares en los tres temas en todos los textos, con algunas excepciones. Como norma general, los procedimientos de expansión son los más numerosos, a los que le siguen los de variación y, por último, los de reducción. Los tipos textuales con mayor número de PRF son el artículo en revista científica y la tesis doctoral. Destacan los casos de PRF por expansión en la entrevista en informática, los procedimientos por variación en informática en el artículo en revista científica o en la tesis doctoral en medicina.

Para resumir este apartado final, se sintetizan las características prototípicas fruto del análisis de los elementos más frecuentes en el corpus para los diferentes tipos textuales (*vid. infra* Tabla 102).

Tipo textual	Subemisor	Subdestinatario	Clase	IDT	IR
Artículo en revista científica	Especialista investigador	Especialista investigador	Especializado	0,248	85,048
Blog	Semiespecialista / Especialista no investigador	Instruido lego formado / Lego no formado	Semiespecializado	0,251	20,887
Entrevista	Semiespecialista / Especialista no investigador	Instruido lego formado / Lego no formado	Semiespecializado	0,157	63,169
Noticia	Semiespecialista	Lego no formado	No especializado	0,187	18,893
Participación en foro	Instruido lego formado / Lego no formado	Instruido lego formado	No especializado	0,152	21,332
Revista o web de divulgación especializada	Semiespecialista / Especialista no investigador	Instruido lego formado	Semiespecializado	0,235	25,651
Tesis doctoral	Especialista investigador	Especialista investigador	Especializado	0,292	70,698

Tabla 102: Características más representativas para tipos textuales

1.4. Análisis del ÍNDICE DE DENSIDAD TERMINOLÓGICA

Este apartado se centra en una de las variables numéricas de análisis. La densidad terminológica permite conocer el número de términos que existe en un texto en relación con el número total de unidades léxicas. Por el contrario a los apartados dedicadas al cómputo de formas léxicas especializadas, este apartado se centra en comparar la densidad terminológica según las diferentes variables.

1.4.1. Relación con el TEMA

En primer lugar, puede afirmarse que la densidad terminológica no permite discriminar el tema de un texto. Como se observa en el Gráfico 66, el promedio de IDT global para cada tema es muy próximo al promedio de IDT global (que se sitúa en 0,217), lo que no permite establecer diferencias representativas entre temas basadas en la densidad terminológica textual. Así, la medicina es el tema con mayor densidad terminológica (0,230), seguido del derecho (0,220) y de la informática (0,201).

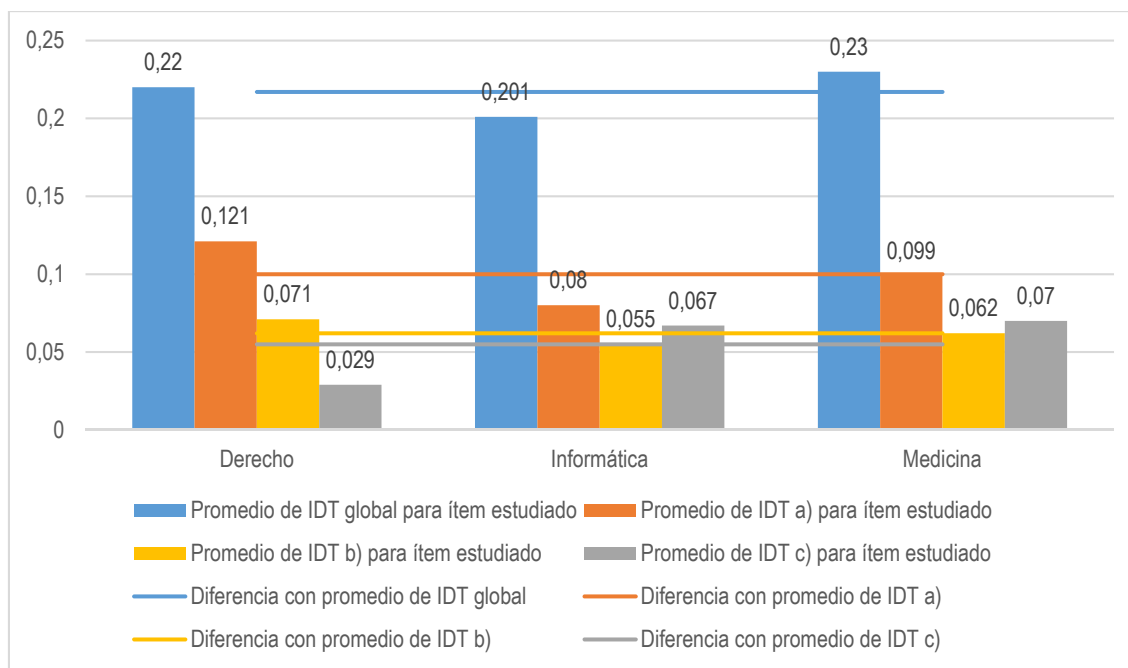


Gráfico 66: Promedio de IDT global y tipos según tema y diferencias con promedio de IDT global y tipos

En cuanto a los diferentes tipos de densidades según cada tipo de término, la mayor densidad para las formas délficas, es decir, aquellas reconocidas como frecuentes en el discurso común, la ostenta el derecho (0,121). De hecho, supera en 0,04 puntos a la informática (0,08) y 0,02 a la medicina (0,099), y se sitúa por encima de la media de IDT a) (0,1) (de hecho, la medicina coincide con el promedio de IDT de términos délficos). Estos datos proporcionales al total de formas léxicas coinciden con la distribución del total de formas délficas del Gráfico 20 (*vid. supra*, p. 266), en los que el orden, de mayor a menor, parte del derecho, pasando por la medicina hasta el menor valor de términos délficos en la informática.

Al margen de los tributos habituales (normalmente impuestos de pago periódico), una de las tasas más comunes en la Administración es la de derechos de examen. En este sentido queremos mencionar el servicio de *Inscripción en Pruebas Selectivas*. La inscripción en pruebas selectivas es un servicio de la AGE que facilita a los ciudadanos la inscripción por Internet en las pruebas selectivas de acceso a la Administración. El ciudadano puede hacer la inscripción completamente en línea rellenando el formulario, realizando el pago electrónico y registrando electrónicamente la solicitud. El ciudadano, si es que así lo desea, también puede rellenar el formulario, imprimirlo y posteriormente, presentarlo en ventanilla previo

pago en la entidad bancaria. La unidad de RRHH tiene acceso a todas las solicitudes registradas, a servicios de verificación de datos (fecha de nacimiento, titulación, familia numerosa, desempleado, discapacidad), a los datos estadísticos, entre otras funcionalidades. También ofrece un servicio para el alta de las solicitudes recibidas de modo presencial a partir de reconocimiento del código de puntos que figura en la parte inferior izquierda de la solicitud del ciudadano.

(DER BLOG12) (Estudiante)

Tabla 103: Ejemplo de texto con IDT a) elevado

Los otros dos tipos no coinciden con los datos analizados en el Gráfico 20. En primer lugar, el IDT de términos semicrípticos es mayor en el derecho, lo que coincide con los datos de número de FLE (27 371 formas semicrípticas). El siguiente valor corresponde con la medicina (0,062), que, sin embargo, es el valor menos elevado en número de términos semicrípticos (23 278) si bien se diferencia por muy poco de la informática (23 563 formas semicrípticas), que cuenta con un IDT de 0,055. En este caso, de nuevo, la medicina coincide con el promedio de IDT de términos semicrípticos (0,062).

Asimilar lo expresado por el artículo 63 en comentario con otras modalidades de vinculación indirecta de trabajadores, como está ocurriendo en algunos casos, va en contravía de las normas vigentes que tienen claramente reguladas las modalidades de vinculación de trabajadores asalariados por vía directa con su empleador a través de contratos a término indefinido, término fijo, por obra o labor determinado y por labores ocasionales o transitorias, o por medio del sector solidario de la economía con las características propias de un trabajador cooperado al que le aplica un régimen especial por fuera del laboral. Plantear otras modalidades de vinculación de trabajadores carece de fundamento.

(DER DIV12) (Usuario medio)

Tabla 104: Ejemplo de texto con IDT b) elevado

Por último, en el caso del IDT de términos crípticos, la relación entre informática y medicina es muy similar. En efecto, los datos, tanto de FLE y de IDT, son prácticamente idénticos en ambos casos. La medicina cuenta con el mayor valor de IDT c) (0,07) (29 915 formas crípticas), seguido de muy cerca del valor de la informática (IDT = 0,067) (27 880). El derecho, por el contrario, es el que cuenta con menor número de términos crípticos (11 506 formas) y menor IDT (0,029). Así, de esta forma, solo es posible defender que la densidad terminológica de las formas crípticas permite distinguir el derecho, dada la ausencia de estas unidades en los textos si son comparadas con la medicina o la informática.

—El dominio N-terminal (residuos 1-206) que forma una estructura conocida como plegamiento de Rossmann, con seis hojas β paralelas (S1-S6) y seis hélices (H1-H6) alternadas. Este dominio se encuentra involucrado en la unión del GTP/GDP.

—El dominio central (residuos 207-384) que está formado por cuatro hojas β paralelas (S7-S10) y tres hélices (H8-H10) entremezcladas. Este dominio es el encargado de los contactos laterales entre los monómeros de α,β -tubulina que van a originar los protofilamentos.

—El dominio C-terminal (a partir del residuo 385) que consiste en dos hélices α antiparalelas (H11 y H12). Este dominio está implicado en las interacciones con las proteínas asociadas a microtúbulos (MAPs), tales como la proteína tau.

(MED TES01) (Traductor)

Tabla 105: Ejemplo de texto con IDT c) elevado

1.4.2. Relación con la CLASE

En el caso de la clase según el grado de especialización, puede identificarse una vinculación directa: a mayor grado de especialización, mayor índice de densidad terminológica (*vid. infra* Gráfico 67), lo cual permite defender que la densidad terminológica contribuye a discriminar el grado de especialización textual. De hecho, esta relación no solo se establece entre la clase y el IDT global, sino también entre la clase y el resto de tipos de densidades para cada tipo de término.

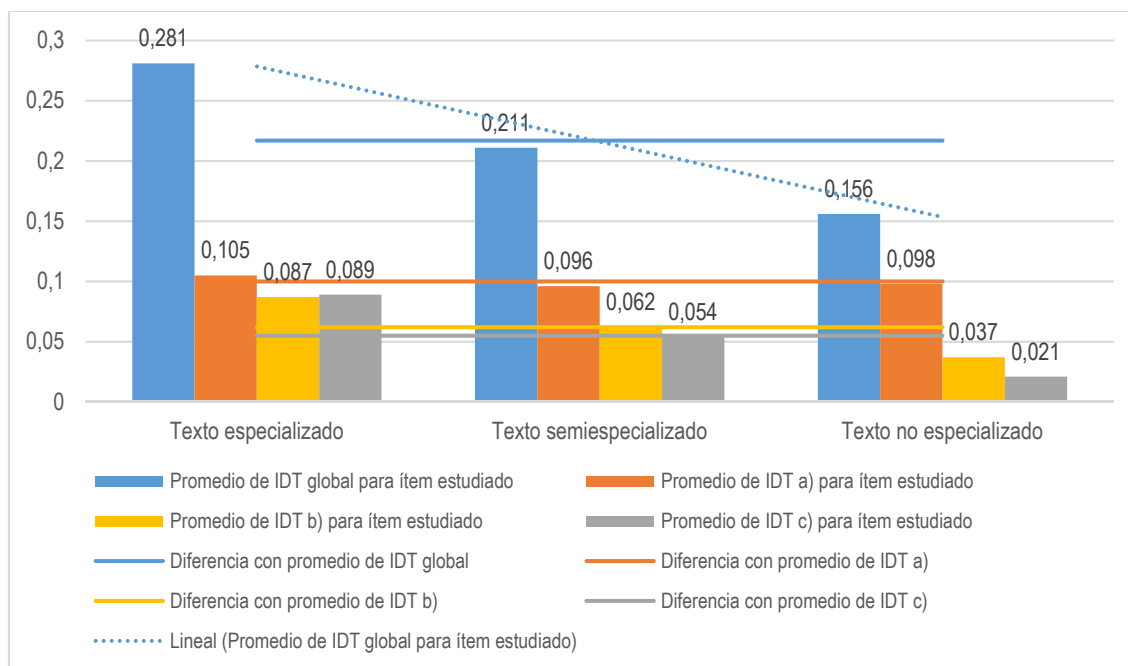


Gráfico 67: Promedio de IDT global y tipos según clase y diferencias con promedio de IDT global y tipos

En este sentido, debe mencionarse que el texto especializado es la única de las clases que se sitúa por encima del promedio de IDT global (0,217), ya que cuenta con un IDT de 0,281. De hecho, tanto el texto semiespecializado como el no especializado se sitúan por debajo de la media, si bien el primero está muy próximo a ella (0,211), no siendo comparable el caso del texto no especializado, que se encuentra a 0,061 puntos por debajo (0,156). Esta relación también se establece en el caso del IDT b) e IDT c) pero no en el caso del IDT a), que es similar en las tres clases: el promedio es 0,1 y las tres clases, de mayor a menor especialización, se sitúan en torno a dicho promedio, correspondiendo, respectivamente, con los valores 0,105, 0,096 y 0,098.

En este último caso se observa cómo existe una ligera diferencia a favor del texto no especializado en el IDT a) con respecto al texto no especializado. A nuestro juicio dicha diferencia no es representativa ni permite diferenciar las clases.

1.4.3. Relación con el TIPO DE TEXTO

De los siete tipos textuales que conforman el corpus, cuatro de ellos superan el promedio de IDT global (0,217) (*vid. infra* Gráfico 68). En orden de mayor a menor IDT, estos tipos textuales son la tesis doctoral (0,292), el blog (0,251), el artículo en revista científica (0,248) y la revista o web de divulgación especializada (0,235). En los cuatro casos, los tres tipos de FLE superan la media para cada tipo. Además, los valores para cada tipo de IDT son muy próximos en los cuatro tipos textuales, si bien la tesis doctoral supera ampliamente a todos los tipos textuales en todos los IDT y la revista o web de divulgación especializada es la que menor IDT global presenta de este grupo.

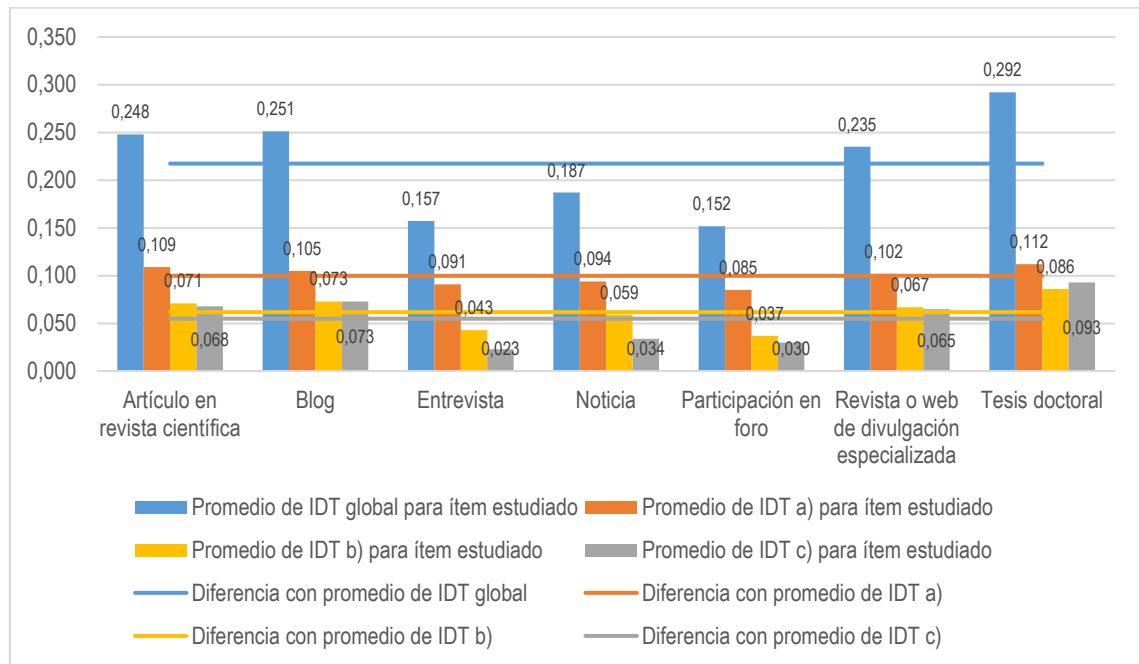


Gráfico 68: Promedio de IDT global y tipos según tipo de texto y diferencias con promedio de IDT global y tipos

Una de las cuestiones que más llama la atención es el hecho de que el blog cuenta con un valor de IDT mayor al del artículo en revista científica (aunque la diferencia se sitúa tan solo en 0,003 puntos), a pesar de que en diversos apartados (*vid. supra* IV. 1.2.1., Gráfico 23; IV. 1.3.1., Gráfico 24; IV. 1.3.2., Gráfico 25, pp. 273 y ss.) se ha vinculado repetidamente el artículo en revista científica con el subemisor y el subdestinatario especialista investigador y con la clase de texto especializado, lo que no ha sido posible establecer para el blog. En efecto, la ligera diferencia se establece en cuanto a la mayor densidad de términos crípticos sobre el total de unidades léxicas, lo que permite justificar la ligera diferencia a favor del blog.

Gracias a los datos expuestos en la Tabla 95 (*vid. supra*, p. 293), se concluía que un valor reducido de IDT c) se asociaba al subdestinatario lego no formado en contraposición

a la subcategoría de destinatario inmediatamente superior (instruido lego formado). De hecho, el blog o la revista o web de divulgación especializada son dos de los tipos textuales que muestran estas diferencias (*vid. supra* Tabla 95, p. 293).

Los otros tres tipos textuales, en orden de mayor a menor IDT, corresponde con la noticia, la entrevista y la participación en foro (0,187, 0,157 y 0,152 respectivamente), que se alejan del promedio de IDT global en varias décimas. Lo mismo sucede con el resto de los tipos de IDT, que se sitúan por debajo de las respectivas medias (aunque la noticia, con un IDT b) de 0,059 se sitúa a tan solo 0,003 puntos de la media (0,062). Los casos que cuentan con menor IDT global son los mismos que cuentan con menores valores de IDT c), es decir, la entrevista, la noticia y la participación en foro.

A pesar de estas diferencias, se aprecia que el IDT a) es muy similar en todos los tipos textuales; tan solo la participación en foro se aleja del valor promedio. De hecho, la principal diferencia que se observa entre todos los tipos textuales se halla en el IDT b) y el IDT c), que son muy similares en aquellos tipos textuales que cuentan con valores de IDT más elevados, como son el artículo, el blog, la revista o web de divulgación especializada o la tesis doctoral, puesto que los otros tres tipos textuales presentan valores más alejados de la media. En todo el corpus, solo en el caso de la tesis doctoral, el IDT de términos crípticos supera al IDT de términos semicrípticos (en el blog los dos valores coinciden (IDT b) = IDT c) = 0,73).

Si atendemos a los tres temas del corpus según los diferentes tipos textuales, puede comprobarse que los datos que se relacionan entre el Gráfico 69 y el Gráfico 75 (*vid. infra*) son muy similares a los que presentan los tipos de FLE según temas para cada tipo textual (*vid. supra* Gráfico 52 a Gráfico 58). De hecho, las diferencias que se identificaban entre los temas y entre los tipos textuales coinciden con las que se presentan para la densidad terminológica.

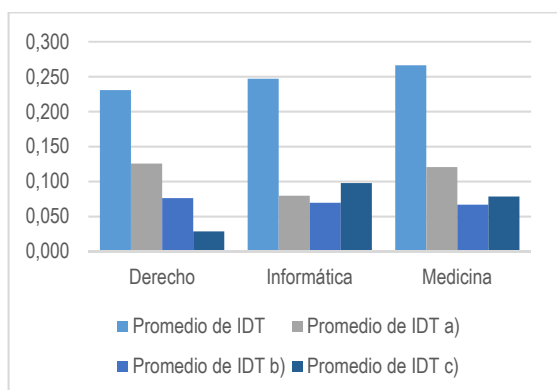


Gráfico 69: Tipos de IDT según temas para el artículo en revista científica

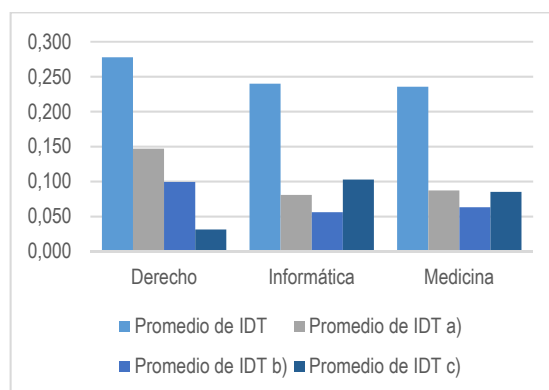


Gráfico 70: Tipos de IDT según temas para el blog

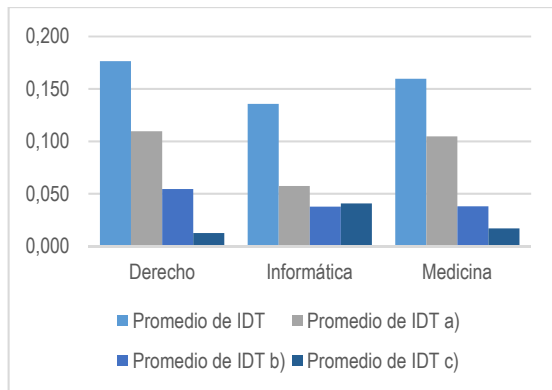


Gráfico 71: Tipos de IDT según temas para la entrevista

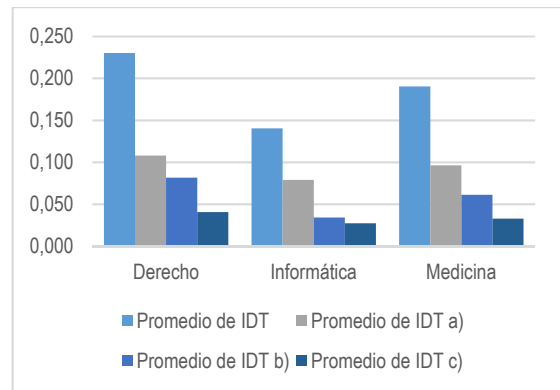


Gráfico 72: Tipos de IDT según temas para la noticia

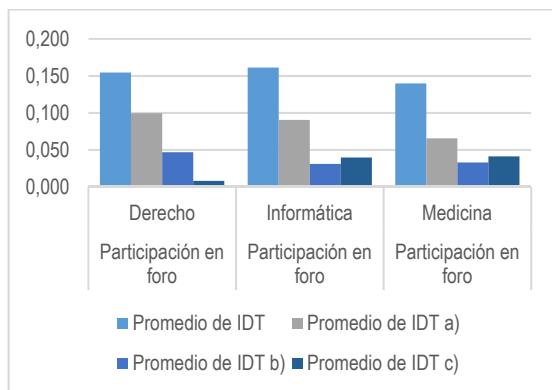


Gráfico 73: Tipos de IDT según temas para la participación en foro

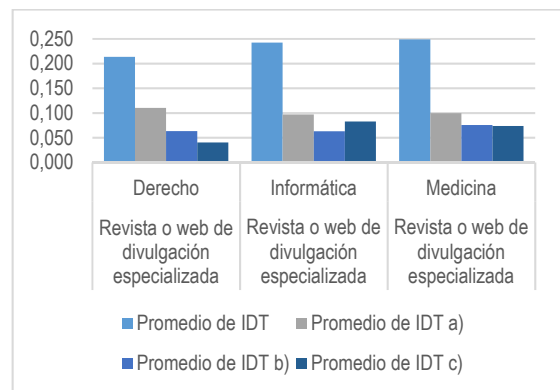


Gráfico 74: Tipos de IDT según temas para la revista o web de divulgación especializada

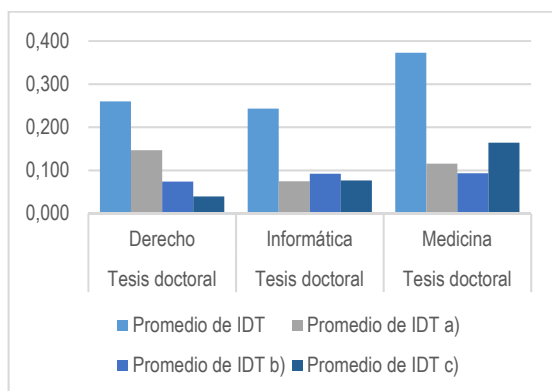


Gráfico 75: Tipos de IDT según temas para la tesis doctoral

1.4.4. Relación con el EMISOR y el SUBEMISOR

En cuanto al emisor y subemisor, la única categoría que supera el promedio es el especialista (IDT = 0,262). De hecho, en cuanto a los subdestinatarios, el especialista investigador (IDT = 0,276) y el especialista no investigador (IDT = 0,239) son las únicas categorías que destacan sobre el promedio de IDT (0,217) por varias décimas (*vid. infra* Gráfico 76).

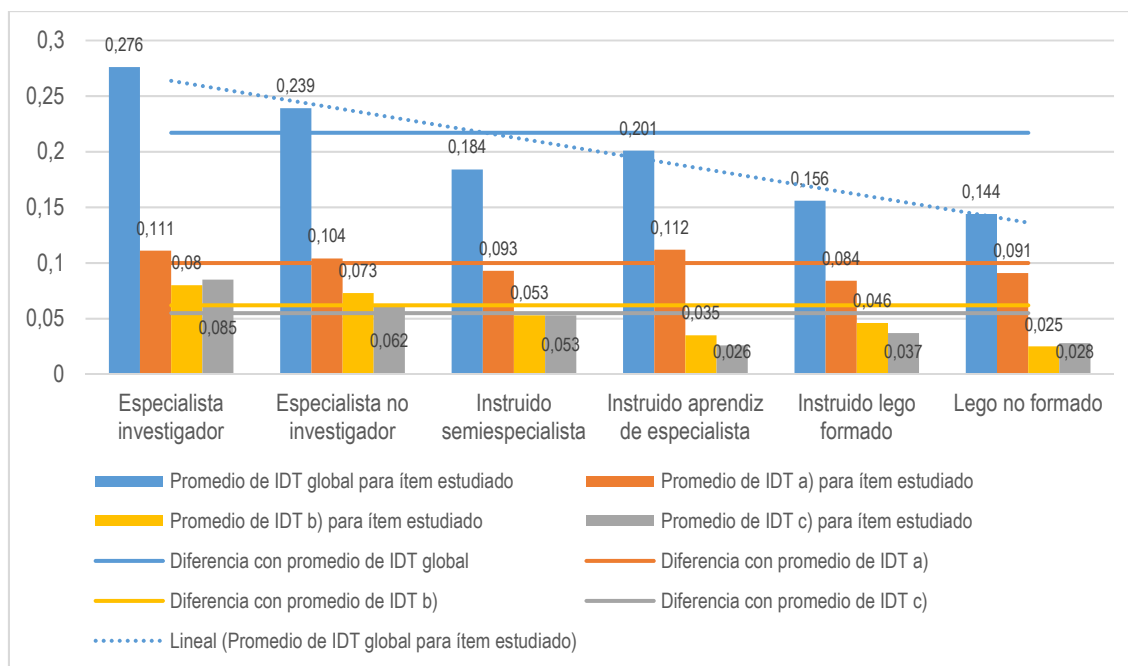


Gráfico 76: Promedio de IDT global y tipos según subemisor y diferencias con promedio de IDT global y tipos

El IDT global, al igual que en el caso de las clases (*vid. supra* I.V. 1.4.2., p. 310), tiene una relación directa con la competencia cognitivo-comunicativa del subemisor. A excepción del caso instruido semiespecialista (IDT = 0,184) e instruido aprendiz de especialista (IDT = 0,201), que aparecen como invertidos. A nuestro juicio, esto se debe, principalmente, al escaso número de textos emitidos por un instruido aprendiz de especialista (27 textos; frente a 135 emitidos por semiespecialista), lo que no permite contar con suficientes patrones para definir de forma precisa el IDT para dicho subemisor. No obstante, si atendemos a las densidades para los términos semicríticos y los términos críticos (especialmente a esta última), se comprueba que destacan en el caso del subemisor semiespecialista, y no en el del instruido aprendiz de especialista. A pesar de esta leve diferencia, consideramos que sí puede establecerse una relación directa entre el emisor de un texto según su competencia cognitivo-comunicativa y el IDT que presente el texto.

En cuanto al resto de tipos de IDT, los niveles de IDT de términos délficos en los seis subemisores es similar, como se describía anteriormente en el apartado IV. 1.4.3. (p. 311). No obstante, el mayor valor se vincula al instruido aprendiz de especialista y no al especialista, como podría creerse *a priori*, lo que puede justificarse a partir del reducido número de patrones clasificados como aprendiz de especialista. Por otra parte, el lego no formado presenta un valor de IDT de términos délficos similar al del semiespecialista. No obstante, como se mencionó en los apartados anteriores (*vid. supra* IV. 1.4.2, Gráfico 67 y I.V. 1.4.3, Gráfico 68, p. 310 y 311), el IDT de términos délficos no parece contribuir a

establecer diferencias entre patrones. En contraste, el caso del IDT de términos semicrípticos e IDT de términos crípticos sí parece establecer una relación directa (además de gradual) con los perfiles de los usuarios según su competencia cognitivo-comunicativa, a pesar de los valores excepcionales del instruido aprendiz de especialista. En estos casos de IDT, tan solo el perfil de especialista supera los respectivos promedios, el semiespecialista se acerca por debajo al promedio y el resto de los usuarios se alejan de él. Cabe destacar, además, que el usuario especialista es el único que cuenta con mayor IDT c) que IDT b), lo cual se ajusta a las características observadas para la tesis doctoral (*vid. supra* Gráfico 68, p. 311), especialmente en medicina (*vid. supra* Gráfico 75, p. 313), como se muestra en el ejemplo de la Tabla 106.

1.5.2 Dinámica de microtúbulos

La dinámica de polimerización y despolimerización de microtúbulos juega un papel crucial en la función celular. El proceso de polimerización comienza cuando un nuevo dímero de tubulina se añade al extremo (+) del microtúbulo en crecimiento, el dominio catalítico de la α -tubulina contacta con el GTP de la β -tubulina recién incorporada y provoca su hidrólisis a GDP. Como resultado, el cuerpo del microtúbulo está compuesto por β -tubulina GDP, salvo en el extremo (+), en el que existe una capa de dímeros de tubulina con la subunidad β unida a GTP, denominada usualmente tapón (*cap*) de GTP. La pérdida del tapón de GTP provoca la despolimerización del microtúbulo, liberándose los heterodímeros de α , β -tubulina al medio celular.

(MED TES01) ^(Traductor)

Tabla 106: Ejemplo de IDT c) elevado en tesis doctoral sobre medicina

1.4.5. Relación con el DESTINATARIO y el SUBDESTINATARIO

El caso del destinatario es, en cierta forma, contrario al emisor, pues cuatro de las seis subcategorías de destinatario se encuentran por encima del promedio de IDT (*vid. infra* Gráfico 77); tan solo las dos categorías de usuarios con menor competencia cognitivo-comunicativa se sitúan por debajo del promedio. En este caso, la relación entre IDT y usuarios según la competencia cognitivo-comunicativa es directa puesto que se observa cómo el IDT disminuye a medida que lo hacen también los diferentes perfiles. Este patrón y la mencionada relación directa también se establece en el IDT de términos semicrípticos e IDT de términos crípticos pero no en el IDT de términos délficos. De hecho, el valor más elevado para el IDT a) corresponde con el especialista investigador (*vid. infra* Tabla 107), al que le sigue el lego no formado (*vid. infra* Tabla 126, p. 359), de forma contraria a como sucede en el caso del IDT global. Como en los casos mencionados anteriormente, esto puede deberse al número de textos de dicho subcorpus. Aun observando dicha relación directa, en este caso no es posible defender que ninguno de los tipos de densidad

terminológica contribuya a la discriminación de textos según la competencia cognitivo-comunicativa del destinatario.

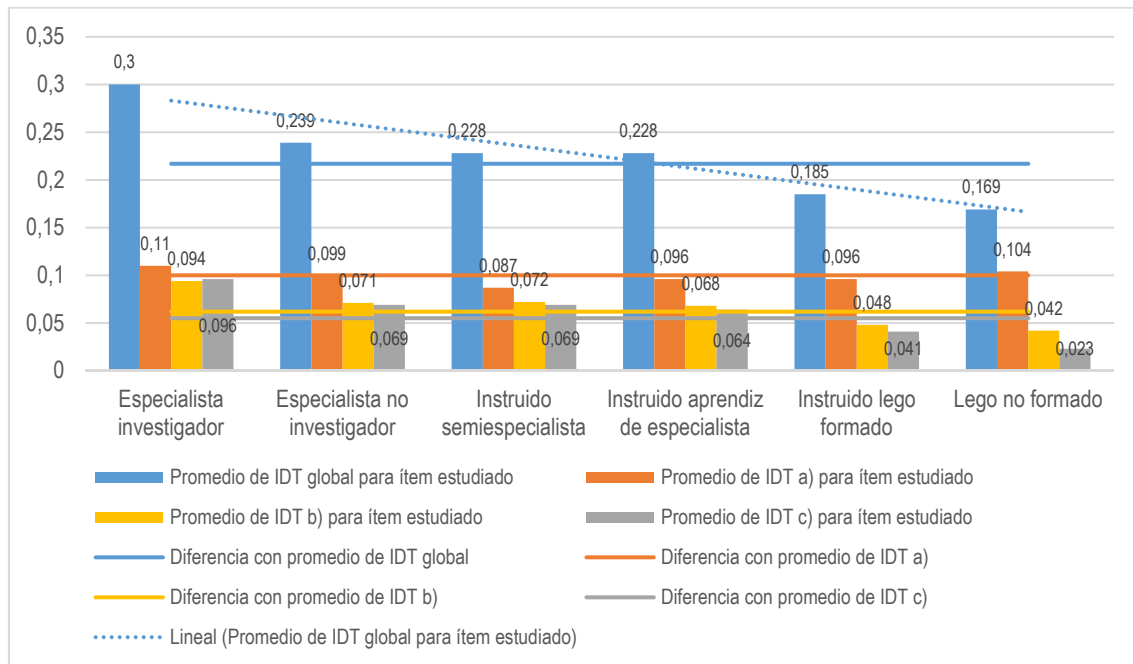


Gráfico 77: Promedio de IDT global y tipos según subdestinatario y diferencias con promedio de IDT global y tipos

Así por ejemplo los autores anglosajones suelen excluir la multa del catálogo de alternativas, ya que no se promueve como una real alternativa a la prisión al tratarse de una pena reservada a delitos leves. También suelen excluir todas aquellas penas que pese a ser penas no privativas de libertad, siempre son accesorias, como la de inhabilitación o la de comiso debido a que no constituyen una alternativa a la pena, sino un complemento a la misma. Por otra parte, a veces se le niega a un instrumento la calidad de “alternativa” por su origen, si es que dicho instrumento no es hijo del movimiento descarceratorio de los años 60, ya que se argumenta que estas penas “presuponen” la prisión (Larrauri 1991, 49). Por ejemplo se puede notar en Vass (1990, 2), quien distingue tres categorías de penas alternativas: aquellas diseñadas para sustituir la prisión (TBC, suspensión de la pena de prisión, probation con condiciones), aquellas que son no privativas de libertad pero no fueron diseñadas específicamente como sustitutos de la prisión pero que pueden ser usadas y/o promovidas como alternativas a la prisión, y las que reducen la estancia en prisión ofreciendo una alternativa (libertad condicional).

(DER TES06) (Estudiante)

Tabla 107: Ejemplo de IDT a) elevado en tesis doctoral con destinatario especialista investigador

¿Y qué pasa con la vacuna?

Siguiendo las recomendaciones de la Asociación Española de Pediatría, la vacuna de la varicela se ha incorporado recientemente a los calendarios vacunales de las diferentes comunidades autónomas, con una primera dosis a los 15 meses y otra entre los 3-4 años de edad. La vacuna de la varicela (Varivax®, Varilrix®) contiene virus vivo atenuado, por lo que cabe la posibilidad de que un niño correctamente vacunado pueda más adelante desarrollar un herpes zóster. Sin embargo, parece que la incidencia de herpes zóster en niños inmunocompetentes vacunados de la varicela es menor que en no vacunados.

(MED BLOG13) (Estudiante)

Tabla 108: Ejemplo de IDT a) elevado en blog con destinatario lego no formado

1.4.6. Relación SUBEMISOR-CLASE-IDT

La relación directa entre el IDT global (y los diferentes tipos de IDT) y la clase ya se estableció en el apartado IV. 1.4.2. (*vid. supra* Gráfico 67, p. 310). Esta relación directa también se observa si se incorporan los valores para el subemisor (*vid. infra* Gráfico 78). En dicha gráfica puede comprobarse cómo los valores de IDT global decrecen a medida que disminuye el grado de especialización del texto en cada uno de los subemisores¹⁶².

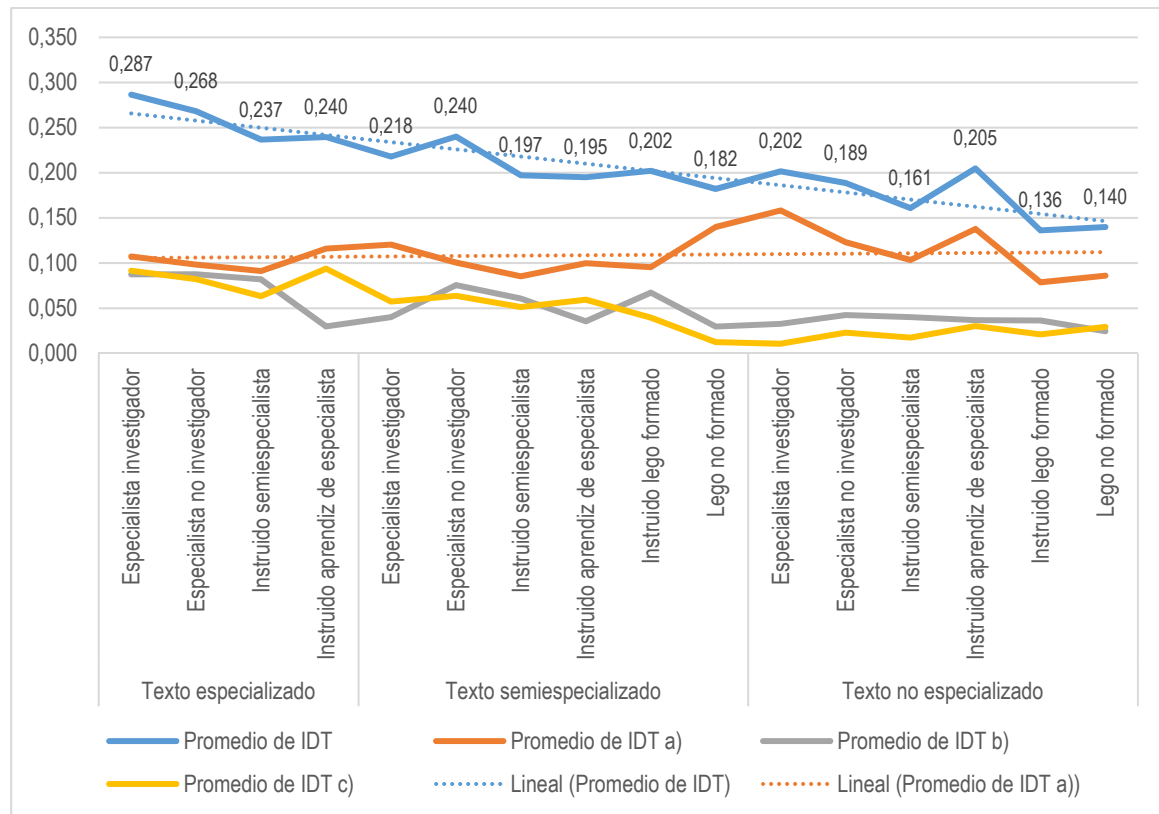


Gráfico 78: Relación de datos globales para la correspondencia subemisor-clase según IDT global y tipos

Por otra parte, el IDT b) y el IDT c) siguen el mismo patrón de distribución que el IDT global, excepto en el caso del instruido aprendiz de especialista en el texto especializado. Además, el IDT de términos crípticos suele situarse ligeramente por debajo del IDT de términos semicrípticos en todos los casos.

Esta distribución contrasta con la del IDT a), que, al contrario que el resto de los tipos, disminuye en los textos especializados y aumenta en los textos no especializados (a excepción del subemisor instruido lego formado y lego no formado).

¹⁶² La diferencia entre texto semiespecializado (IDT = 0,195) y texto no especializado (IDT = 0,205) en el instruido aprendiz de especialista no es suficientemente representativa, especialmente si se tiene en cuenta el reducido número de textos para cada clase (17 y 8 respectivamente).

1.4.7. Relación SUBDESTINATARIO-CLASE-IDT

El caso del subdestinatario parece divergir ligeramente de la interpretación realizada en el apartado anterior. Si bien es cierto que la relación directa entre subdestinatario según la competencia cognitivo-comunicativa y la clase se mantiene (*vid. infra* lineal del Gráfico 79), las diferencias son mucho más pronunciadas, sobre todo entre texto semiespecializado y texto no especializado.

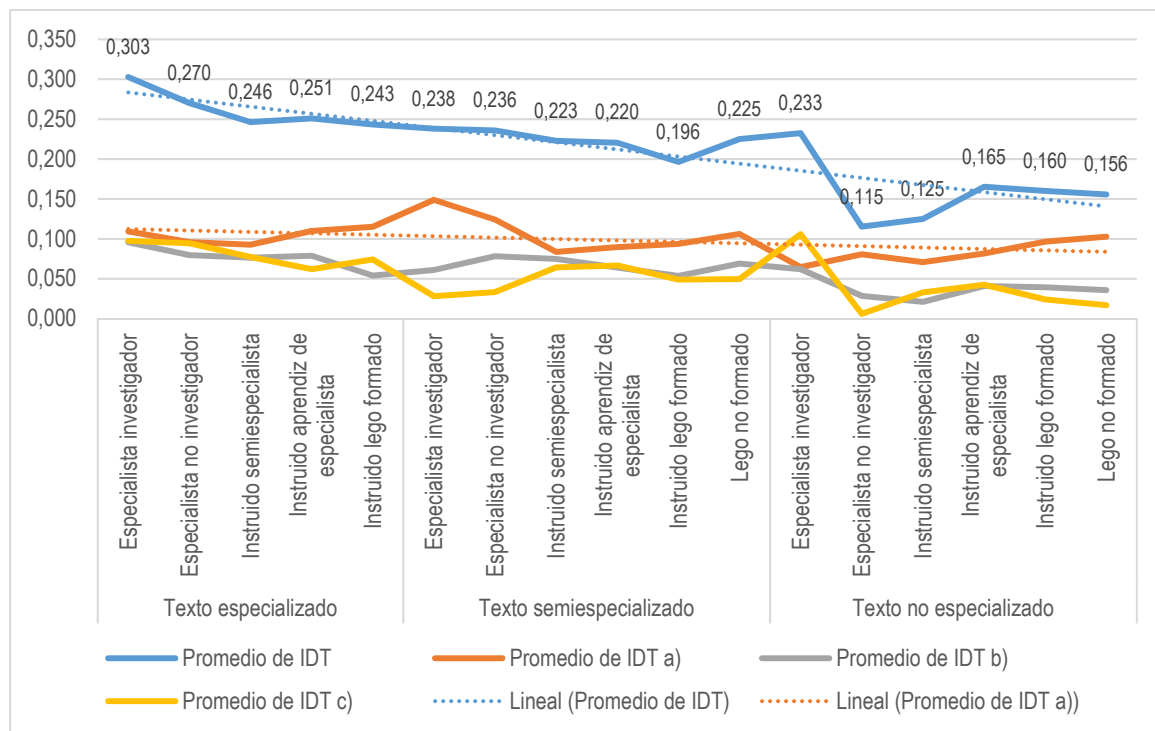


Gráfico 79: Relación de datos globales para la correspondencia subdestinatario-clase según IDT global y tipos

Aunque el IDT de términos semicríticos y el IDT de términos críticos conservan el patrón del IDT global y mantiene las correspondientes reducciones graduales de valores hacia el texto no especializado, el segundo presenta un valor excepcional cuando el destinatario es un especialista no investigador y el texto se clasifica como no especializado. De los 17 textos del corpus con subdestinatario especialista no investigador clasificados como texto no especializado, 16 corresponden con participaciones en foro y 1 con una entrevista (12 en derecho y 5 en medicina). Normalmente, el subemisor de estos textos es lego no formado o, en su defecto, instruido lego formado. El reducido número de textos y las particularidades de este tipo textual pueden explicar este valor excepcional que se comenta, puesto que normalmente las intervenciones de los usuarios se dirigen a resolver dudas sobre el sistema conceptual en el que se inserta un conjunto de denominaciones cuyo significado desconoce, lo que aumenta el número de FLE sobre el total de FL que presenta el texto (*vid.* Tabla 127).

Estimada amiga,

le indico que ciertos datos de los que nos cuentan, sí me parecen que son dignos de estudio y de "ocupación" por su parte, aunque nunca preocupación.

- 1) Esa pérdida de 38 kilos en un año (imagino sin causa justificada aparente).
- 2) el CEA elevado, dado que se trata de un **marcador tumoral**.

La leve **linfocitosis** que nos indica aparece en situaciones normales, por una **viriasis**, por el mismo crecimiento, etc. No es llamativa. Tampoco lo es la TSH, que se mantiene en rango normal, por lo que su **tiroides** sigue funcionando con **hormonas** a nivel normal. Tan solo repítase **analíticas** anuales, hasta un hipotiroidismo clínico o una TSH por encima de 10 no se pone tratamiento.

Por lo que me dice, ya le están haciendo estudios: **colonoscopia**, **endoscopia**, el **TAC**.

Esto es lo verdaderamente importante. Descartar que ese **marcador** sea algo más que secundario a una inflamación (le indico que efectivamente es muy inespecífico y no solo lo elevan los **tumores**).

En resumen, los **síntomas** y **signos** que indica, son dignos de mantener un estudio amplio, pero sin preocuparse porque pueden corresponder a **patologías benignas**.

Un saludo cordial

Dra. Martín

(MED FOR06) ^(Traductor)

Tabla 109: Ejemplo de explicaciones sobre la terminología en una participación en foro no especializada sobre medicina

Según los datos proporcionados por el cruce de valores entre subdestinatario y los diferentes tipos de IDT, puede afirmarse que se mantiene la relación directa entre IDT (sobre todo, global, IDT b) e IDT c)) y subdestinatario de un texto.

1.4.8. Relación SUBEMISOR-SUBDESTINATARIO-IDT

La combinación de las diferentes categorías de emisor y destinatario favorece la idea de relación directa entre usuario e IDT (como ya se ha defendido en los apartados IV. 1.4.6. y IV. 1.4.7.), si bien se identifican numerosas excepciones. Como en otros apartados, en la parte inferior del eje horizontal se indica el subemisor y justo encima se indica el subdestinatario del texto.

Por ejemplo, en el caso del emisor especialista (*vid. infra* Gráfico 80), la vinculación con las categorías de destinatario especialista no es solo muy frecuente (*vid. supra* Gráfico 10, apartado IV. 1.1.4., p. 250) y relacionada con el texto especializado (*vid. supra* Gráfico 27; apartado IV. 1.2.4., p. 281), sino que, además, es la que cuenta con mayor IDT global.

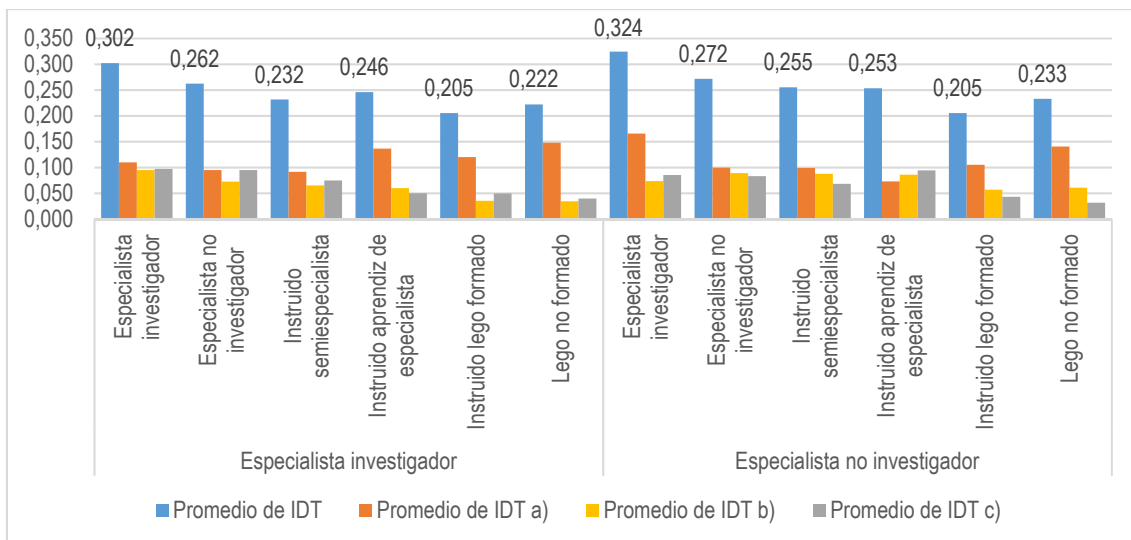


Gráfico 80: Relación de datos globales para la correspondencia emisor especialista-tipo de subdestinatario según IDT global y tipos

En términos generales, el IDT disminuye a medida que disminuye el perfil cognitivo-comunicativo del subdestinatario, como ya se avanzó en el apartado IV. 1.4.7. Los valores del IDT de términos délficos parecen ser mayores cuando el especialista investigador se dirige a los perfiles con menor competencia y, en el caso del especialista no investigador, no se identifica una relación directa. En el caso del IDT de términos semicrípticos establece una relación gradual y más o menos directa con el perfil del destinatario, al igual que el IDT de términos crípticos, que presenta valores que destacan cuando el destinatario es especialista.

Además, destaca el contraste entre tipos de IDT que existe entre subemisor especialista investigador y subdestinatario lego no formado, donde el IDT a) (0,148) destaca muy por encima del IDT b) (0,035) o IDT c) (0,040). Se trata de un total de 8 textos del corpus (7 entrevistas y 1 blog; 1 sobre informática y 7 sobre medicina), todos ellos clasificados como textos no especializados, en los que destacan principalmente aquellos términos que son muy frecuentes en el discurso común, lo cual puede asociarse con las características del tipo textual, más accesible para el conjunto de la población.

1. Fuiste el primer cirujano en hacer el primer trasplante hepático infantil en nuestro país, pero eso fue en junio del año 1985. Desde ese momento, la supervivencia ha mejorado, llegándose a situar en un 85-95% al año. ¿A qué se debe esta mejora?

Hay muchos puntos que lo podrían explicar pero, posiblemente, los más importantes están relacionados con la adquisición de experiencia por parte del equipo de trasplante. El mejor manejo preoperatorio de las complicaciones hepáticas y el apoyo nutricional pretrasplante permite que el paciente llegue en mejor situación a la cirugía. La aplicación de nuevas técnicas quirúrgicas para expandir el pool de donantes (especialmente el donante vivo) ha hecho disminuir el tiempo de espera. Además, las mejoras en los tratamientos inmunosupresores (que cada vez son más efectivos y menos agresivos) y los cambios en los cuidados postoperatorios, por parte de las unidades de intensivistas pediátricos, permiten “cuidar” mejor estos órganos.

(MED ENT12) (Usuario medio)

Tabla 110: Ejemplo de FLE a) en entrevista sobre medicina

¿Qué variables intervienen en la probabilidad de curación de una quemadura grave?

La extensión, la profundidad y la edad de la persona, pues a partir de los 35 años las posibilidades de curación van disminuyendo.

Una persona aún es joven a esa edad...

Efectivamente lo es, pero sus órganos quizás no tanto, o al menos no lo suficiente para poder aguantar con éxito todo el tratamiento que requiere una quemadura grave y que esté extendida. De hecho, en un caso como el tuyo, con el 90% del cuerpo quemado, se sabe que, por debajo de los 35 años de edad, la curación es posible en la mitad de los casos, mientras que a partir de esa edad, esta estadística empieza a reducirse progresivamente. Aunque seamos jóvenes, a partir de los 35 años, las posibilidades de curación de una quemadura grave y extendida van disminuyendo.

(MED ENT11) (Estudiante)

Tabla 111: Ejemplo de FLE a) en entrevista sobre medicina

En cuanto al emisor instruido (*vid. infra* Gráfico 81), existen tres tipos de relaciones: directa en el caso del subemisor instruido semiespecialista (el IDT decrece con el perfil del destinatario); inversa en el subemisor aprendiz de especialista (el IDT global aumenta ligeramente en los tres casos); y directa en el instruido lego formado (con excepción de los perfiles de subdestinatario más inferiores). El emisor instruido no es solo complejo si atendemos al IDT global, sino que también es particular en la distribución del IDT a), que se organiza de forma más errática que en el emisor especialista (*vid. supra* Gráfico 80) o también en la de los otros dos tipos de IDT. De hecho, por ejemplo, el caso del IDT b) del subemisor instruido semiespecialista y subdestinatario especialista investigador (0,133) supera el valor máximo que se observa para IDT b) entre especialistas en el Gráfico 80 (0,095).

En algunos casos concretos, destaca el contraste que existe entre los diferentes tipos de IDT, como en el subemisor instruido semiespecialista y subdestinatario especialista no investigador (IDT a) = 0,147; IDT b) = 0,060; IDT c) = 0,024) o el subemisor aprendiz de especialista y subdestinatario lego no formado (IDT a) = 0,153; IDT b) = 0,037; IDT c) = 0,023).

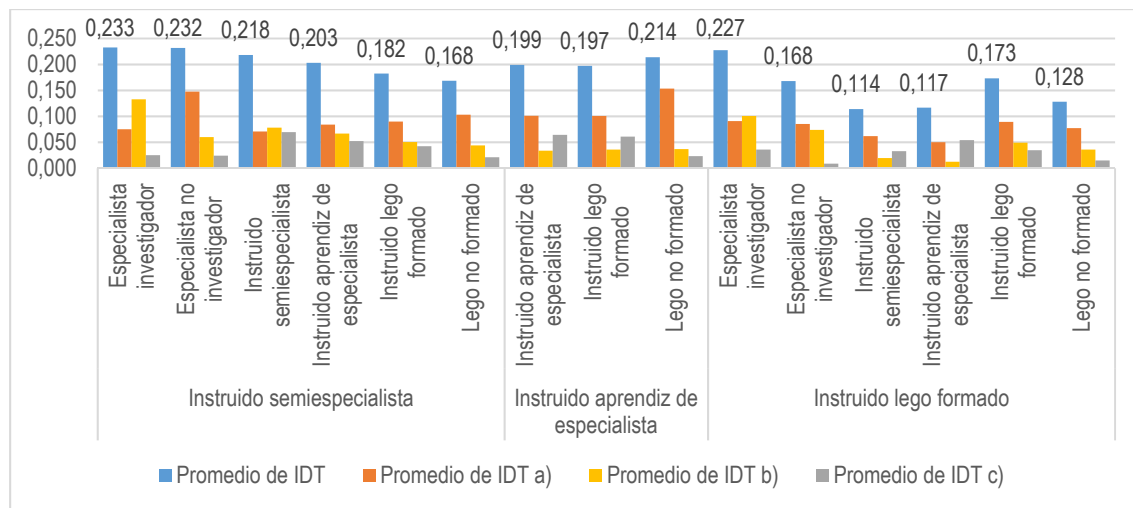


Gráfico 81: Relación de datos globales para la correspondencia emisor instruido-tipo de subdestinatario según IDT global y tipos

Por último, en el caso del emisor lego no se establece una relación clara, puesto que, aunque el mayor IDT se vincula al subdestinatario especialista investigador, es cierto que el resto de subdestinatarios son similares en IDT global. Tampoco existe en el resto de casos, donde la distribución es muy dispar. Por norma general, se observa que el IDT b) e IDT c), es decir, los términos menos frecuentes en el discurso común, son escasos cuando el emisor es un lego no formado, si bien aumentan cuando se dirigen a un subdestinatario especialista investigador, como el caso identificado previamente sobre la participación en foro (*vid. supra*; Gráfico 79; apartado IV. 1.4.7., p. 318).

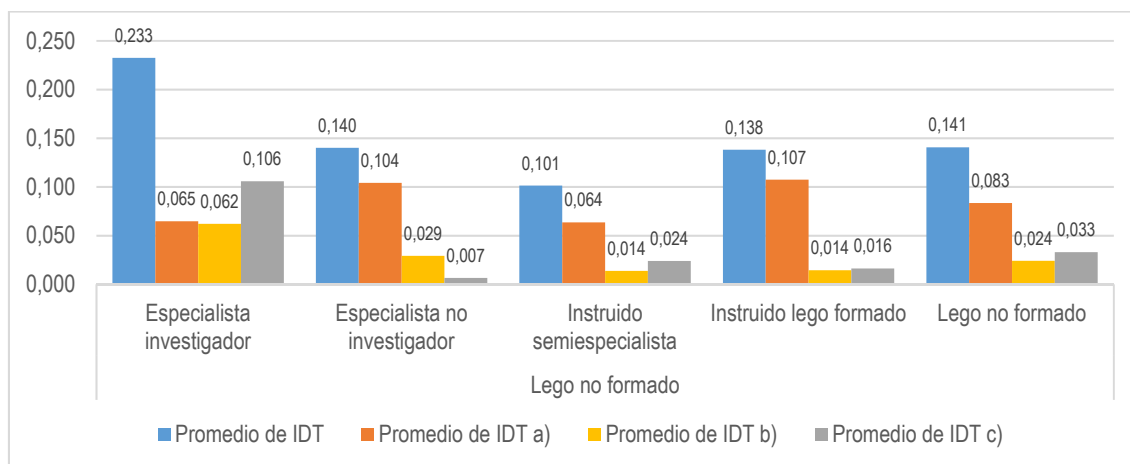


Gráfico 82: Relación de datos globales para la correspondencia emisor lego-tipo de subdestinatario según IDT global y tipos

1.4.9. Conclusiones parciales sobre la DENSIDAD TERMINOLÓGICA

A continuación, se relaciona una recapitulación de las interpretaciones más relevantes sobre los datos relacionados con la densidad terminológica de los textos del corpus.

- a) La densidad terminológica no permite discriminar el tema de un texto. No obstante, según los datos obtenidos, un IDT de términos crípticos reducido permite identificar los textos sobre derecho, pero no discriminar los que tratan sobre informática o medicina. El IDT de términos délficos, por norma general, no contribuye a discriminar ninguna clase, tipo textual o relación discursiva.
- b) En cuanto a la clase, existe una relación directa entre el IDT global de un texto y su grado de especialización, lo cual permite defender que la densidad terminológica contribuye a discriminar el grado de especialización textual. Tan solo el IDT de términos délficos de un texto no contribuye a la discriminación de clases según el grado de especialización.
- c) En relación con los tipos de texto, se pueden distinguir dos grupos de tipos textuales según la densidad terminológica: (i) los que superan el promedio de IDT global, como son, en orden de mayor a menor IDT, la tesis doctoral, el blog, el artículo en revista científica y la revista o web de divulgación especializada; y (ii) los que no alcanzan el promedio de IDT global, como son, en orden de menor a mayor IDT, la entrevista, la participación en foro y la noticia. Estos últimos son, además, los que cuentan con menor IDT de términos crípticos, en contraposición con los tipos textuales con mayor IDT global, que también se asocian a mayor IDT de términos crípticos. Además, las similitudes y diferencias de densidad terminológica de los tipos textuales según los temas coinciden con las expuestas para el número total de formas léxicas especializadas.
- d) Atendiendo al emisor y al subemisor, es posible identificar una relación directa entre el IDT y el tipo de emisor o subemisor según la competencia cognitivo-comunicativa del usuario, ya que, por norma general, el IDT global disminuye a medida que dicha competencia lo hace. Tan solo el emisor especialista y sus subcategorías se posicionan por encima del promedio de IDT global.
- e) En cuanto al destinatario y el subdestinatario, de forma contraria al emisor, tan solo las subcategorías de instruido lego formado y lego no formado se sitúan por debajo del promedio de IDT global. En este caso, la relación entre IDT y usuarios

según la competencia cognitivo-comunicativa es directa puesto que se observa cómo el IDT disminuye a medida que lo hacen también los diferentes perfiles. Este patrón y la mencionada relación directa también se establece en el IDT de términos semicrípticos e IDT de términos crípticos pero no en el IDT de términos délficos. Aun así, teniendo en cuenta que las diferencias son muy sutiles, no es posible defender que ninguno de los tipos de densidad terminológica contribuya a la discriminación de textos según la competencia cognitivo-comunicativa del destinatario.

- f) La relación directa entre IDT y clase también se mantiene si se vincula con el emisor, por una parte, o con el destinatario, por otra. Particularmente, por norma general, el IDT de términos semicrípticos y el IDT de términos crípticos siguen el mismo patrón de distribución que el IDT global.
- g) Por último, si se relacionan subemisor y subdestinatario, se identifica una relación directa entre emisor especialista y los diferentes subdestinatarios, relaciones diversas y complejas cuando el emisor es un instruido y una ausencia de relación clara cuando el subemisor es un lego no formado.

1.5. Análisis del ÍNDICE DE REFORMULACIÓN

Como en el caso anterior, el índice de reformulación permite conocer el grado en el que se reformula un texto atendiendo a la densidad terminológica de cada texto. En primer lugar, se analizará el IR según cada variable de análisis y, en segundo lugar, se cruzarán dichos datos con el IDT ya analizado para comprobar la relación que existe entre estas dos variables. Debe recordarse que no existe correspondencia entre IDT a), IDT b) o IDT c) e IR a), IR b) e IR c), ya que el segundo grupo no pretende ilustrar el grado de reformulación del primero, sino atender a diferentes procedimientos de reformulación, es decir, a) procedimientos por expansión, b) procedimientos por reducción y c) procedimientos por variación. Como se observa en este apartado, todos los tipos de IR suelen distribuirse y comportarse de forma análoga en todas las variables, si bien el IR por reducción (IR b)) es el que presenta una distribución más errática y, a veces, poco relevante.

1.5.1. Relación con el TEMA

Entre las principales similitudes que pueden hallarse entre los temas, se observa que el patrón de distribución suele ser similar (*vid. infra* Gráfico 83). No obstante, la informática presenta mayor reformulación, seguida de la medicina y esta del derecho. En concreto, los dos primeros temas presentan un IR global, IR a) e IR c) sobre la media. El derecho, por el contrario, se sitúa por debajo del promedio de todos los tipos de IR.

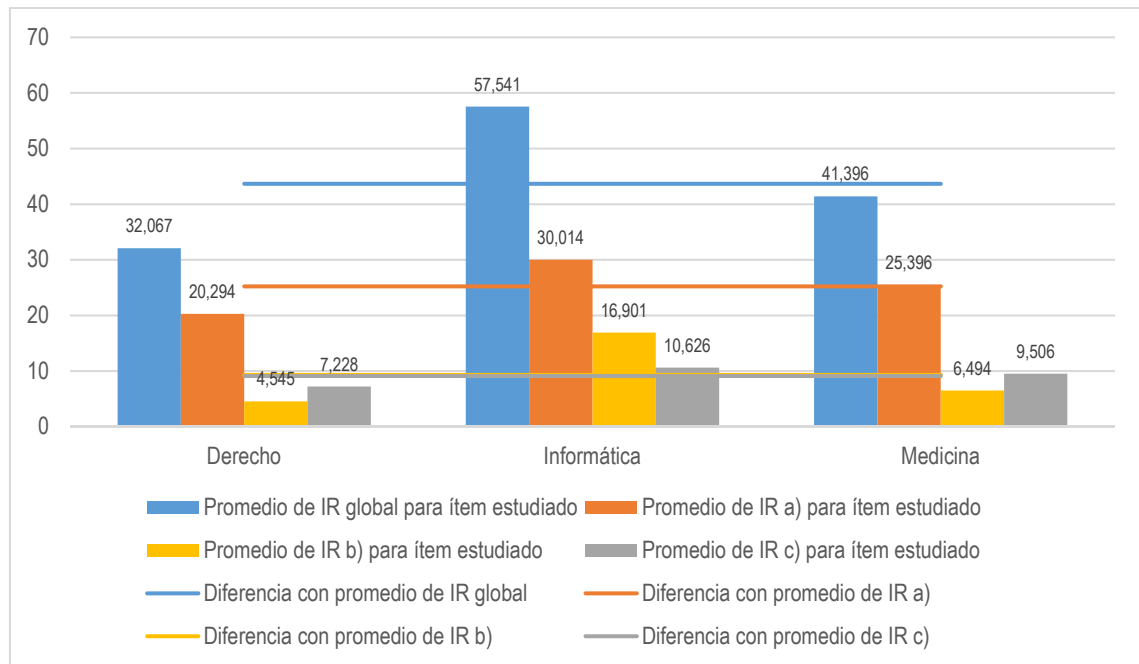


Gráfico 83: Promedio de IR global y tipos según tema y diferencias con promedio de IR global y tipos

En relación con los PRF vinculados a cada tema (*vid. supra* Gráfico 22; apartado IV. 1.1.8., p. 269), el Gráfico 84 sirve de comparación entre el número total de procedimientos de reformulación y el IR de cada subcorpus, para lo cual se ha normalizado cada conjunto de valores (*vid. supra* apartado III. 5.3., p. 206, para conocer el procedimiento de normalización), ya que al contrastar los datos del Gráfico 83 y el Gráfico 22 se observa una diferencia notable en la informática, como se puede comprobar en el Gráfico 84.

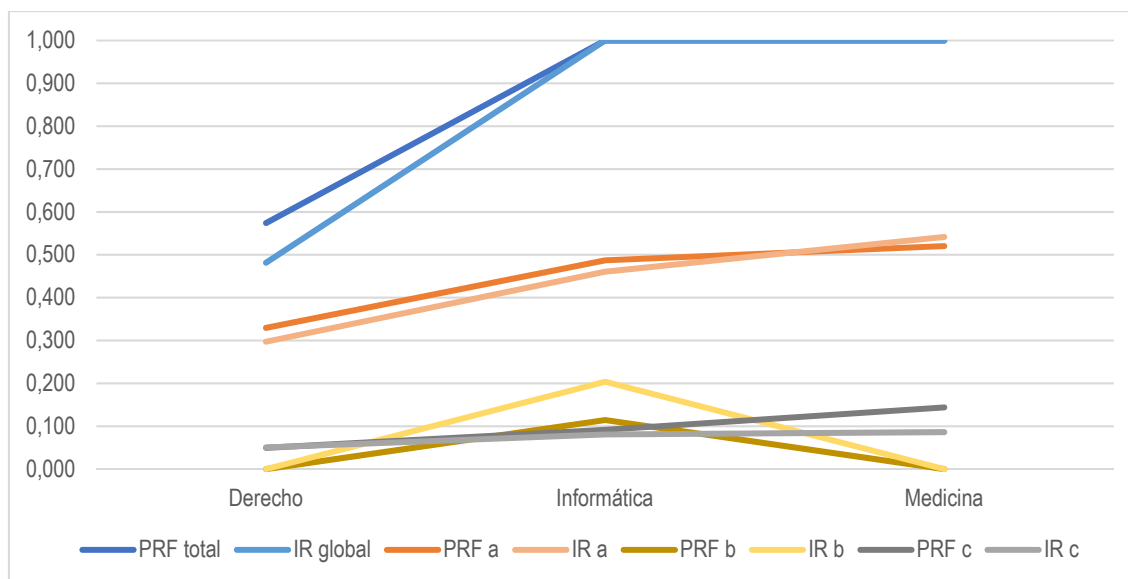


Gráfico 84: Superposición de datos normalizados de tipos de IR y tipos de PRF según tema

En efecto, los PRF y el IR son proporcionales y coinciden, en mayor o menor medida, en todos los temas. No obstante, es posible hallar que el IR por reducción de la informática es proporcionalmente mayor que el conjunto de PRF por reducción en el mismo tema, lo cual revela que la terminología de los textos sobre informática sufre mayor número de procedimientos de reformulación por reducción, sobre todo, si se compara con la medicina o el derecho, que se encuentran muy por debajo de la media del IR por reducción.

De la misma forma, se advierte una diferencia entre los PRF totales y el IR en derecho, que puede interpretarse de forma contraria a la informática: los textos en derecho no solo cuentan con menor cantidad de procedimientos de reformulación, sino que, proporcionalmente a la terminología del subcorpus, son textos que sufren menor reformulación en general.

A pesar de que cada tema parece presentar un patrón de distribución más o menos homogéneo, no puede considerarse que el IR pueda contribuir a la discriminación de los temas de un texto (al igual que tampoco lo hace el IDT, *vid. supra* IV. 1.4.1., p. 308).

1.5.2. Relación con la CLASE

Como en el caso de la densidad terminológica, existe una relación directa entre la clase según el grado de especialización textual y la densidad reformulativa de un texto,

como se comprueba en el Gráfico 85 (*vid. infra*), especialmente si se observa la lineal¹⁶³ para IR global y para IR a).

Así, puede afirmarse que la densidad reformulativa de un texto permite discriminar el grado de especialización textual, ya que cuanto mayor reformulación asociada a la terminología existe en un texto, mayor es el grado de especialización de este. Esto se aplica específicamente al IR global, al IR por expansión (IR a)) y al IR por variación (IR c)) pero no al IR por reducción, que es inferior en el caso del texto especializado y del texto no especializado. Asimismo, igual que ocurre con el IDT global, el IR global es superior a la media en el texto especializado (63,17), pero no en el semiespecializado (39,588) o el no especializado (26,945). Esta diferencia también está presente en el IR por expansión y el IR por variación.

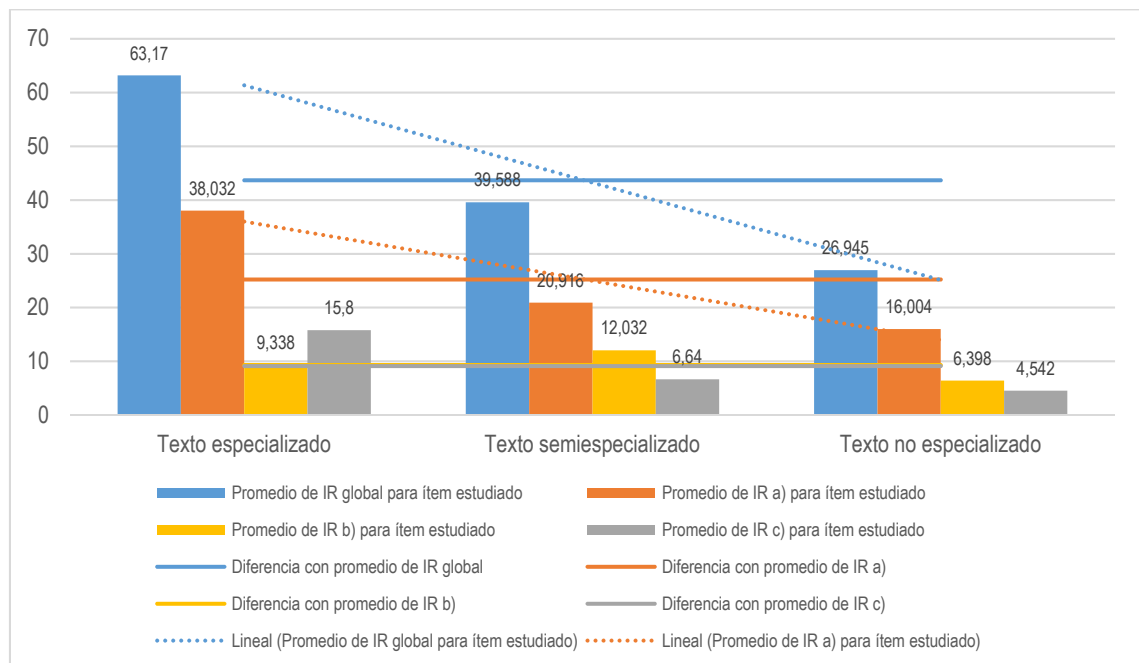


Gráfico 85: Promedio de IR global y tipos según clase y diferencias con promedio de IR global y tipos

1.5.3. Relación con el TIPO DE TEXTO

Tres son los tipos textuales que se hallan por encima del promedio del IR global, IR por expansión e IR por variación, en orden de mayor a menor: el artículo en revista científica, la tesis doctoral y la entrevista. Solo dos de estos textos coinciden con el caso del IDT (*vid. supra* Gráfico 68; apartado IV. 1.4.3., p. 311): la tesis doctoral y el artículo en revista científica, los cuales son los textos con mayor densidad terminológica del corpus.

¹⁶³ Una línea de tendencia lineal muestra el aumento o reducción constante a través del tiempo o de una variable ordinal.

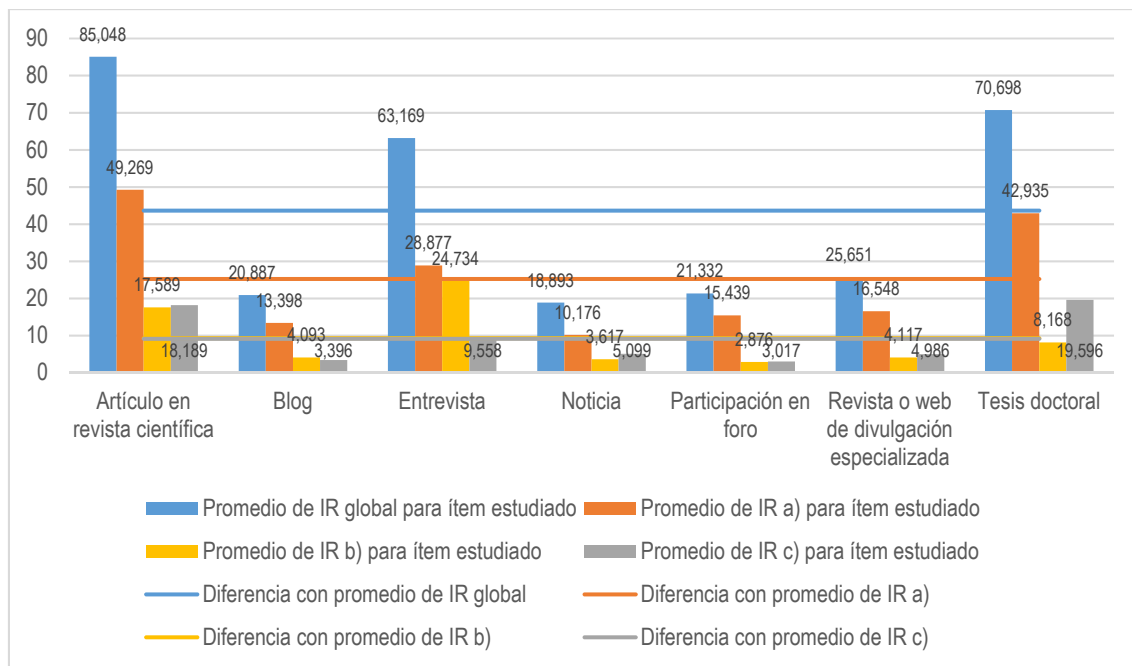


Gráfico 86: Promedio de IR global y tipos según tipo de texto y diferencias con promedio de IR global y tipos

La distribución y proporción de los distintos tipos de IR en el resto de los tipos textuales (blog, noticia, participación en foro y revista o web de divulgación especializada) es similar y las diferencias entre tipos textuales y entre tipos de IR no es significativa en ninguno de los casos, por lo que en este caso presentan también más similitudes que diferencias. En contraste con los datos de IDT (*vid. supra* Gráfico 68; apartado IV. 1.4.3., p. 311), debe destacarse que el blog y la revista o web de divulgación especializada son textos que, pese a presentar una densidad terminológica similar al del artículo en revista científica, cuentan con terminología que se reformula poco, lo cual puede llegar a ser llamativo si tenemos en cuenta que se trata de tipos textuales que, aunque sirven como medio de divulgación de masas, están restringidos a un público concreto interesado en ciertos temas.

Si atendemos a los tres temas del corpus según los diferentes tipos textuales, puede comprobarse que los datos que se relacionan entre el Gráfico 87 y el Gráfico 93 (*vid. infra*) son muy similares a los que presentan los tipos de PRF según temas para cada tipo textual (*vid. supra* Gráfico 59 a Gráfico 65, p. 305). De hecho, las diferencias que se identificaban entre los temas y entre los tipos textuales coinciden con las que se presentan para la densidad reformulativa.

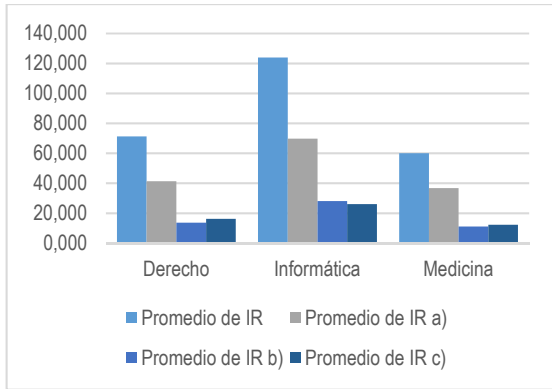


Gráfico 87: Tipos de IR según temas para el artículo en revista científica

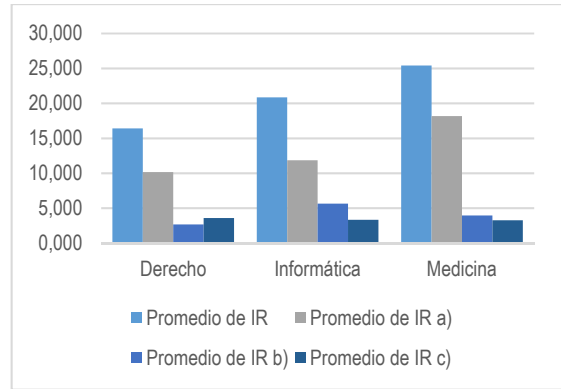


Gráfico 88: Tipos de IR según temas para el blog

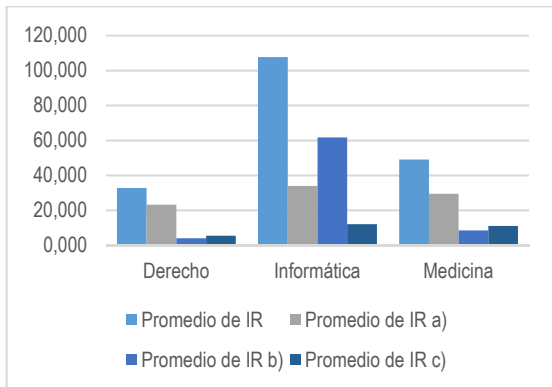


Gráfico 89: Tipos de IR según temas para la entrevista

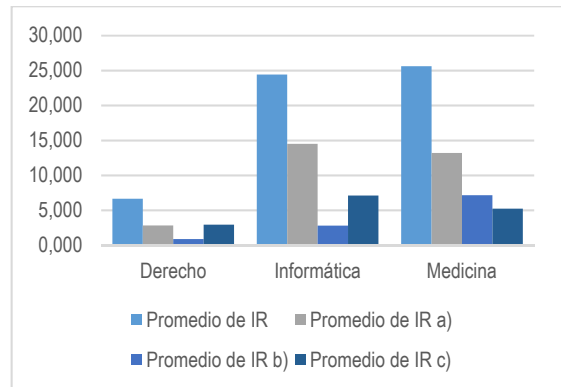


Gráfico 90: Tipos de IR según temas para la noticia

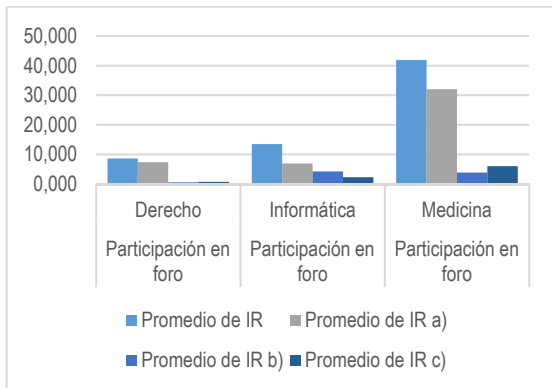


Gráfico 91: Tipos de IR según temas para la participación en foro

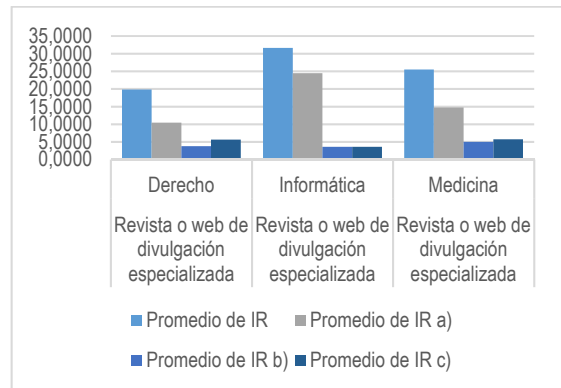


Gráfico 92: Tipos de IR según temas para la revista o web de divulgación especializada

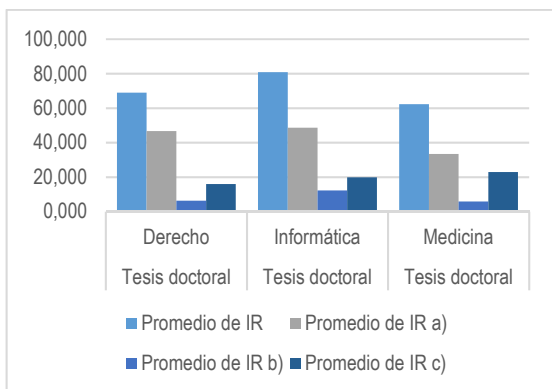


Gráfico 93: Tipos de IR según temas para la tesis doctoral

1.5.4. Relación con el EMISOR y el SUBEMISOR

En contraste con el IDT (*vid. supra* Gráfico 76; apartado IV. 1.4.4., p. 314), el emisor que se encuentra por encima del promedio de la reformulación global es exclusivamente el especialista investigador (IR = 66,595) (*vid. infra* Gráfico 94), y no los dos tipos de especialista, como sucede con el IDT global. De hecho, la reformulación global en el especialista investigador casi duplica la del especialista no investigador (IR = 38,142).

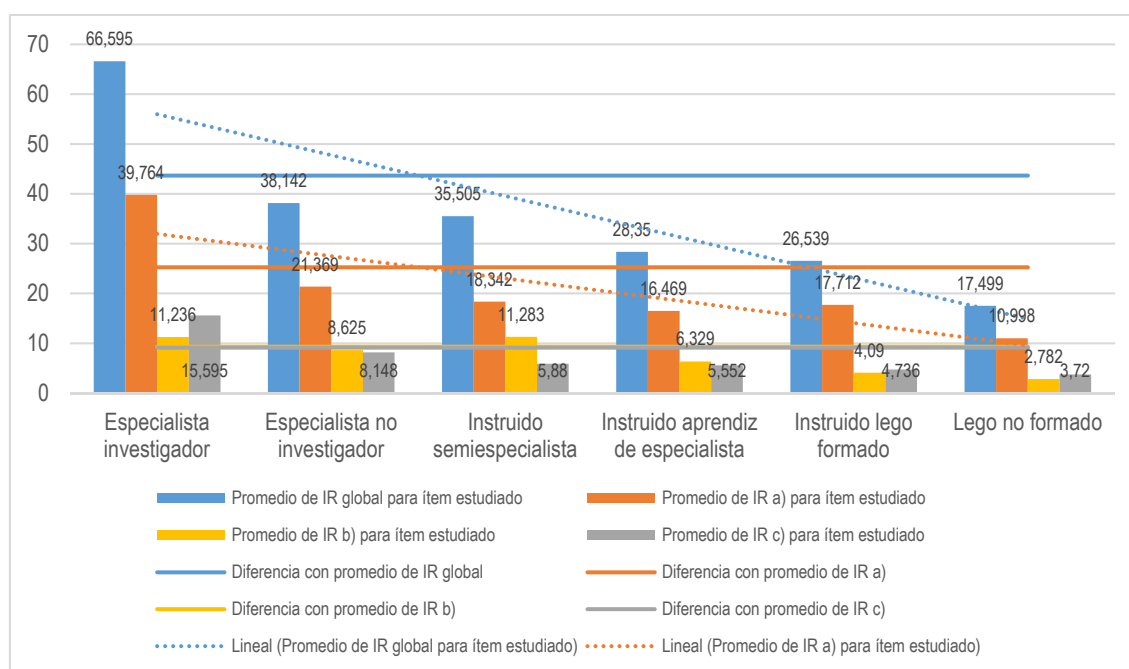


Gráfico 94: Promedio de IR global y tipos según subemisor y diferencias con promedio de IR global y tipos

El resto de subemisores cuenta con un valor que desciende gradualmente desde el usuario con mayor competencia cognitivo-comunicativa al usuario con menor competencia. Así, la relación entre el IR y la competencia del emisor es directa, como indica la lineal del Gráfico 87 para el IR global y para el IR a). A pesar de esta gradación, los valores de todos los tipos de IR de los subemisores intermedios no se diferencian lo suficiente como para contribuir a discriminar las diferentes subcategorías de emisores.

Por otro lado, siguiendo el patrón de los tipos textuales con mayor reformulación (artículo en revista científica y tesis doctoral), que, además, son los que se asocian a un texto especializado (*vid. supra* Gráfico 23; apartado IV. 1.2.1., p. 273), el subemisor especialista investigador es el que cuenta con un IR por variación (IR c)) más elevado y representativo que en el resto de subemisores.

En cuando al IR b) e IR c), los valores son tan similares en todos los subemisores, que dichas variables tampoco permiten distinguir las categorías.

Para Garaigordobil (2011a) el **ciberbullying** consiste en la utilización de las **TIC**, Internet, el teléfono móvil y los videojuegos online principalmente, para ejercer un **acoso psicológico** entre iguales, y precisa que el rápido desarrollo del mismo genera la necesidad de estudiarlo. Por último, Álvarez-García et al. (2011), con un buen análisis en el principio de su artículo, determinan que para ellos **la violencia a través de las TIC** nombre que utilizan para referirse al **ciberbullying** debido a la falta de consenso sobre el término- sería “aquella conducta intencionada con la que se causa un daño o un perjuicio a través de las **tecnologías de la información y la comunicación**, principalmente el teléfono móvil e Internet” (p. 222).

Desde la perspectiva de Calmaestra (2011) el **ciberbullying** es una nueva forma de **bullying** que implica el uso de teléfonos móviles, Internet u otras **TIC** para **acosar, amenazar o intimidar** deliberadamente a alguien. Añade que al considerarlo un tipo de **bullying** debe de mantener sus mismas características, aunque con matices que ofrecen las **TIC**, especialmente el desequilibrio de poder que se puede producir al no poder eliminar el contenido dañino por parte de la **víctima** o en el anonimato del **agresor**, y la repetición entendida tanto como la repetición del **acoso** como la lectura o visionado continuo de la **agresión**.

(DER TES14) (Estudiante)

Tabla 112: Ejemplo de reformulaciones en tesis doctoral sobre derecho con subemisor especialista investigador

¿Tiene España grandes profesionales en esta área?

[...]

Los «**hackers**» son **investigadores de seguridad**, los **ciberdelincuentes** son los que cometen delitos y los «**hacktivistas**» son los que indican ideología por medio de las técnicas de «**hacking**»

(INF ENT05) (Estudiante)

Tabla 113: Ejemplo de reformulaciones en entrevista sobre informática con subdestinatario lego no formado

1.5.5. Relación con el DESTINATARIO y el SUBDESTINATARIO

Medir el comportamiento de la reformulación según el destinatario es complejo (*vid. infra* Gráfico 95), puesto que la relación directa no es tan clara como en casos anteriores (por ejemplo, en el caso del IDT y el subdestinatario, *vid. supra* Gráfico 77; apartado IV. 1.4.5., p. 316).

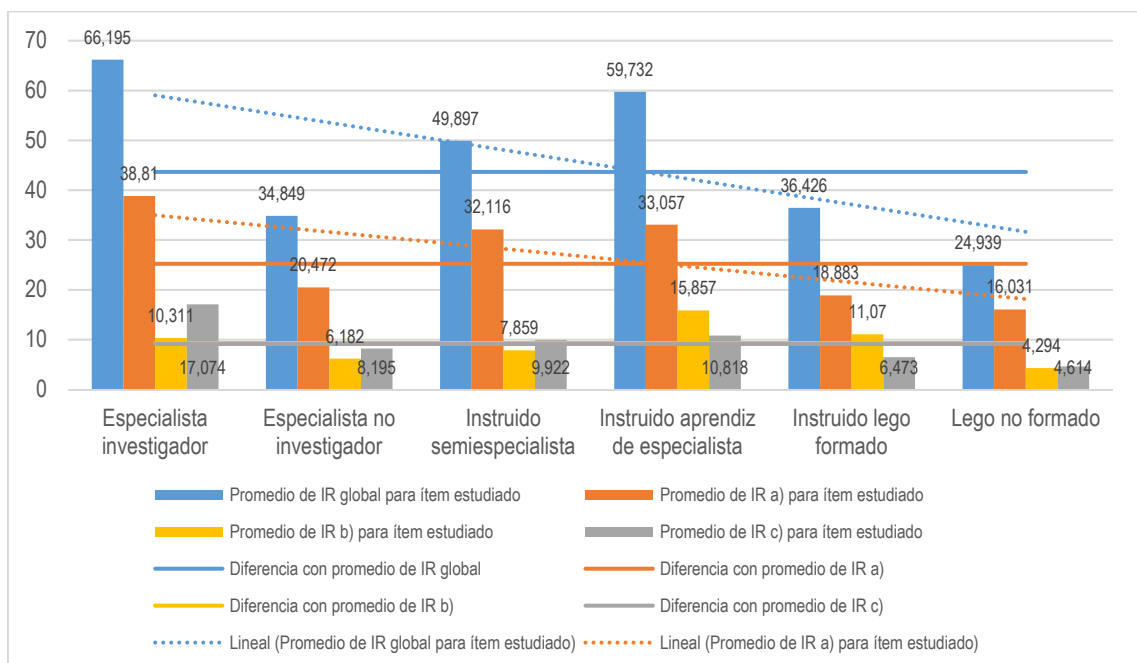


Gráfico 95: Promedio de IR global y tipos según subdestinatario y diferencias con promedio de IR global y tipos

A pesar de que la lineal para IR global y para IR a) desciende a medida que lo hace la competencia cognitivo-comunicativa de los subdestinatarios, se advierte que la reformulación decrece en el subdestinatario especialista no investigador (solo 879 PRF, el valor mínimo del total de PRF según los subdestinatarios) y, de nuevo, en el instruido lego formado (1634 PRF, el segundo valor máximo de PRF según subdestinatarios, por detrás del especialista investigador). En un principio, podría considerarse que la falta de conocimiento sobre ciertos temas implica el aumento de la reformulación de las formas léxicas especializadas atendiendo a cada destinatario, por eso la reformulación no decrece, sino que aumenta entre el especialista no investigador y el instruido aprendiz de especialista. Dado que a partir de este subdestinatario el valor de IR decrece de nuevo hasta alcanzar el valor mínimo en el lego no formado, no puede asociarse una relación inversa entre IR y perfil cognitivo-comunicativo del destinatario.

Dado que los dos últimos subdestinatarios coinciden con los subdestinatarios por debajo del promedio de IDT (*vid. supra* Gráfico 77; apartado IV. 1.4.5., p. 316), podría tenerse en cuenta la reducida densidad terminológica asociada a un número escaso de procedimientos de reformulación formal, si bien esto no se aplica al subdestinatario instruido lego formado, cuyos PRF alcanzan 1634 procedimientos (sí al lego no formado, que cuenta solo con 901 PRF). De hecho, puede considerarse que el caso del subdestinatario especialista no investigador y el subdestinatario instruido lego formado son casos inversos, ya que su similar IR se debe a que el primero cuenta con un IDT muy elevado pero un número de PRF escaso; y el segundo cuenta con un IDT reducido pero una gran cantidad de PRF (no comparable a la del especialista investigador, que alcanza los 4713 procedimientos).

El caso del IR a) sí se aproxima más a la mencionada relación directa entre reformulación y perfil del destinatario, puesto que, a pesar del valor del subdestinatario especialista no investigador, el resto de los valores para los subdestinatarios se mantiene similar en el caso del semiespecialista y del aprendiz de especialista, y desciende en el lego formado, que comparte un valor próximo con el lego no formado. Así, sí puede defenderse el papel que desempeña la reformulación por expansión según el destinatario, que aumenta a medida que aumenta el perfil cognitivo-comunicativo del destinatario (*vid. infra* Tabla 114).

Desde el punto de vista del **Derecho penal** y **penitenciario** – y con independencia de muchas otras cuestiones problemáticas que suscita el paradigma de la prevención en el que se inserta esta preocupación por la **peligrosidad** de los **delincuentes** (cfr., muy críticamente, Harcourt, 2007 o Albrecht, 2004) – una de las cuestiones que más interesa acerca de estos **instrumentos** de **valoración** del **riesgo** de **violencia** es la

capacidad predictiva que tienen, esto es, con qué grado de acierto efectúan los pronósticos sobre la conducta futura de las personas. En otras palabras, qué probabilidad hay de equivocarse cuando los operadores jurídicos toman decisiones basadas en estos pronósticos, ya sea por haber considerado de alto riesgo a un sujeto que en realidad luego no vuelve a delinquir (falsos positivos), o por haber considerado de riesgo bajo a una persona que sin embargo más tarde sí comete nuevos delitos violentos (falsos negativos).

(DER ART09) (Estudiante)

Tabla 114: Ejemplo de reformulaciones por expansión en artículo en revista científica sobre derecho con subdestinatario especialista investigador

El caso de la reformulación por reducción no es relevante según el subdestinatario, puesto que no existe una relación clara y, además, las diferencias tampoco son suficientes para establecer conclusiones relevantes. No obstante, a pesar de que en el caso del IR por variación las diferencias tampoco son muy acusadas, debe mencionarse que destaca, principalmente, en los casos especialista investigador (*vid. infra* Tabla 115) a aprendiz de especialista, y resulta reducido en el lego formado y el lego no formado, lo que revela que los procedimientos de variación denominativa o acronimia, que son los principalmente identificados entre los procedimientos por variación, no son frecuentes en los textos que recibe un usuario con una competencia cognitivo-comunicativa menor.

Se proponen las PBN como metodología para modelar un proceso de fabricación de ensamblaje automatizado, utilizándolas para generar datos cuantitativos que contribuyan en la evaluación de la incidencia de fallas en un Análisis de Modo de Falla y Efectos para Diseño (DFMEA, por sus siglas en inglés).

[...]

DMAIC son las siglas en inglés para Definir, Medir, Analizar, Mejorar y Controlar.

[...]

Este paquete de software fue diseñado por un grupo de investigadores de la Universidad de Birmingham (Reino Unido), para la construcción y el análisis formal de una amplia gama de propiedades cuantitativas de varios tipos de modelos probabilísticos, incluyendo Cadenas de Markov de tiempo discreto (DTMC, por sus siglas en inglés) y continuo (CTMC, por sus siglas en inglés), Procesos de Decisión de Markov (MDP, por sus siglas en inglés) y Autómatas de Tiempo Probabilísticos (PTA, por sus siglas en inglés).

(INF TES14) (Estudiante)

Tabla 115: Ejemplo de reformulaciones por variación en tesis doctoral sobre informática con subdestinatario especialista investigador

1.5.6. Relación SUBEMISOR-CLASE-IR

El Gráfico 96 muestra la relación de los distintos IR entre los seis tipos de subemisores con cada una de las clases textuales según el grado de especialización.

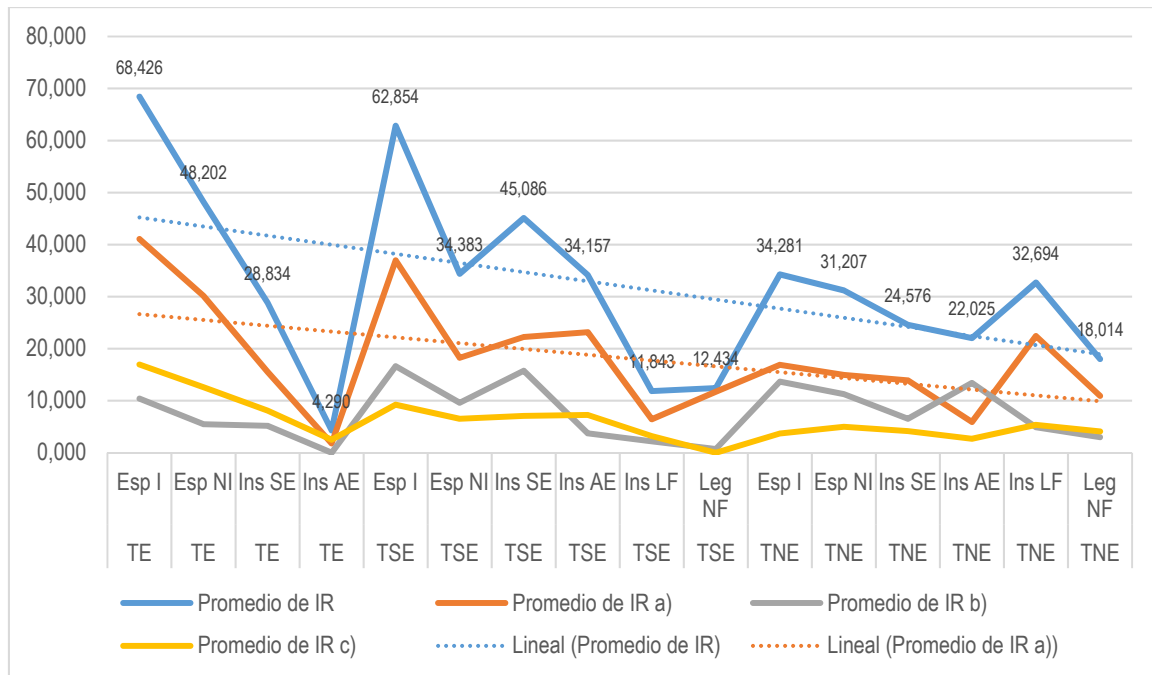


Gráfico 96: Relación de datos globales para la correspondencia subemisor-clase según IR global y tipos

Así, se identifican tres grupos de subemisores.

- a) En el primer bloque se incluyen el subemisor especialista investigador y el especialista no investigador. Se observa que en los dos casos el IR global disminuye a medida que decrece el grado de especialización de los textos. Esto también sucede en el IR por expansión, que parece ser una réplica del IR global, pero no en el caso del IR por reducción o el IR por variación.
- b) En el segundo bloque se incluyen el subemisor instruido semiespecialista y el subemisor instruido aprendiz de especialista. En este caso, la reformulación es mayor en el texto semiespecializado (recuérdese la estrecha relación entre el subemisor semiespecialista y el texto semiespecializado (*vid. supra* Gráfico 24, p. 278). En el caso del aprendiz de especialista, la relevancia del texto semiespecializado queda reducida al restringido número de patrones que conformen dicho fenómeno, ya que no suponen más del 1,35 % del corpus. Tanto el IR global como el IR por expansión se comportan de forma similar.
- c) El último bloque lo conforman los dos subemisores restantes, el instruido lego formado y el lego no formado, que cuentan con mayor reformulación en el texto no especializado.

De esta forma, cabe defender que el IR se vincula no solo con la clase, sino también con el emisor del texto, ya que la reformulación aumenta cuando la conjunción de emisor

y clase aumentan hasta el máximo grado y disminuye a medida que ambos lo hacen, como sintetiza la siguiente tabla.

Subemisor	Clase	IR
Especialista investigador	Texto especializado	máx.
Especialista no investigador		
Instruido semiespecialista	Texto semiespecializado	↓
Instruido aprendiz de especialista		
Instruido lego formado	Texto no especializado	
Lego no formado		

Tabla 116: Relación entre subemisor, clase e IR

1.5.6.1. Relación SUBEMISOR-CLASE-IR-IDT

Con el objetivo de comprobar cómo se relacionan el IDT y el IR según la clase y el subemisor de cada texto, se han diseñado diferentes gráficos que conjugan dichos valores. Para poder representar visualmente en un gráfico de líneas la relación entre IDT e IR, se ha procedido a normalizar sus índices, ya que cada índice cuenta con escalas diferentes: el IDT cuenta con valores decimales de 0 a 1 y el IR, con valores enteros, a partir de 0. Para ello, a cada conjunto de datos se le ha aplicado la siguiente fórmula: $x^1 = \frac{x-x_{\min}}{x_{\max}-x_{\min}}$. Así, para obtener el valor escalado, se opera con el conjunto de valores de cada índice (es decir, la operación se realiza con todos los valores para IDT, y con todos los valores para IR, de forma separada). Se parte del valor que pretende escalarse, al que se resta el valor mínimo del conjunto. El valor obtenido es dividido entre la diferencia del valor máximo y el valor mínimo del conjunto. De esta forma, se obtiene un valor que oscila entre 0 y 1 para cada conjunto (es decir, para IDT y para IR), de forma que pueden representarse gráficamente mediante un gráfico de líneas y observar así el grado de coincidencia y diferencia de los valores.

En el Gráfico 97, se comprueba que, en general, el IDT y el IR disminuyen de forma pareja conforme disminuye el grado de especialización, así como a la competencia cognitivo-comunicativa del emisor. Por una parte, en la clase de texto especializado prima el IDT sobre el IR, mientras que en el texto no especializado prima la reformulación sobre la densidad terminológica. Esta relación, que se invierte en el texto semiespecializado (lo que sigue proporcionando pruebas de su complejidad como objeto de transición), no implica que los textos no especializados estén más reformulados que el resto de textos: debe tenerse en cuenta que tanto el IR como el IDT son más reducidos en los subemisores

con menor competencia, especialmente en los textos con menor grado de especialización (*vid. supra* Gráfico 76 y Gráfico 94, pp. 314 y 330).

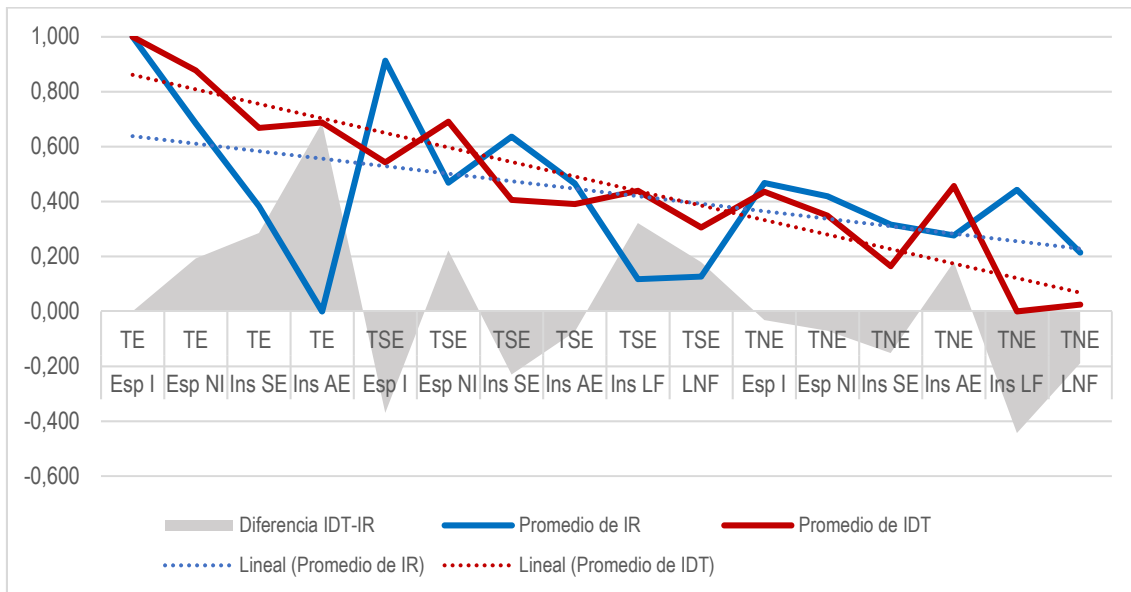


Gráfico 97: Superposición de datos normalizados de IDT global e IR global para subemisor y clase

En este Gráfico 97 distinguimos tres comportamientos de la relación IDT-IR, que coinciden, en mayor o menor medida, con las clases según el grado de especialización:

- a) En el texto especializado la disminución gradual y pareja entre IDT e IR tiene lugar hasta el emisor instruido aprendiz de especialista, que cuenta con una reformulación prácticamente nula¹⁶⁴. En esta clase prima la densidad terminológica sobre la reformulación que afecta a dicha terminología.
- b) En el texto semiespecializado, las relaciones IDT-IR se establecen de forma alterna y errática entre los diferentes subemisores.
- c) En el texto no especializado la reformulación se vuelve un factor más relevante que la densidad terminológica, a excepción del caso con emisor instruido aprendiz de especialista¹⁶⁵.

A continuación, se presentan tres gráficos (Gráfico 98 a Gráfico 100) que conjugan el IR con los diferentes tipos de IDT.

¹⁶⁴ En efecto, los textos especializados emitidos por un instruido aprendiz de especialista son solo dos en todo el corpus (dos participaciones en foro sobre informática: INF FOR03 e INF FOR04), que suman, a su vez, tan solo dos PRF en total, por lo que la diferencia entre IDT e IR es notable. En este caso. Dichos patrones no son significativos en el análisis.

¹⁶⁵ En este caso también puede hablar de ocho patrones poco significativos sobre el total del corpus.

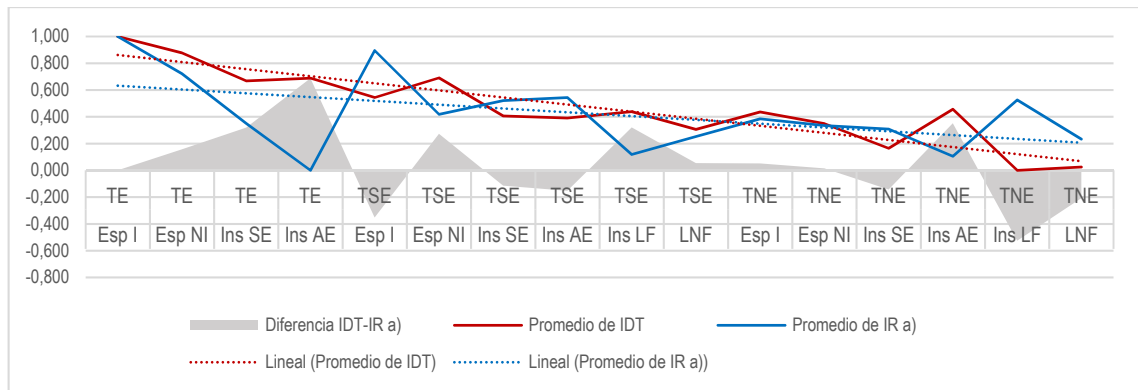


Gráfico 98: Superposición de datos normalizados de IDT global e IR a) para subemisor y clase

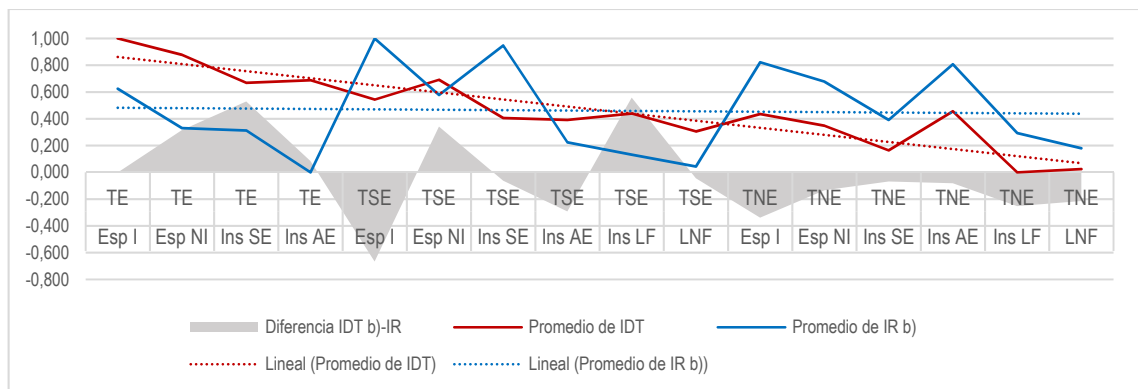


Gráfico 99: Superposición de datos normalizados de IDT global e IR b) para subemisor y clase

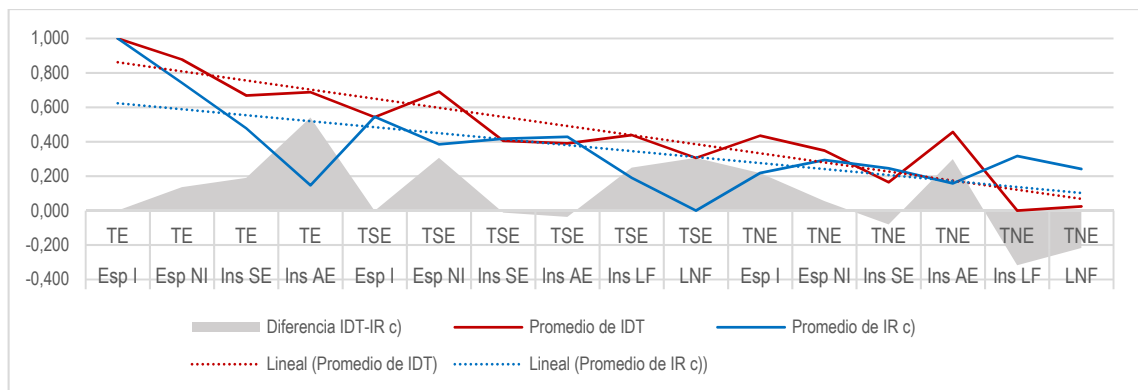


Gráfico 100: Superposición de datos normalizados de IDT global e IR c) para subemisor y clase

En los siguientes puntos se relacionan las características más representativas de dichos gráficos en cuanto al IDT en relación con los diferentes tipos de IR:

- a) En el Gráfico 98, se comprueba una ligera diferencia entre IDT e IR por expansión. Ambos valores disminuyen con la clase y la competencia del emisor, pero se invierten especialmente en los textos no especializados emitidos por lego formado y lego no formado. La relación en el texto semiespecializado y los emisores con mayor competencia del texto no especializado presentan una distribución más impredecible.

- b) En el Gráfico 99, la reformulación por reducción predomina por encima de la relevancia que tiene la densidad terminológica en el texto no especializado, al contrario de lo que sucede en el texto especializado. El texto semiespecializado, de nuevo, es el que cuenta con más alternancias IDT-IR. En este caso, la lineal del IR b) es la única de los tres gráficos en permanecer más o menos estable a lo largo de las clases y los subemisores y no disminuir con ellos.
- c) Por último, el Gráfico 100 muestra una tendencia clara, tanto del IDT como del IR, a disminuir conforme disminuyen la clase según el grado de especialización y la competencia cognitivo-comunicativa del emisor. Además, la reformulación (en este caso, por variación) es también menor que la densidad terminológica, generalmente, en todos los casos excepto en el instruido lego formado y lego no formado. En efecto, de esta forma puede afirmarse que los procedimientos de presentación de variantes denominativas y formas acronímicas van ligadas a la densidad terminológica del texto (*vid. supra* Tabla 115, p. 333).

1.5.7. Relación SUBDESTINATARIO-CLASE-IR

Como en el caso anterior, el Gráfico 101 muestra la correspondencia entre los distintos IR entre los seis tipos de subdestinatarios con cada una de las clases textuales según el grado de especialización. En este gráfico se observa que la distribución de todos los tipos de IR es similar, salvaguardando las diferencias que puedan existir, especialmente para el caso IR por expansión, que ya se vinculó estrechamente con el IR global. Además, consecuentemente, se mantienen las relaciones complejas que se avanzaron en el apartado IV. 1.5.5. (p. 331).

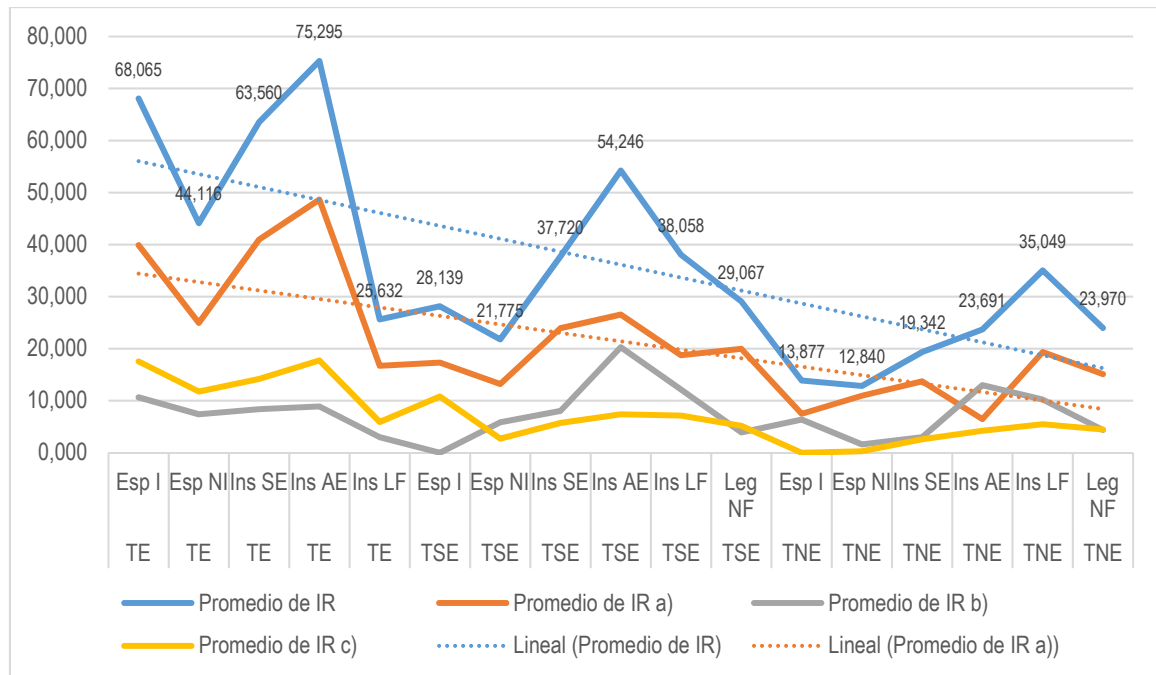


Gráfico 101: Relación de datos globales para la correspondencia subdestinatario-clase según IR global y tipos

A pesar de que se establece una relación directa para el IR entre la clase y el destinatario de un texto, ya que la densidad reformulativa decrece conforme disminuye el grado de especialización y la competencia del destinatario, se identifican diferencias con el caso del emisor (*vid. supra* Gráfico 96, p. 334). Tampoco se pueden establecer bloques que aúnen similitudes, como en el caso del emisor, puesto que la distribución del promedio de IR global e IR por expansión no sigue un patrón reconocible. Aun así, por cada clase, se distingue que el destinatario determina la reformulación de un texto cuando se asocia a la clase:

- a) En el texto especializado, la reformulación disminuye cuando el destinatario es un especialista no investigador o un lego no formado. Así, podría afirmarse que el emisor de un texto reconfigura el discurso y aumenta la reformulación de las formas léxicas especializadas si su destinatario es alguien que cuente con cierto conocimiento sobre el tema. Con *cierto conocimiento* se establece un límite muy amplio y ambiguo de definir, puesto que puede inferirse que, como revela el Gráfico 101, el destinatario especialista no investigador cuenta con conocimiento suficiente para comprender la terminología y, en el caso del instruido lego formado, que la reformulación no es una vía directa para la comprensión de la terminología. Posiblemente esto se deba a la clasificación de texto especializado, en el que prima la representación del conocimiento sobre la propia función comunicativa.

- b) En el texto semiespecializado, la reformulación está asociada al instruido aprendiz de especialista, junto con el instruido semiespecialista y lego formado. La justificación de la reformulación se repite en este caso para el texto semiespecializado.
- c) En el texto no especializado, la relación es más clara, puesto que la reformulación crece a medida que disminuye la competencia del destinatario. En contraste con los casos anteriores, en el texto no especializado se comprueba que los procedimientos de reformulación se asocian de forma frecuente a la terminología, puesto que puede interpretarse que el destinatario no cuenta con la información conceptual o denominativa necesaria. No obstante, la curva decrece hacia el lego no formado, en lugar de aumentar, como sería lógico con esta interpretación. En este caso podría deberse al IDT de términos délficos, que en el caso del destinatario lego no formado se halla por encima de la media (no en los casos de instruido semiespecialista, instruido aprendiz de especialista o instruido lego formado) (*vid. supra* Gráfico 77; apartado IV. 1.4.5., p. 315) y este tipo de formas léxicas especializadas son las más frecuentes y comprensibles para el grueso de la población.

1.5.7.1. Relación SUBDESTINATARIO-CLASE-IR-IDT

Como en el caso de la relación subemisor-clase-IR-IDT (*vid. supra* IV. 1.5.6.1., p. 335), el Gráfico 102 (*vid. infra*), presenta los valores normalizados de IDT e IR para subdestinatario y clase.

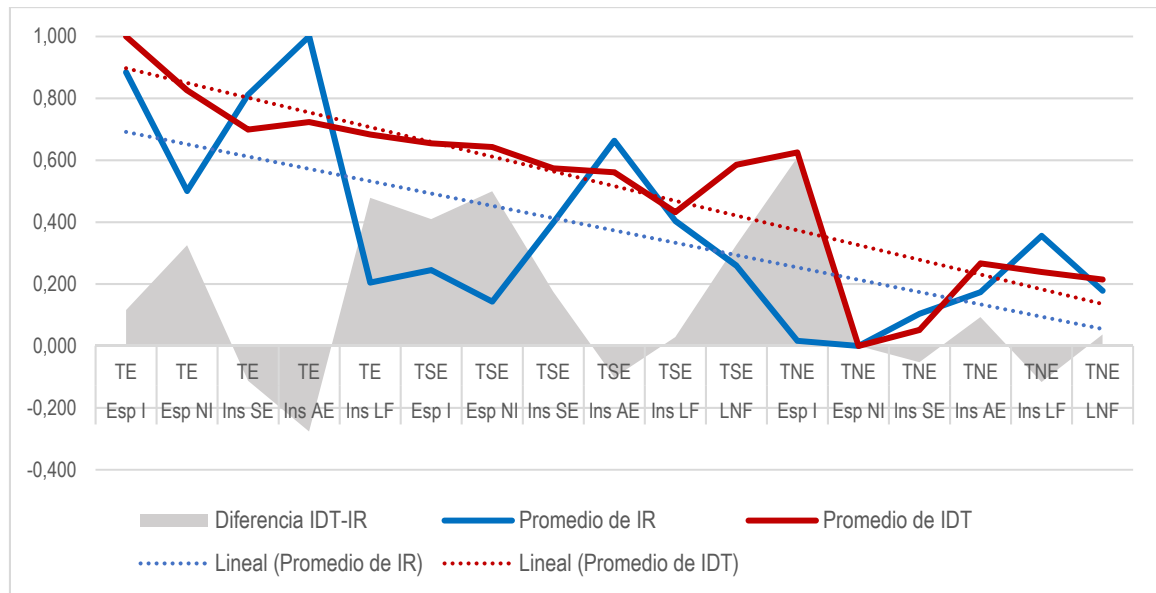


Gráfico 102: Superposición de datos normalizados de IDT global e IR global para subdestinatario y clase

Lo primero que cabe mencionar respecto a este cruce de datos es que, atendiendo al subdestinatario, el IDT prima sobre la reformulación en prácticamente todo el corpus y la tendencia de esta relación no se invierte en el texto no especializado, como sí lo hacía en el caso del subemisor (*vid. supra* Gráfico 97, p. 336). Además, tanto IDT como IR mantienen una relación directa con la clase y la competencia cognitivo-comunicativa del subdestinatario (*vid. supra* apartados IV. 1.4.5. y IV. 1.5.5., pp. 315 y 331). Las mayores diferencias entre IDT e IR se establecen en tres puntos:

- a) En el texto especializado, a favor de la reformulación en el subdestinatario instruido semiespecialista y el instruido aprendiz de especialista.
- b) En el texto semiespecializado, a favor de la densidad terminológica en los casos de subdestinatario especialista e instruido semiespecialista.
- c) En el texto no especializado, a favor de la densidad terminológica en el caso del especialista investigador.

De hecho, cruzar los datos para IDT e IR permite contar con mayor respaldo al argumento que se esgrima en el apartado anterior (*vid. supra* IV. 1.5.7.), ya que se observa que el destinatario condiciona el uso de la reformulación sobre la terminología pero estrechamente asociado a la clase. Con todo, parece ser el subdestinatario instruido el que condiciona el uso de procedimientos de reformulación; ni siquiera el lego no formado, que, *a priori*, parecería ser el subdestinatario que más necesite de la reformulación para acceder al sistema conceptual de un texto). En efecto, tanto en el caso del subemisor (*vid. supra* Gráfico 97, p. 336) como en el caso del subdestinatario (*vid. supra* Gráfico 102) la densidad

reformulativa asociada a la terminología decrece en el usuario lego no formado, por lo que no es posible defender que cuanto mayor sea la falta de conocimiento de un hablante sobre un tema, mayor será la reformulación a la que esté sujeto el discurso sobre dicho tema.

En los siguientes casos (Gráfico 103 a Gráfico 105), en los que se compara la relación entre el IR global y los diferentes IDT, se observa que la relación entre IDT e IR es pareja (a excepción del Gráfico 104, cuya relación comienza a invertirse en el texto no especializado).

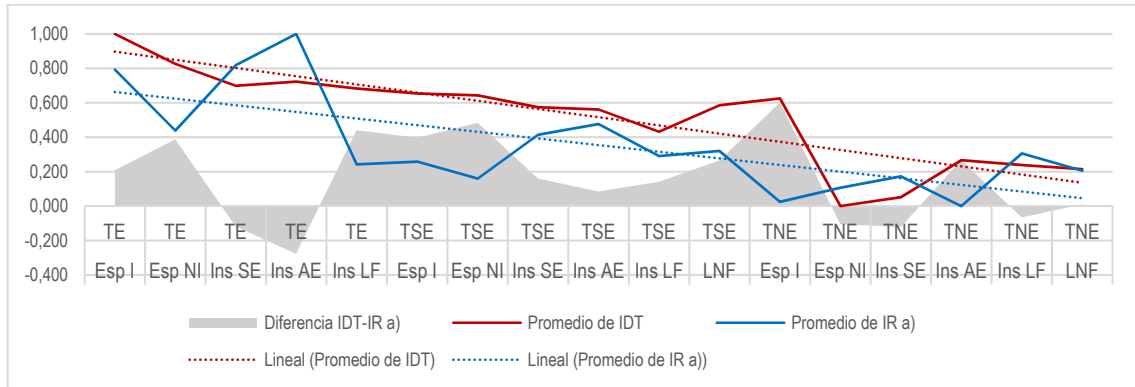


Gráfico 103: Superposición de datos normalizados de IDT global e IR a) para subdestinatario y clase

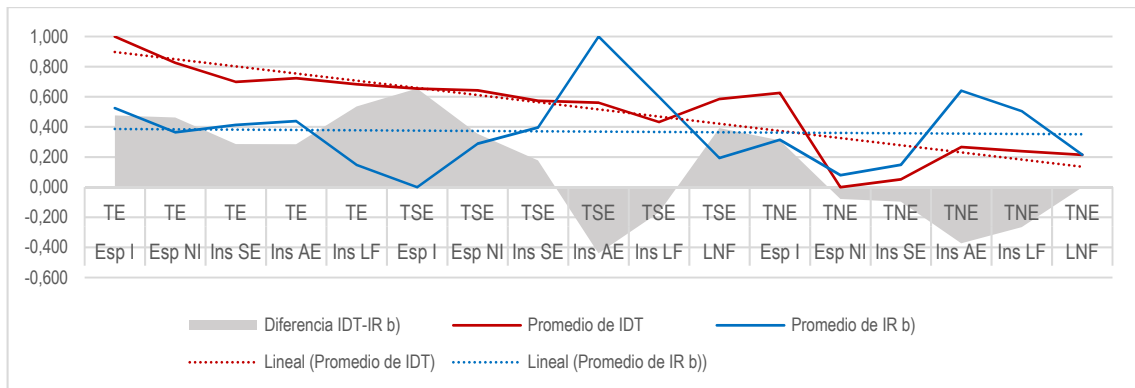


Gráfico 104: Superposición de datos normalizados de IDT global e IR b) para subdestinatario y clase

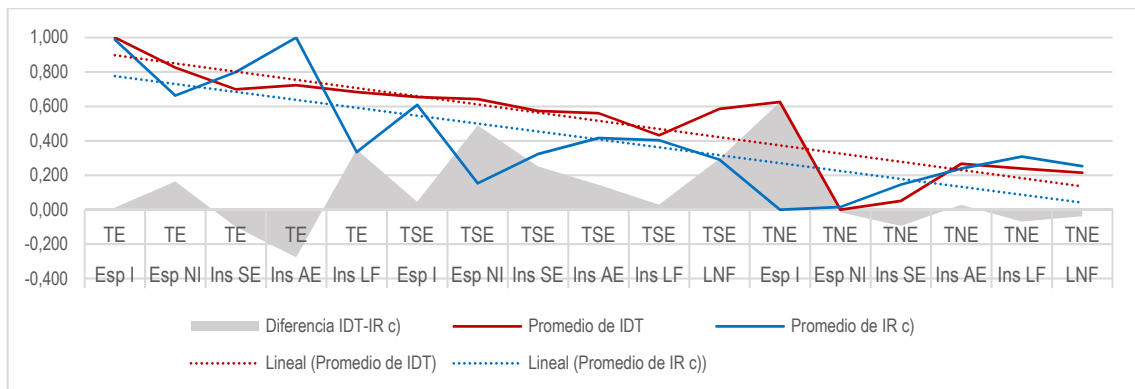


Gráfico 105: Superposición de datos normalizados de IDT global e IR c) para subdestinatario y clase

En los siguientes puntos se relacionan las características más representativas de dichos gráficos en cuanto al IR en relación con los diferentes tipos de IDT:

- a) El Gráfico 103 es muy similar al Gráfico 102. De hecho, en otras ocasiones ya se ha advertido la similitud de la distribución del IR por expansión con el IR global, lo que tiene sentido si se tiene en cuenta que el número total de PRF por expansión del corpus (6312 PRF a)) (*vid. supra* Gráfico 21, p. 268) constituye casi el 60 % del total de procedimientos de reformulación. El IR por expansión destaca sobre el IDT de igual forma que lo hace el IR global. De hecho, la interpretación de dicho gráfico se ajusta también a la realidad para el Gráfico 102 (p. 341).
- b) El Gráfico 104 destaca especialmente en los destinatarios instruidos del texto semiespecializado y en el destinatario instruido del texto no especializado. Si filtramos aquellos textos que cumplen estas características y que superan la media de IR por reducción, se observa que, aunque se incluyen los siete tipos textuales del corpus, tres son los que ostentan el 70 % de la representación: el blog, la entrevista y la revista o web de divulgación especializada. Esta representación resulta coherente si tenemos en cuenta que estos son los tipos textuales que destacan para el destinatario aprendiz de especialista y lego formado (*vid. supra* Gráfico 39, p. 291) y son, de hecho, los que se comparan con el destinatario lego no formado (*vid. supra* Tabla 94) y quedan vinculados a una densidad terminológica más elevada. Si se investiga la cantidad de PRF por reducción con los que cuentan estos tipos textuales, se halla que la entrevista sobre informática destaca muy por encima de cualquier otro tema o cualquier otro tipo textual (*vid. supra* Gráfico 61, p. 305) en el número de procedimientos por reducción con los que cuentan dichos textos.
- c) En último lugar, el IR por variación en comparación con el IDT mostrado en el Gráfico 105, destaca principalmente en el artículo en revista científica (especialmente en informática; *vid. infra* Tabla 117) y en la tesis doctoral (principalmente en medicina). El caso de que la reformulación por variación sea prácticamente nula en el texto no especializado con destinatario especialista, se ajusta a los argumentos aportados anteriormente que relacionan destinatario asociado a clase con la reformulación.

Karaboga presentó en 2005 un algoritmo de enjambre de abejas para resolver problemas numéricos de optimización [7], el cual es conocido como el método “artificial bee colony” (ABC).

(INF ART11) (Estudiante)

Tabla 117: Ejemplo de reformulación por reducción en artículo en revista científica sobre informática

1.5.8. Relación SUBEMISOR-SUBDESTINATARIO-IR

Al combinar los datos de IR para subemisor y subdestinatario, se observa la relevancia con la que cuenta la competencia cognitivo-comunicativa del destinatario instruido como factor regulador de la reformulación.

En el caso del emisor especialista (*vid. infra* Gráfico 106), los destinatarios más relevantes son los instruidos. Debe resaltarse un caso que, desde nuestro punto de vista, puede considerarse especial. Se trata de la reformulación elevada que existe entre especialista investigador-especialista y, por ende, en el texto especializado (*vid. supra* Gráfico 27, p. 281), especialmente en el artículo en revista científica y la tesis doctoral (*vid. supra* Gráfico 40 y Gráfico 46, pp. 295 y 296). Esta relación puede vincularse directamente con la elevada densidad terminológica y con las propias características del discurso científico, entre las que se encuentran la repetición, la descripción y la explicación, etc.

En este caso también destaca la relevancia de la reformulación por variación de los cuatro primeros subdestinatarios para el subemisor especialista investigador, lo que también está asociado al tipo textual de artículo en revista científica y tesis doctoral.

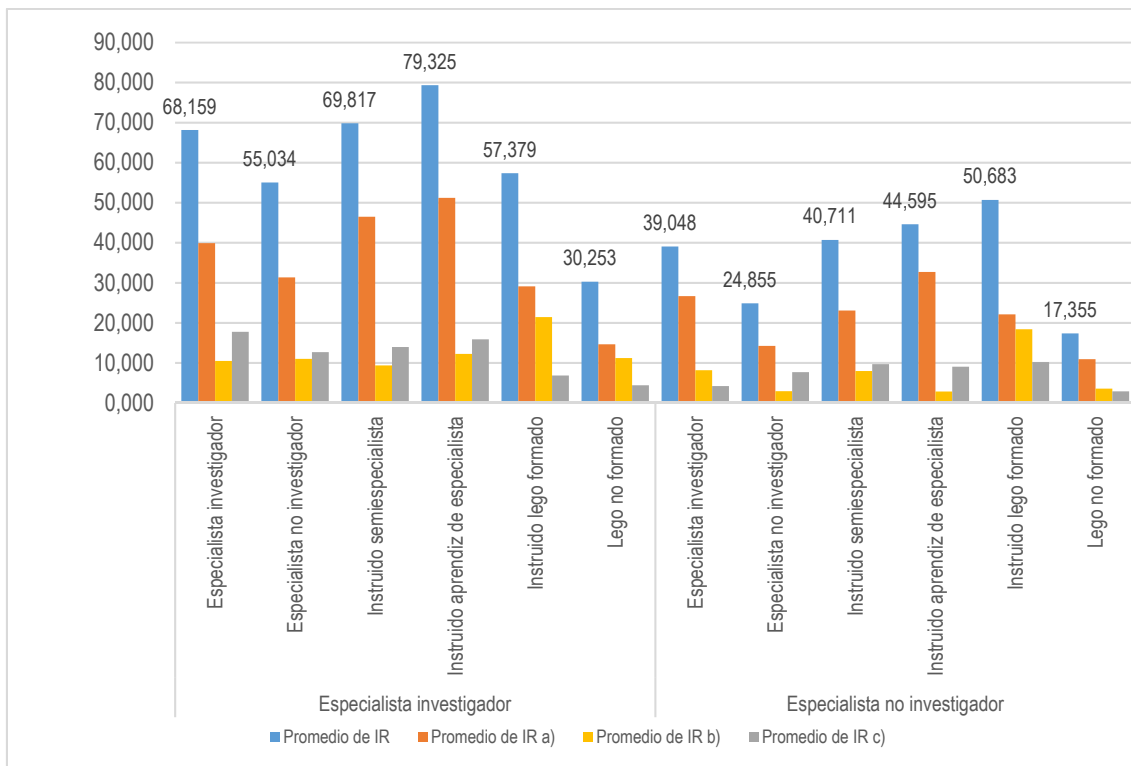


Gráfico 106: Relación de datos globales para la correspondencia emisor especialista-tipo de subdestinatario según IR global y tipos

El caso del emisor instruido (*vid. infra* Gráfico 107) es algo más complejo, puesto que las relaciones no quedan claras para todos los subdestinatarios. En los casos del subemisor semiespecialista y el subemisor aprendiz de especialista, se identifica la reformulación asociada, principalmente al subdestinatario instruido, si bien en el caso del semiespecialista-especialista investigador la reformulación sigue siendo elevada.

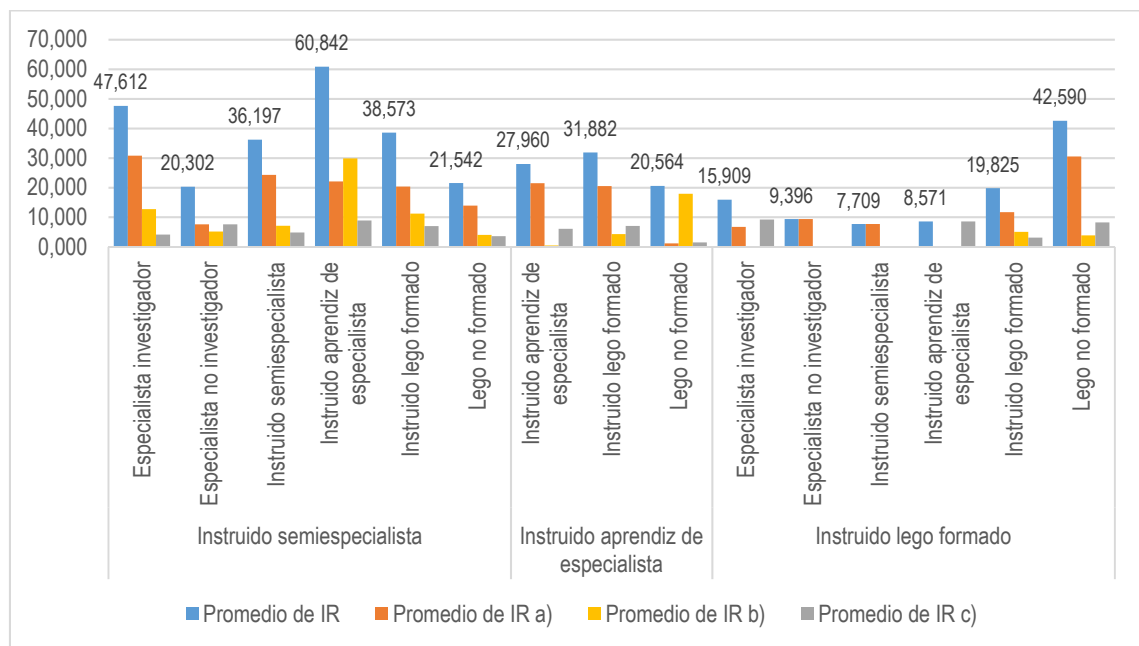


Gráfico 107: Relación de datos globales para la correspondencia emisor instruido-tipo de subdestinatario según IR global y tipos

A nuestro juicio, esto sigue vinculándose con el tipo textual, puesto que la gran mayoría de textos con dicha situación discursiva pertenecen a los tipos textuales clasificados como especializados. En contraste, la relación lego formado-lego no formado destaca especialmente, dado que dicho IR es el segundo más elevado del subcorpus subemisor instruido-tipo de subdestinatario. Estos textos corresponden, principalmente, con entrevistas o participaciones en foro no especializadas. Dado que la reformulación más elevada, de hecho, es la reformulación por expansión, puede entenderse que el objetivo explicativo (tan propio de los tipos textuales participación en foro y entrevista) hacia el lego no formado se manifiesta a través del procedimiento de término seguido de descripción o explicación.

Hola, Yo tambien sufro de Neuralgia Intercostal, segun lo que me Explico mi Medico es la secuela originada por un tratamiento de radioterapia por consecuencia de un cancer de Mamás que tuve hace dos años, es muy doloroso y los calmante a veces no hacen efecto, me gustaría escribirme con personas que sufren de la misma dolencia Gracias

(MED FOR01) (Usuario medio)

Tabla 118: Ejemplo de reformulación por expansión de un término críptico en una participación en foro sobre medicina con destinatario lego no formado

Por último, en el Gráfico 108 destaca el destinatario instruido lego formado. Todos los textos que cumplen la relación lego no formado-lego formado son participaciones en foro no especializadas donde se pone en marcha el procedimiento de término seguido de descripción (*vid. infra* Tabla 119).

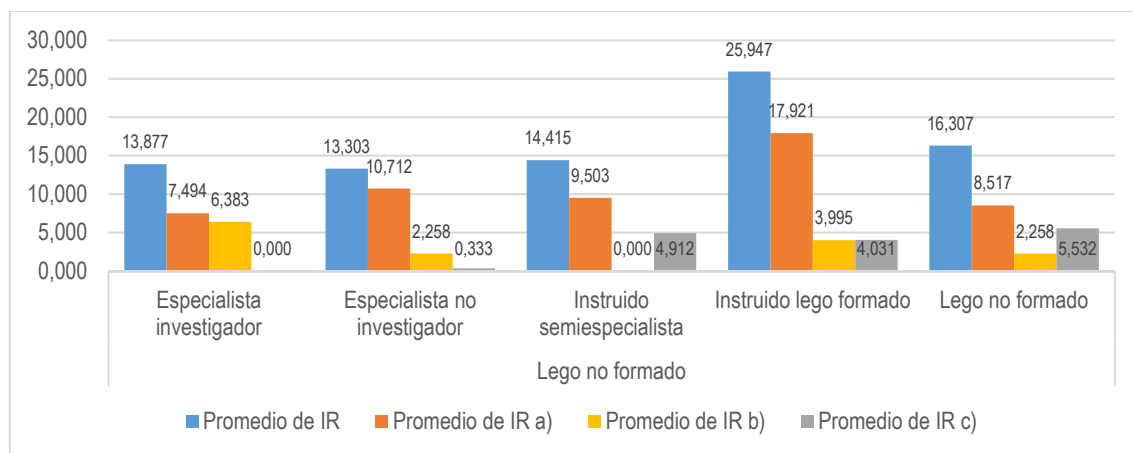


Gráfico 108: Relación de datos globales para la correspondencia emisor lego-tipo de subdestinatario según IR global y tipos

Hace 20 años sufrí un trauma craneoencefálico severo, estuve en estado de coma durante 4 días, presente pérdida temporal de memoria y en el periodo de convalecencia abusé de estimulantes (modafinil, cafeína, coca-cola, aspirinas) para mantenerme despierta en la noche, toda vez que si llegaba a dormir olvidaría lo que había estudiado.

[...]

El cansancio o agotamiento es un síntoma frecuente que toda persona tiene en algunas ocasiones ya sea por sueño, por trabajo intenso, un viaie agotador o por la práctica excesiva de deporte.

[...]

El cansancio es la sensación de falta de energía que disminuye la capacidad de trabajo y dificulta a los músculos realizar sus funciones con la facilidad habitual; es normal que este síntoma se presente después de largas jornadas de actividad física o mental, al no dormir lo suficiente, al padecer anemia o depresión o no alimentarse adecuadamente, entre otros factores.

(MED FOR08) (Traductor)

Tabla 119: Ejemplo de reformulación por expansión en una participación en foro sobre medicina con emisor lego no formado

1.5.9. Relación TIPO DE TEXTO-CLASE-IDT-IR

Si se relacionan los valores del IDT y del IR con la clase de mayor a menor grado de especialización y a los siete tipos textuales se distinguen dos bloques en dicha relación (*vid. infra* Gráfico 109).

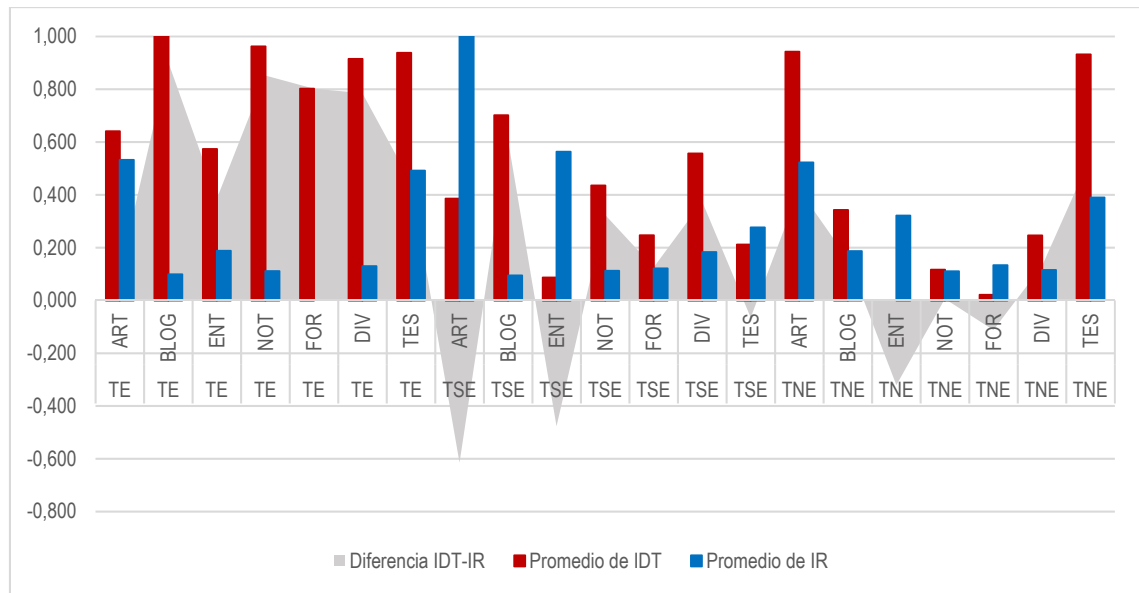


Gráfico 109: Superposición de datos normalizados de IDT global e IR global para tipo de texto y clase

El primer bloque corresponde con el texto especializado, en el cual la densidad terminológica prima sobre la reformulación en todos los tipos textuales. De hecho, los textos más reformulados en el texto especializado son aquellos que cuentan con mayor densidad terminológica, es decir, el artículo en revista científica y la tesis doctoral.

El segundo bloque, más complejo, engloba tanto el texto semiespecializado como el texto no especializado y presenta diferencias en los diversos tipos textuales. Uno de los tipos que más llama la atención es el elevado contraste entre el artículo en revista científica como texto semiespecializado (donde la reformulación cuenta con el valor más elevado del corpus) y como texto no especializado. Destaca también la entrevista, con valores de IDT muy reducidos en el corpus, pero con un IR elevado, de forma que puede interpretarse que la mayoría de la terminología se reformula (*vid. infra* Tabla 120).

En este sentido, debemos poner el foco en el **paciente frágil**, aquel especialmente vulnerable a presentar **deterioro funcional** y **eventos adversos**, y no tanto en el **manejo** de **patologías concretas** de forma individual. Evidentemente el **manejo** de **patologías crónicas** que afectan a la **función** (**fundamentalmente osteomusculares, neurodegenerativas y cardiovasculares**) resulta muy importante para evitar el **deterioro funcional**. Además en el anciano siempre debemos estar atentos a los denominados **síndromes geriátricos** (**caídas, delirium y deterioro cognitivo, disfagia, etc.**) que condicionan de forma significativa la calidad de vida de nuestros **pacientes**.

(MED ENT06) ^(Traductor)

Tabla 120: Ejemplo de reformulación por expansión en una entrevista sobre medicina

En general, la tendencia de la densidad terminológica y de la densidad reformulativa es disminuir conforme lo hace el grado de especialización textual (como ya se defendió en los apartados IV. 1.4.2. y IV. 1.5.2., pp. 310 y 326).

Si se desglosan los datos para los tipos de IR y se relacionan con el tipo textual y la clase, llama la atención que los gráficos (Gráfico 110 a Gráfico 112) sean casi idénticos. De hecho, las principales diferencias entre tipos textuales coinciden con las ya comentadas en el apartado IV. 1.5.3. (*vid. supra* Gráfico 86, p. 328).

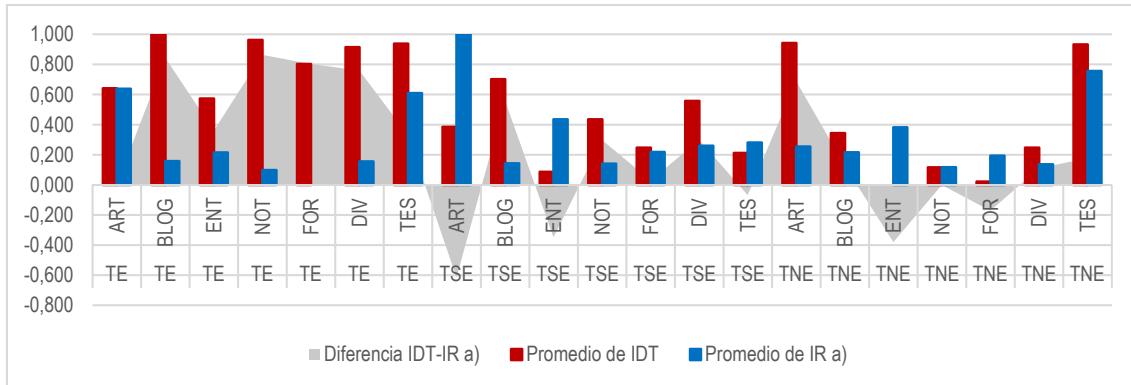


Gráfico 110: Superposición de datos normalizados de IDT global e IR a) para tipo de texto y clase

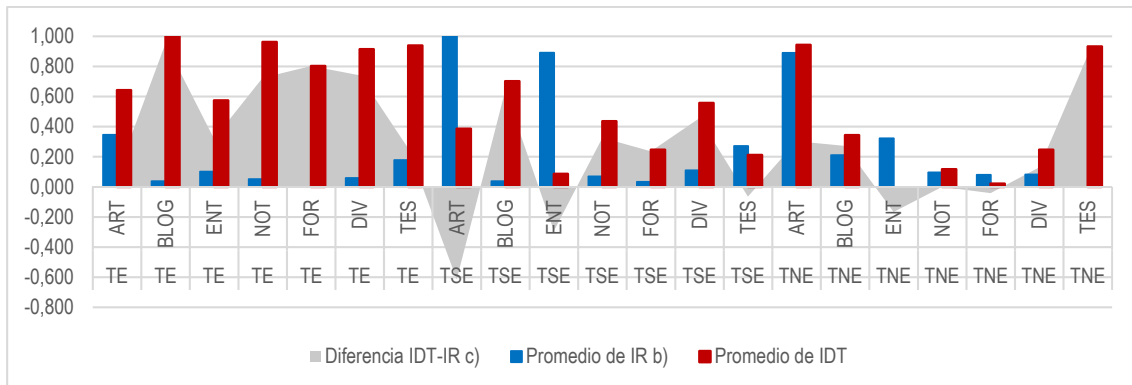


Gráfico 111: Superposición de datos normalizados de IDT global e IR b) para tipo de texto y clase

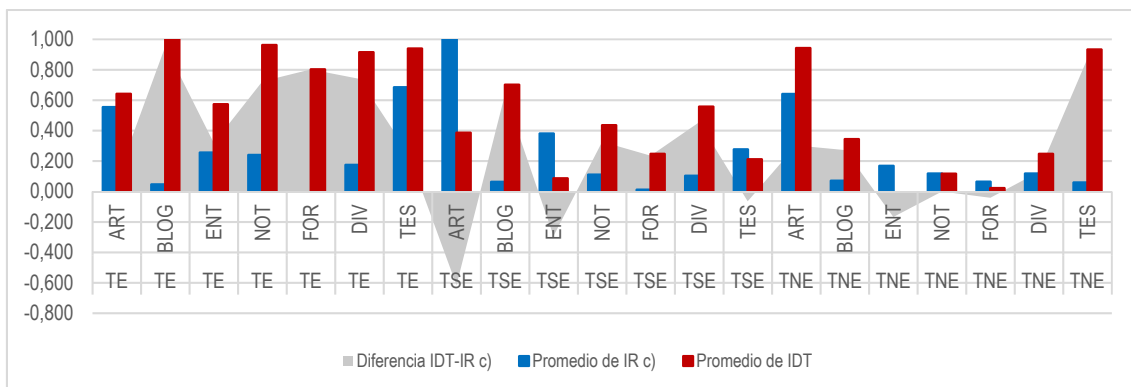


Gráfico 112: Superposición de datos normalizados de IDT global e IR c) para subdestinatario y clase

1.5.10. Conclusiones parciales sobre la REFORMULACIÓN

Como síntesis de esta sección, los párrafos siguientes resumen las descripciones y explicaciones más representativas sobre la densidad reformulativa asociada a la terminología de los textos del corpus.

- a) El índice de reformulación no contribuye a discriminar los temas de un texto. Tan solo es posible defender que el derecho es el tema menos reformulado, en contraste con la informática. Además, esta última presenta una densidad reformulativa por reducción mucho más elevada que el resto de temas.
- b) Si se atiende a la clase, se identifica una relación directa entre el IR global y el IR por expansión de un texto y su grado de especialización, lo cual permite defender que la densidad reformulativa asociada a la terminología contribuye a discriminar el grado de especialización textual (de igual forma que sucede con el IDT). Tan solo el IR b), que opera con los procedimientos por reducción, no contribuye a la discriminación de clases según el grado de especialización.
- c) En cuanto a los tipos de texto, distinguimos dos grupos de tipos textuales según la densidad reformulativa: (i) el primero está compuesto por aquellos tipos textuales que superan ampliamente el promedio de IR, como son el artículo en revista científica, la tesis doctoral y la entrevista; (ii) el segundo está compuesto por aquellos tipos textuales que se alejan del promedio, como son el blog, la noticia, la participación en foro y la revista o web de divulgación especializada. Las diferencias entre temas y tipos textuales no son especialmente significativas y coinciden con las expuestas para los tipos de PRF.
- d) En relación con el emisor y el subemisor, se establece una relación directa entre el IR y la competencia cognitivo-comunicativa del emisor y el subemisor. No obstante, dada la similitud de los valores entre los diferentes subemisores, no es posible defender que el IR contribuya significativamente a diferenciar las subcategorías de emisores.
- e) Las relaciones entre IR y destinatario y subdestinatario son muy complejas. Dado que no se establece una relación clara y que los subdestinatarios que cuentan con un IR más elevado corresponden con el especialista investigador, el instruido semiespecialista y el instruido aprendiz de especialista, se ha considerado que la falta de conocimiento sobre ciertos temas implica el aumento de la reformulación de las formas léxicas especializadas atendiendo a cada destinatario, por eso la

reformulación no decrece, sino que aumenta entre el especialista no investigador y el instruido aprendiz de especialista. Dado que a partir de este subdestinatario el valor de IR decrece de nuevo hasta alcanzar el valor mínimo en el lego no formado, no puede asociarse una relación inversa entre IR y perfil cognitivo-comunicativo del destinatario. Con todo, sí se advierte que el papel que desempeña la reformulación por expansión contribuye a discriminar los subdestinatarios, puesto que la primera aumenta cuando la competencia del segundo lo hace. Además, también se advierte que los procedimientos de variación denominativa o acronimia (procedimientos por variación), que son los principales identificados entre los procedimientos por variación, no son frecuentes en los textos que recibe un usuario con una competencia cognitivo-comunicativa menor.

- f) En relación con el subemisor y la clase, es posible afirmar que el IR se vincula no solo con la clase, también con el emisor del texto, ya que la reformulación aumenta cuando la conjunción de emisor y clase aumentan hasta el máximo grado y disminuye a medida que ambos lo hacen. En cuanto a la relación entre IDT e IR con la clase y el subemisor, se identifica una relación directa entre los dos grupos, si bien existe un punto de inversión entre IDT e IR para el texto no especializado: el IDT prima, por lo general, en el texto especializado y semiespecializado, mientras que el IR prima en el texto no especializado.
- g) Atendiendo a la relación entre IR con la clase y el subdestinatario (así como con el IDT), se distingue que el destinatario determina la reformulación de un texto cuando se asocia a la clase.
- h) La competencia del destinatario instruido desempeña un papel relevante como factor regulador de la reformulación, especialmente en los casos en los que se opera con el procedimiento de término seguido de descripción o explicación.

1.6. Análisis de la VARIEDAD LÉXICA

1.6.1. Según TEMA e IDT

El promedio de TTR estandarizada es similar en los tres temas del corpus, rondando la media de 65,5 % de variedad léxica (*vid. infra* Gráfico 113). A pesar de que gráficamente, dicho gráfico pueda mostrar un desnivel significativo, este tan solo se ajusta a la escala del

gráfico, siendo las diferencias entre columnas de apenas algunas décimas. A pesar de que existe una relación directa entre la TTR estandarizada y el IDT, este dato refleja que la variedad léxica de los textos del corpus no permite diferenciar el tema de los textos.

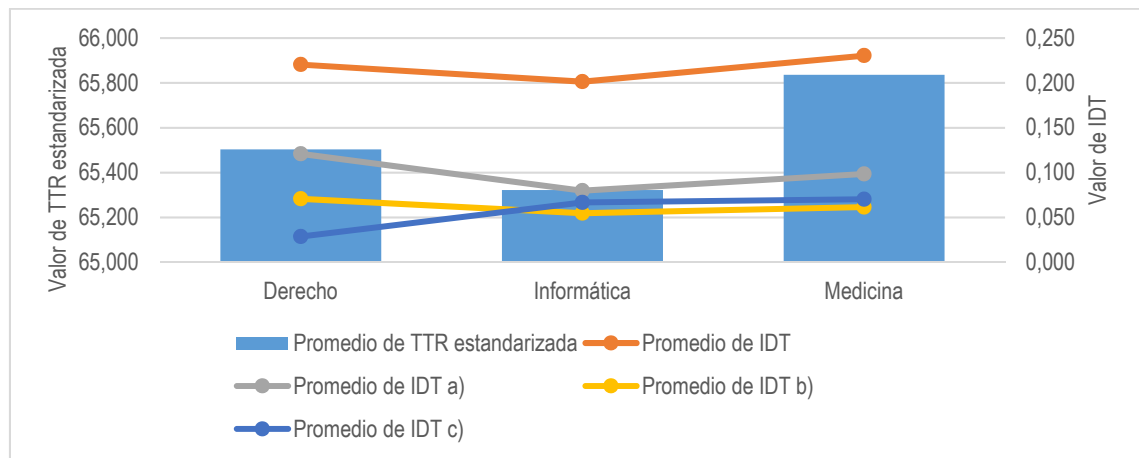


Gráfico 113: Datos de promedio de TTR estandarizada según tema e IDT

El índice de densidad terminológica se comporta de manera similar a la TTR estandarizada, puesto que el primero disminuye cuando la segunda lo hace, lo que puede explicarse teniendo en cuenta que el IDT se calcula teniendo en cuenta el número de formas léxicas. Al desglosar los tipos de IDT, se observa que el patrón sigue siendo similar, si bien es más elevado en los casos IDT a) e IDT b) en derecho que en medicina, siendo las diferencias menos acusadas y más sutiles que en el caso del IDT global. La mayor diferencia se encuentra en el promedio de IDT c), que se trata de la terminología poco frecuente y más críptica, que rompe al patrón identificado entre TTR estandarizada e IDT global, puesto que es menor que la media en derecho y mayor en informática y medicina.

1.6.2. Según TEMA e IR

El Gráfico 114 no revela ninguna relación clara entre la TTR estandarizada y el IR de cada subcorpus según el tema. De hecho, aunque pareciese identificarse una relación inversa, la tabla de datos que acompaña al gráfico no revela que esto suceda en todos los temas. Tan solo el tema establece una relación inversa entre la mínima TTR estandarizada y el máximo valor para IR global, IR b) e IR c).

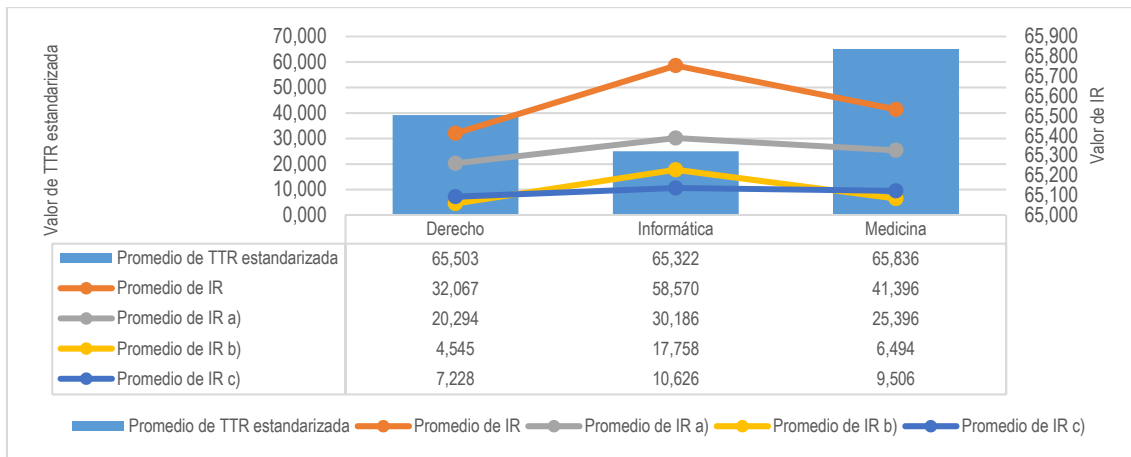


Gráfico 114: Datos globales de promedio de TTR estandarizada según tema e IR

1.6.3. Según CLASE e IDT

En primer lugar, como indica el Gráfico 115, la relación entre TTR estandarizada y grado de especialización textual es inversa, puesto que a medida que aumenta la primera, disminuye la segunda. Dado que entre IDT y grado de especialización existe una relación directa (*vid. supra* Gráfico 67; apartado IV. 1.4.2., p. 310), por pura transitividad, la relación entre TTR estandarizada y clase es inversa: la variedad léxica de un texto disminuye cuando aumenta su grado de especialización. Cabe destacar la similitud de variedad léxica que comparten el texto semiespecializado y el texto no especializado, que apenas se diferencian 4 puntos.

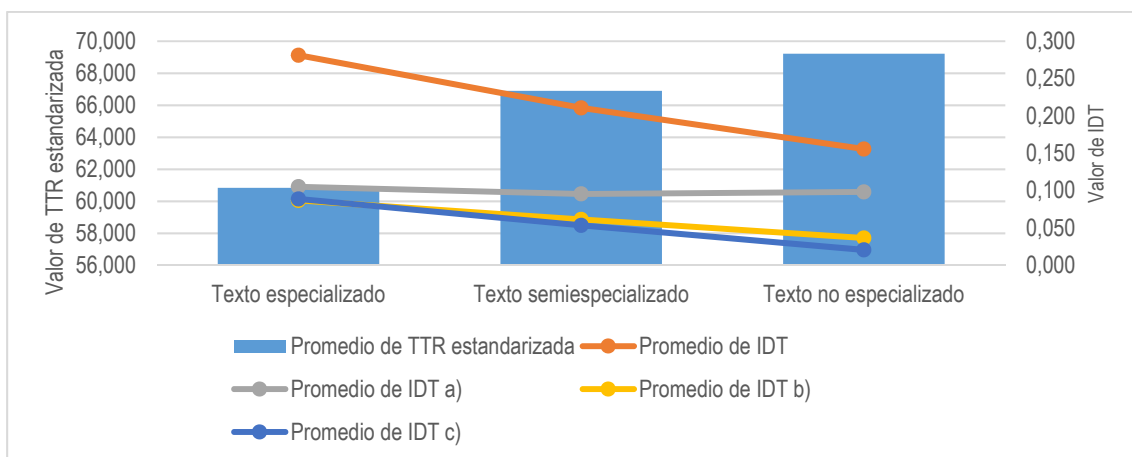


Gráfico 115: Datos globales de promedio de TTR estandarizada según clase de texto e IDT

“Los **algoritmos** no son necesariamente equitativos, puesto que es la persona que los crea la que define su funcionamiento y sus resultados”, subraya Cathy O’Neil, una científica especializada en datos informáticos y autora del libro *Weapons of Math Destruction* (Armas matemáticas de destrucción), que critica el uso generalizado y opaco de los **algoritmos**.

En un reciente foro organizado por el centro de reflexión New America, O’Neil alertó sobre la “confianza ciega” depositada en los **algoritmos** para obtener resultados objetivos.

(INF NOT14) (Traductor)

Tabla 121: Ejemplo de texto no especializado con TTR elevada

Descrito por primera vez por Rudolf Virchow en 1854 como “tuberculum dolorosum”, hay que decir que los **leiomiomas cutáneos** son **tumores benignos** poco frecuentes, seguramente infradiagnosticados, y que la localización **cutánea** es la más frecuente después de la **uterina**. Los **cutáneos** afectan ambos sexos por igual (algunos dicen que con cierta predominancia femenina). La concurrencia de **leiomiomas uterinos** y **cutáneos** se describió como una condición **hereditaria** en 1973 conocida como **síndrome de Reed**. Posteriormente, en 2001 el **síndrome de Reed** se asoció a **carcinoma renal** en algunos casos, en lo que se conoce hoy en día como **leiomiomatosis hereditaria** y **carcinoma renal**, con más de 180 familias descritas, pero de eso hablaremos otro día.

(MED BLOG10) (Traductor)

Tabla 122: Ejemplo de texto semiespecializado con TTR elevada

Para conseguir una definición de **conceptualización** independiente del **estado mundo**, Guarino y col. introducen una **noción intensional** de **relación** que ellos llamaron **relación conceptual**. Una **relación conceptual** ρ^n de aridad n definida sobre el **espacio del dominio** (D, W) , es una función total $\rho^n : W \rightarrow 2^{D^n}$ del conjunto de posibles **estados** del **mundo** W en el conjunto de todas las posibles **relaciones extensionales** de aridad n de D , de manera que para cada **estado** del **mundo** $w \in W$, $\rho^n(w)$ es la **relación** de R que contiene las **tuplas** de elementos del **dominio** que satisfacen la **relación** ρ^n en ese **estado concreto** del **mundo**. Así, dos agentes distintos, compartirán el significado de una **relación conceptual** si, en presencia del mismo **estado** del **mundo**, seleccionan las mismas **tuplas** como instancias de la **relación extensional asociada**. De no ser así, ambos agentes tendrían diferentes **conceptualizaciones** de la **relación**.

(INF TES13) (Traductor)

Tabla 123: Ejemplo de texto especializado con TTR reducida

En los ejemplos de la Tabla 121, Tabla 122 y Tabla 123 se han marcado en negrita aquellas formas léxicas repetidas en cada texto. Como puede comprobarse, a medida que se reduce la TTR (es decir, hay mayor número de formas léxicas marcadas en negrita), más especializado es el texto, como es el caso del ejemplo de la Tabla 123. Esto se debe a que, dada la función de representar el conocimiento especializado, los textos especializados se sirven de la repetición léxica como forma de organización del contenido informativo, lo que hace que la variedad léxica sea mucho menor que en el texto no especializado.

1.6.4. Según CLASE e IR

Al igual que el caso anterior, por transitividad, la variedad léxica de un texto disminuye a medida que aumenta su grado de especialización (*vid. infra* Gráfico 116). Las similitudes entre texto semiespecializado y texto no especializado en cuanto a la TTR se mantienen.

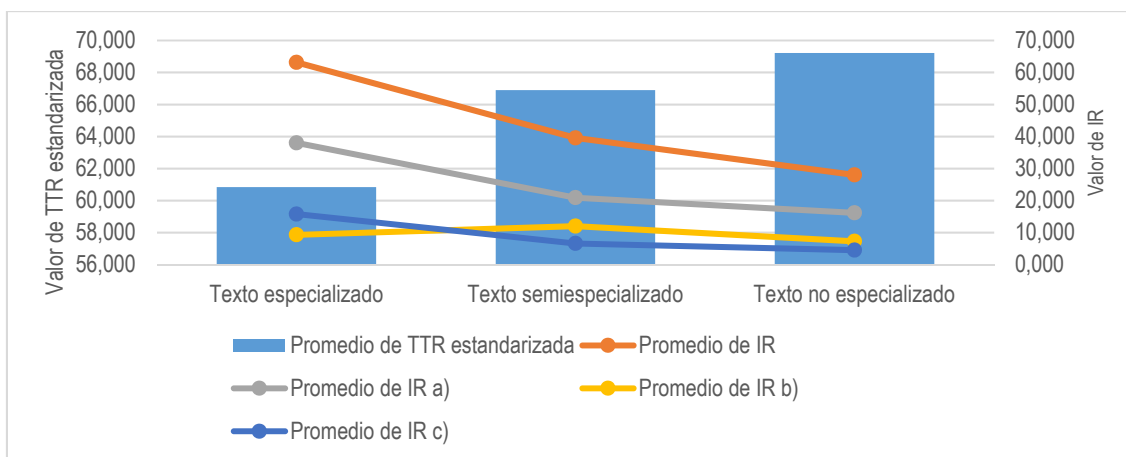


Gráfico 116: Datos globales de promedio de TTR estandarizada según clase de texto e IR

1.6.5. Según TIPO DE TEXTO e IDT

En el Gráfico 117 (*vid. infra*), a pesar de que la variedad léxica es muy similar en varios textos, se observa que a medida que aumenta el valor de TTR estandarizada (por ejemplo, en la entrevista, la noticia o el foro), disminuye el promedio de IDT. En aquellos tipos textuales con valores de TTR estandarizada más bajos (ART = 59,424; y TES = 60,209), la densidad terminológica aumenta (ART = 0,248; TES = 0,292, $\bar{x} = 0,217$).

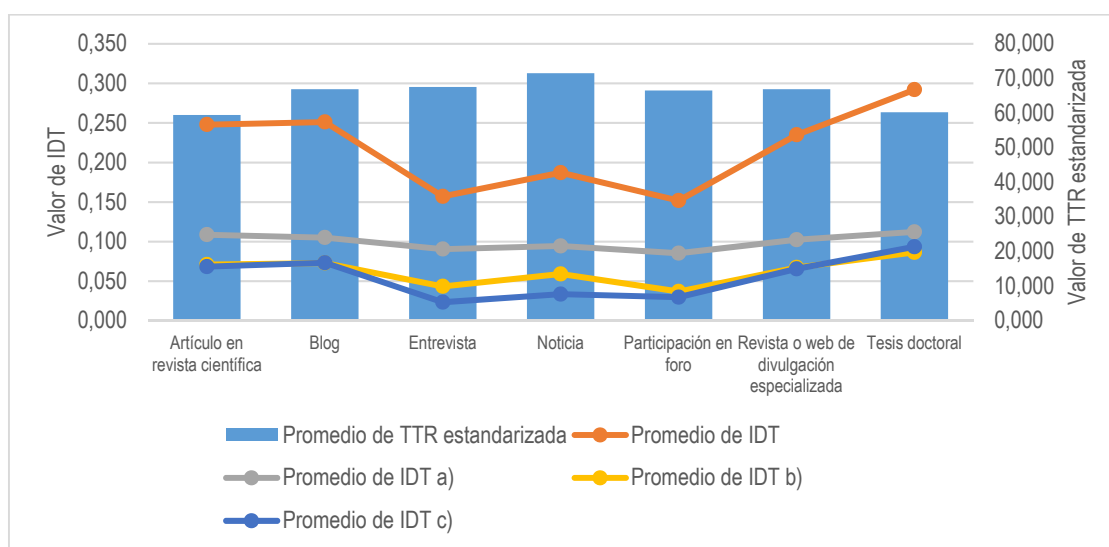


Gráfico 117: Datos globales de promedio de TTR estandarizada según tema e IDT

Cabe destacar los casos del blog y de la revista o web de divulgación especializada. Estos tipos textuales ya se han relacionado previamente como casos particulares (como en la Tabla 87 y la Tabla 88, donde se comparaban las características de estos tipos textuales en las tres clases). Especialmente, se trata de casos con una elevada densidad terminológica

(0,251 y 0,235 respectivamente), siendo los únicos tipos textuales no clasificados representativamente como textos especializados que acompañan al artículo en revista científica y la tesis doctoral como tipos textuales de elevada densidad terminológica. Así, consideramos que el público al que se suele dirigir el blog o la revista de divulgación especializada aúne tanto las características de elevada variedad léxica y de densidad terminológica (*vid. infra* Tabla 124).

El dolor lumbar es una de las causas más frecuentes de consulta al médico (la 5ª en EEUU)(1) con una prevalencia del 70% en países desarrollados a lo largo de la vida(2) y afectando al 33% de la población adulta en el Reino Unido cada año(3).

Es reconocido que el número de radiografías y resonancias magnéticas que se realizan por dolor lumbar inespecífico es excesivo y probablemente perjudicial. Un estudio reciente, robusto metodológicamente, realizado por la Universidad de Ottawa (4) establece que solo el 44% de las RNM que se realizan por dolor lumbar tenían una indicación adecuada.

La evidencia actual en forma de guías clínicas (NICE, Cochrane, American College of Physicians, American Pain Society y British Orthopaedic Association, entre otras) es clara en cuanto a la indicación de pruebas de imagen para la lumbalgia no específica.

En esta entrada resumo algunos conceptos y las conclusiones de estas guías clínicas para aplicarlas a nuestra práctica diaria tanto desde la Atención Primaria como Hospitalaria en cuanto a la solicitud de pruebas de imagen en la lumbalgia no específica.

(MED BLOG06) ^(Traductor)

Tabla 124: Ejemplo de blog con TTR e IDT elevados

1.6.6. Según TIPO DE TEXTO e IR

De forma similar al caso anterior, y relacionado con la transitividad que existe entre IDT e IR, en el Gráfico 118 se presentan los valores más bajos de TTR estandarizada para el artículo en revista científica y la tesis doctoral, a los cuales corresponde los valores más elevados de IR. El resto de tipos textuales presenta un IR inferior, a excepción de la entrevista, que, como ya se mencionó en el Gráfico 86 (*vid. supra* apartado IV.1.5.3., p. 328), es el tipo textual más reformulado junto con los tipos textuales clasificados como texto especializado, especialmente en informática (*vid. supra* Gráfico 89).

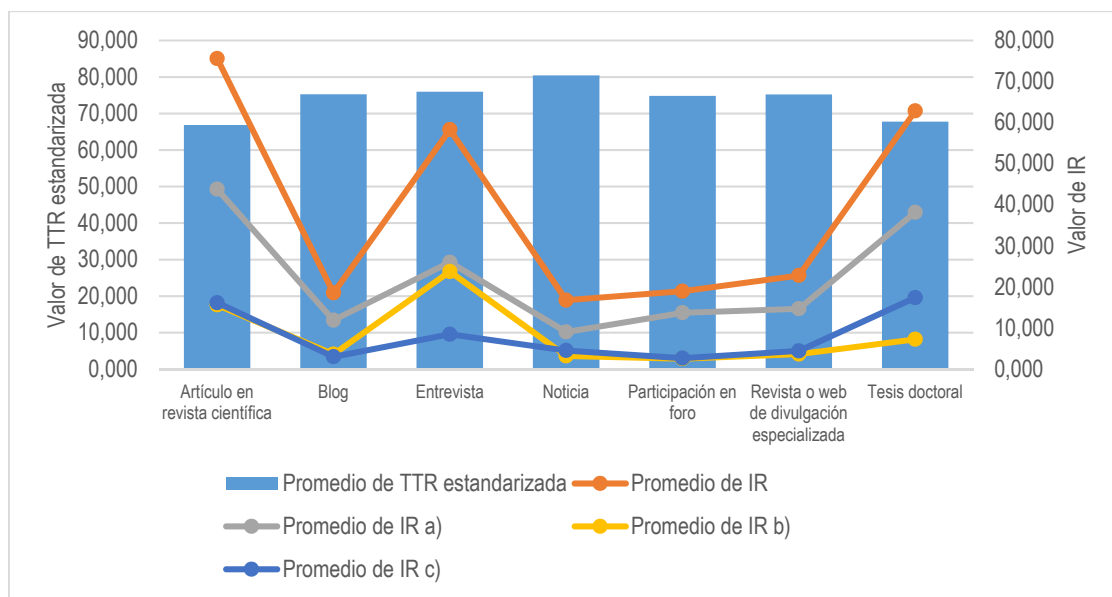


Gráfico 118: Datos globales de promedio de TTR estandarizada según tipo de texto e IR

1.6.7. Conclusiones parciales sobre la VARIEDAD LÉXICA

En definitiva, las siguientes son las conclusiones parciales en torno al análisis de la variedad léxica, en conjunción con el tema, la clase, el tipo de texto, el IR y el IDT:

- La variedad léxica de un texto no permite distinguir el tema del que versa.
- La variedad léxica de un texto permite distinguir su grado de especialización, ya que existe una relación inversa, en la que la variedad léxica disminuye a medida que aumenta el grado de especialización. Esta relación inversa también se establece con la densidad terminológica y la densidad reformulativa de un texto y también se aplica a los tipos textuales que se clasifican en cada clase, teniendo en cuenta, asimismo, las particularidades que suponen el blog y la revista o web de divulgación especializada para el IDT y la entrevista para el IR.

1.7. Resumen de análisis correlacional descriptivo global

La siguiente tabla sintetiza los valores más representativos de cada una de las variables analizadas en todo el subcorpus. Para facilitar la lectura de los datos, dicha tabla presenta la información en tres bloques, uno por tema, y cada bloque atiende a los siete tipos textuales. En las variables que cuentan con valores cuantitativos se ha marcado en verde aquel valor más elevado entre los tres temas, y en rojo, el valor más reducido. Los valores cuantitativos en color negro son los valores intermedios. Para el IR se usan sus

valores escalados. Como se puede apreciar a simple vista en la Tabla 125, el derecho y la informática son los temas que más contrastan: mientras que el derecho mantiene valores elevados de densidad terminológica, muestra valores reducidos de densidad reformulativa, de forma inversa a la informática. Para facilitar la interpretación de datos de esta tabla, el anexo 6 cuenta con el listado de abreviaturas empleado.

Tema y tipo texto	TTR	IDT				IR				Emisor	Subemisor	Destinatario	Subdestinatario	Clase
		Global	a	b	c	Global	a	b	c					
Derecho														
ART	63,493	0,231	0,126	0,076	0,029	0,090	0,105	0,023	0,144	Esp	Esp I	Esp	Esp I	TE
BLOG	63,935	0,278	0,147	0,099	0,032	0,021	0,026	0,004	0,032	Esp	Esp NI	Ins	Leg NF	TSE
ENT	67,467	0,214	0,110	0,063	0,040	0,025	0,027	0,006	0,050	Ins	Ins SE	Ins	Leg NF	TNE
NOT	67,471	0,176	0,110	0,054	0,012	0,042	0,059	0,007	0,049	Ins	Ins SE	Leg	Leg NF	TNE
FOR	63,268	0,154	0,100	0,047	0,008	0,011	0,019	0,001	0,006	Leg	Leg NF	Ins	Esp NI / Ins LF	TNE
DIV	68,982	0,230	0,108	0,082	0,041	0,008	0,007	0,001	0,026	Ins	Ins SE	Ins	Leg NF	TSE
TES	63,908	0,260	0,147	0,074	0,040	0,088	0,119	0,011	0,142	Esp	Esp I	Esp	Esp I	TE
Informática														
ART	60,358	0,247	0,080	0,070	0,098	0,158	0,178	0,047	0,233	Esp	Esp I	Esp	Esp I	TE
BLOG	65,187	0,240	0,081	0,056	0,103	0,026	0,030	0,009	0,030	Ins	Ins SE	Ins	Ins LF	TSE
ENT	67,468	0,243	0,097	0,063	0,083	0,040	0,062	0,006	0,032	Ins	Ins SE	Ins	Ins LF	TSE
NOT	68,623	0,136	0,057	0,038	0,041	0,136	0,086	0,102	0,107	Ins	Ins SE	Leg	Leg NF	TNE
FOR	64,095	0,161	0,091	0,031	0,040	0,017	0,018	0,007	0,020	Ins	Ins LF	Ins	Ins LF	TNE
DIV	71,452	0,140	0,079	0,034	0,027	0,031	0,037	0,005	0,063	Ins	Ins SE	Ins	Ins LF	TSE
TES	60,069	0,243	0,075	0,092	0,077	0,103	0,124	0,021	0,178	Esp	Esp I	Esp	Esp I	TE
Medicina														
ART	54,421	0,266	0,121	0,067	0,079	0,076	0,094	0,019	0,110	Esp	Esp I	Esp	Esp I	TE
BLOG	71,555	0,236	0,087	0,063	0,085	0,032	0,046	0,007	0,029	Esp	Ins SE / Esp NI	Leg	Leg NF	TNE
ENT	65,665	0,249	0,100	0,076	0,073	0,032	0,038	0,008	0,051	Ins	Esp I / Esp NI	Ins	Ins LF / Leg NF	TSE
NOT	66,432	0,160	0,105	0,038	0,017	0,062	0,075	0,014	0,099	Ins	Ins SE	Ins	Leg NF	TNE
FOR	72,094	0,140	0,066	0,033	0,041	0,053	0,082	0,006	0,053	Ins / Leg	Leg NF	Leg	Leg NF	TNE
DIV	74,037	0,190	0,096	0,061	0,033	0,032	0,034	0,012	0,047	Esp/Ins	Ins SE	Ins	Ins LF	TSE
TES	56,649	0,373	0,115	0,093	0,164	0,079	0,086	0,010	0,205	Esp	Esp I	Esp	Esp I	TE

Tabla 125: Resumen global de valores más representativos para cada variable según tema y tipo de texto

2. ANÁLISIS CORRELACIONAL DESCRIPTIVO SEGÚN LOS INFORMANTES

El objetivo de esta sección es comparar la percepción de los informantes de los valores de cada una de las variables en las que participan como facilitadores de información, es decir, contrastar los valores que ofrecen para la clase, el emisor, el subemisor, el destinatario, el subdestinatario y las variables relacionadas con la terminología, como son las formas léxicas especializadas y la densidad terminológica. De esta forma, se pretende analizar tanto la percepción analítica como la percepción sintética de los hablantes (*vid. supra* apartado III. 2.5., p. 175).

Debe recordarse el papel de las valoraciones subjetivas que llevan a cabo los hablantes, que parten del contraste de los criterios de ocurrencia-prominencia de los rasgos relevantes para distinguir variedades lingüísticas. Precisamente en este contraste de ocurrencia-prominencia puede atenderse a si estos son rasgos aceptados colectivamente por todos los grupos, por un grupo concreto, o si son fruto del azar (Caravedo, 2014: 125).

Tanto la TTR estandarizada como el índice de reformulación no se incluyen en esta sección, puesto que sus valores dependen del investigador y no son contrastables entre informantes. El caso del IR toma sus valores del IDT para cada grupo de informantes y los PRF son computados por el investigador de la tesis, de forma que el contraste entre informantes no es representativo.

En los apartados que se incluyen en esta sección se desglosan diferentes gráficos que cuentan con los distintos valores según cada informante. Además, para comparar en profundidad la vinculación de estos valores con otros, especialmente el tipo de texto, se desglosan cuatro gráficos (uno por informante) en bloques por cada valor de las variables relacionadas.

2.1. Análisis de la CLASE

En el Gráfico 119 no se aprecian diferencias notables entre informantes en la clasificación de ninguna de las tres clases.

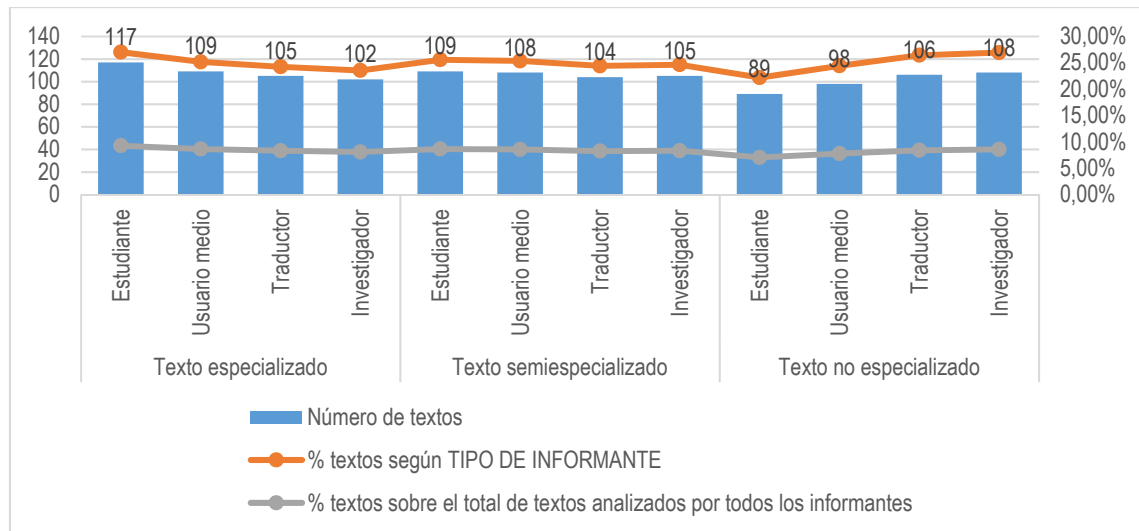


Gráfico 119: Datos de clase de texto según tipo de informante

Las diferencias más acusadas se encuentran entre estudiante e investigador para el texto especializado (117 y 102 textos respectivamente), y entre el estudiante y el investigador para el texto no especializado (89 y 108 textos respectivamente). Una diferencia de 19 textos tan solo supone una diferencia de un 1,5 % sobre el total del corpus y tan solo un 4,4 % sobre el total del subcorpus de textos especializados, por lo que no puede afirmarse que un informante opte por identificar mayor número de textos para una clase.

En la Tabla 126 se relacionan aquellos textos que han sido clasificados como texto especializado por un grupo de informantes, pero no por otro. En este caso, contamos con 37 textos que no coinciden en la clasificación. La amplia mayoría de estos textos está emitida por un especialista.

TE para estudiante y no para investigador			TE para investigador y no para estudiante		
DER	INF	MED	DER	INF	MED
DER DIV02	INF BLOG01	MED BLOG04	DER ART08	INF ART02	MED ART08
DER DIV03	INF BLOG03	MED DIV08	DER ART11	INF ART06	MED BLOG10
DER ENT12	INF BLOG04	MED DIV11	DER ART15	INF ART14	
DER FOR11	INF BLOG05	MED DIV12	DER TES04		
DER NOT02	INF DIV06	MED ENT01	DER TES09		
DER NOT03	INF DIV07		DER TES11		
DER NOT05	INF DIV14				
DER NOT06	INF ENT10				
DER NOT08	INF FOR03				
	INF FOR04				
	INF NOT04				
	INF TES05				

Tabla 126: Diferencias de clasificación de textos especializados entre grupo de informantes de estudiantes y de investigador

Con el objetivo de hallar las diferencias perceptivas para la clasificación, hemos optado por recurrir al índice de densidad terminológica de cada texto, puesto que es el más relacionado directamente con el grado de especialización. Como se muestra en la Tabla 127, casi todos los textos superan la media de IDT global (es decir, 0,217). Aquellos valores que no la superan se marcan en rojo.

TE para estudiante y no para investigador			TE para investigador y no para estudiante		
DER	INF	MED	DER	INF	MED
0,284	0,273	0,336	0,196	0,334	0,185
0,341	0,380	0,351	0,371	0,120	0,333
0,260	0,364	0,410	0,101	0,288	
0,362	0,282	0,340	0,304		
0,175	0,376	0,153	0,312		
0,288	0,363		0,369		
0,274	0,192				
0,211	0,357				
0,380	0,200				
	0,279				
	0,243				
	0,168				

Tabla 127: IDT global para textos especializados no coincidentes entre grupo de informantes de estudiantes y de investigador

Entre el IDT global no se identifica ninguna diferencia relevante entre los dos grupos de textos, por lo que, en segundo lugar, se ha recurrido a consultar los datos de IDT c), que son los que se relacionan en mayor medida con los textos especializados, especialmente con el artículo en revista científica y con la tesis doctoral (*vid. infra* Tabla 128). De igual forma, se indica el IDT c) de cada texto y se marcan en rojo aquellos que se encuentran por debajo de la media de IDT c) (es decir, 0,055).

TE para estudiante y no para investigador			TE para investigador y no para estudiante		
DER	INF	MED	DER	INF	MED
0,101	0,127	0,266	0,005	0,175	0,013
0,066	0,166	0,136	0,007	0,018	0,202
0,002	0,273	0,068	0,025	0,094	
0,083	0,153	0,156	0,092		
0,127	0,121	0,007	0,006		
0,014	0,139		0,009		
0,122	0,018				
0,033	0,205				
0,127	0,054				
	0,133				
	0,038				
	0,024				

Tabla 128: IDT c) para textos especializados no coincidentes entre grupo de informantes de estudiantes y de investigador

A pesar de que en el caso del investigador casi todos los textos cuentan con un IDT de términos crípticos inferior a la media, solo puede afirmarse que las diferencias en la clasificación atienden estrictamente a la percepción de cada informante del conjunto de características discursivas del texto. En otras palabras, que la combinación de perceptos almacenados por los individuos como resultado de la fijación permiten a los informantes tomar una decisión clasificatoria de los textos según el grado de especialización. Así, la densidad terminológica se vincula a otros factores, como podrían ser el emisor o el receptor, según las características lingüísticas observables por los informantes, para poner en marcha las operaciones de generalización (para medir de forma intuitiva *lo frecuente*) y de singularización (para medir de forma intuitiva *lo relevante*, que a veces puede vincularse con un fenómeno propio de un grupo o comunidad). En este caso, *lo relevante* son aquellos elementos que permiten ser vinculados al texto especializado.

2.1.1. Vinculación con el TIPO DE TEXTO

Si prestamos atención a la clasificación que realiza cada informante con respecto a los diferentes tipos textuales, puede comprobarse, de nuevo, que los cuatro grupos de informantes cuentan con más similitudes que diferencias. A continuación, se relacionan tres bloques con cuatro gráficos cada uno, en el que se enumeran los textos clasificados en cada clase según el grado de especialización según cada grupo de informantes.

El primer bloque enumera los textos para la clase de textos especializados (Gráfico 120 a Gráfico 123). La diferencia más relevante se encuentra en el foro y en la noticia, especialmente en el grupo de estudiantes, puesto que son los que asignan mayor número de textos especializados a dichos tipos textuales.

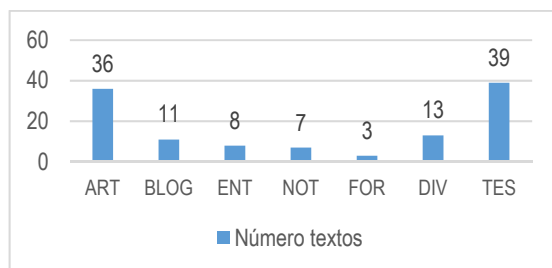


Gráfico 120: Número de textos especializados según grupo de informantes de estudiantes

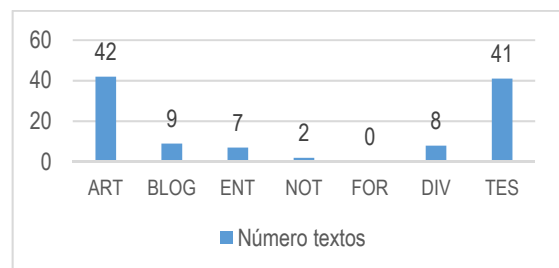


Gráfico 121: Número de textos especializados según grupo de informantes de usuarios medios

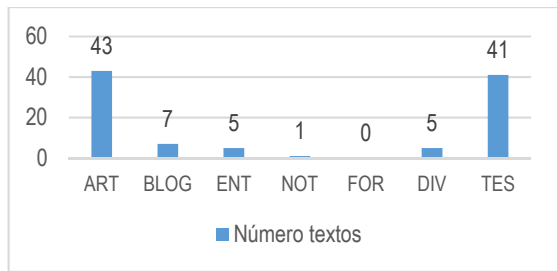


Gráfico 122: Número de textos especializados según investigador

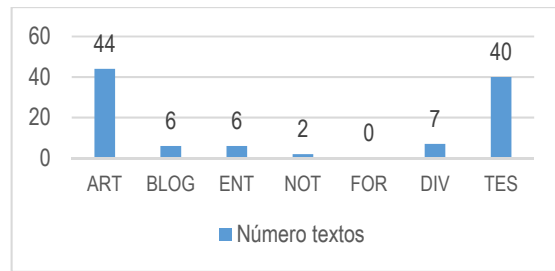


Gráfico 123: Número de textos especializados según grupo de informantes de traductores

En la Tabla 129 se relacionan las 7 noticias y las 3 participaciones en foro clasificadas como texto especializado por parte del grupo de estudiantes. En esta tabla se indican, además, las características que atribuyen el resto de informantes a dichos textos, de forma que puedan identificarse algunas de las razones que justifican que muchos de estos textos sean clasificados como semiespecializados o no especializados. Se ha marcado en negrita la fila dedicada a los informantes estudiantes, de forma que sirva como diferenciación visual para cada texto. Los colores se emplean con el fin de poder contrastar los datos de forma ágil.

Tipo de texto	Texto	Informante	IDT $\bar{x} = 0,217$	IR $\bar{x} = 43,668$	Subemisor	Subdestinatario	Clase
Noticia	DER NOT02	Estudiante	0,175	0,000	Especialista no investigador	Especialista no investigador	Texto especializado
		Us. medio	0,172	0,000	Especialista no investigador	Instruido semiespecialista	Texto semiespecializado
		Investigador	0,185	0,000	Especialista no investigador	Instruido aprendiz de especialista	Texto semiespecializado
		Traductor	0,162	0,000	Especialista no investigador	Instruido aprendiz de especialista	Texto semiespecializado
	DER NOT03	Estudiante	0,288	17,345	Especialista no investigador	Especialista no investigador	Texto especializado
		Us. medio	0,117	42,627	Especialista no investigador	Especialista no investigador	Texto especializado
		Investigador	0,300	16,656	Especialista no investigador	Instruido lego formado	Texto semiespecializado
		Traductor	0,199	25,150	Especialista no investigador	Instruido lego formado	Texto especializado
	DER NOT05	Estudiante	0,274	7,306	Instruido semiespecialista	Instruido lego formado	Texto especializado
		Us. medio	0,232	8,623	Especialista no investigador	Instruido lego formado	Texto semiespecializado
		Investigador	0,228	8,767	Instruido semiespecialista	Lego no formado	Texto semiespecializado
		Traductor	0,247	8,092	Instruido semiespecialista	Lego no formado	Texto semiespecializado
	DER NOT06	Estudiante	0,211	9,465	Especialista no investigador	Instruido semiespecialista	Texto especializado
		Us. medio	0,122	16,390	Instruido semiespecialista	Lego no formado	Texto no especializado
		Investigador	0,208	9,600	Instruido semiespecialista	Lego no formado	Texto no especializado
		Traductor	0,161	12,444	Instruido semiespecialista	Lego no formado	Texto no especializado
	DER NOT08	Estudiante	0,380	2,633	Instruido semiespecialista	Instruido semiespecialista	Texto especializado
		Us. medio	0,266	3,762	Instruido semiespecialista	Instruido lego formado	Texto semiespecializado

Tipo de texto	Texto	Informante	IDT $\bar{x} = 0,217$	IR $\bar{x} = 43,668$	Subemisor	Subdestinatario	Clase	
		Investigador	0,325	3,078	Instruido semiespecialista	Instruido lego formado	Texto semiespecializado	
		Traductor	0,287	3,485	Instruido semiespecialista	Instruido lego formado	Texto semiespecializado	
	DER NOT14	Estudiante	0,505	21,781	Especialista investigador	Especialista investigador	Texto especializado	
		Us. medio	0,416	26,427	Especialista no investigador	Especialista no investigador	Texto especializado	
		Investigador	0,463	23,750	Especialista no investigador	Especialista no investigador	Texto especializado	
		Traductor	0,445	24,734	Especialista no investigador	Especialista no investigador	Texto especializado	
	INF NOT04	Estudiante	0,243	24,719	Especialista no investigador	Instruido semiespecialista	Texto especializado	
		Us. medio	0,187	32,068	Instruido semiespecialista	Instruido aprendiz de especialista	Texto semiespecializado	
		Investigador	0,192	31,224	Instruido semiespecialista	Instruido lego formado	Texto semiespecializado	
		Traductor	0,193	31,020	Instruido semiespecialista	Instruido lego formado	Texto semiespecializado	
	Participación en foro	DER FOR11	Estudiante	0,362	0,000	Especialista investigador	Instruido semiespecialista	Texto especializado
			Us. medio	0,268	0,000	Especialista no investigador	Especialista no investigador	Texto semiespecializado
Investigador			0,298	0,000	Especialista no investigador	Especialista no investigador	Texto semiespecializado	
Traductor			0,298	0,000	Especialista no investigador	Especialista no investigador	Texto semiespecializado	
INF FOR03		Estudiante	0,200	5,000	Instruido aprendiz de especialista	Instruido aprendiz de especialista	Texto especializado	
		Us. medio	0,104	9,600	Instruido aprendiz de especialista	Instruido aprendiz de especialista	Texto no especializado	
		Investigador	0,117	8,571	Instruido lego formado	Instruido aprendiz de especialista	Texto no especializado	
		Traductor	0,117	8,571	Instruido lego formado	Instruido aprendiz de especialista	Texto no especializado	
INF FOR04		Estudiante	0,279	3,581	Instruido aprendiz de especialista	Instruido lego formado	Texto especializado	
		Us. medio	0,214	4,683	Instruido aprendiz de especialista	Instruido lego formado	Texto semiespecializado	
		Investigador	0,234	4,272	Instruido aprendiz de especialista	Instruido lego formado	Texto semiespecializado	
		Traductor	0,218	4,594	Instruido semiespecialista	Instruido lego formado	Texto semiespecializado	

Tabla 129: Características según grupos de informantes para textos de noticia y participación en foro clasificados como textos especializados por el grupo de informantes de estudiantes

En primer lugar, casi todos los informantes coinciden en los valores para emisor y destinatario, así como la clase¹⁶⁶. Las principales diferencias se hallan en el IDT que asigna cada informante al texto. De hecho, el estudiante asigna en todos los casos el valor más elevado para el IDT; a excepción del DER NOT02 y el DER NOT03, cuyo valor más elevado de IDT global lo asigna el investigador (si bien las diferencias con el valor del estudiante son muy reducidas, de solo algunas centésimas). Así, desde nuestro punto de vista, es posible afirmar que el grupo de informantes compuesto por estudiantes vincula de

¹⁶⁶ La clase es el único elemento no coincidente entre los grupos de informantes, ya que el estudiante es el elemento de contraste. Tan solo DER NOT03 y DER NOT14 se clasifican como texto especializado por el resto de grupos.

forma más estrecha la aparición de terminología (sobre todo de FLE délficas) (*vid. infra* apartado IV. 2.5., p. 386) con la clase de texto especializado.

En cuanto al texto semiespecializado (Gráfico 124 a Gráfico 127), las similitudes destacan sobre las divergencias.

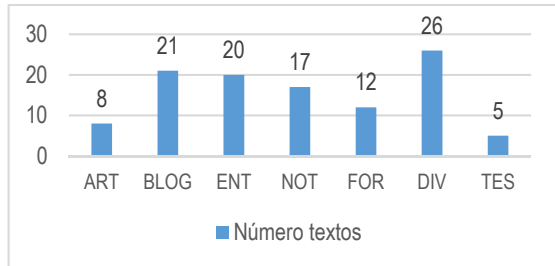


Gráfico 124: Número de textos semiespecializados según grupo de informantes de estudiantes

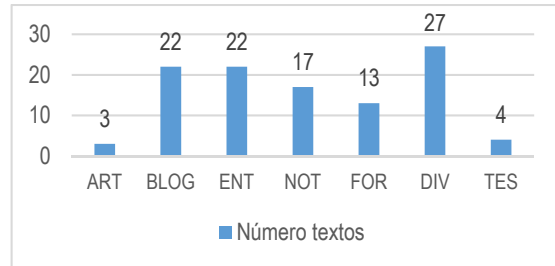


Gráfico 125: Número de textos semiespecializados según grupo de informantes de usuarios medios

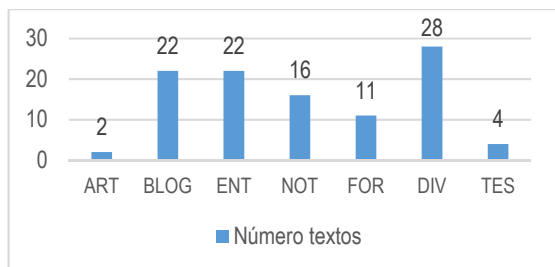


Gráfico 126: Número de textos semiespecializados según investigador

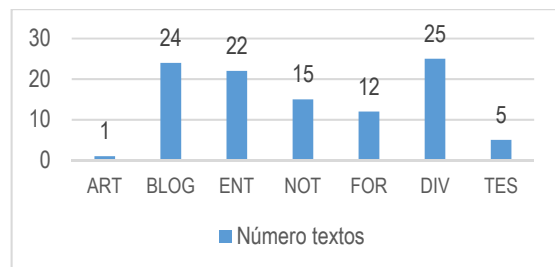


Gráfico 127: Número de textos semiespecializados según grupo de informantes de traductores

Tan solo el grupo de estudiantes identifica un mayor número de artículos en revistas científicas semiespecializados. De hecho, coinciden con aquellos que no fueron clasificados por los estudiantes como textos especializados (y sí por el resto de los informantes: DER ART11, DER ART15, INF ART02, INF ART06 y MED ART08). En la Tabla 130 se muestran las características para los artículos en revista científica clasificados como semiespecializados por todos los informantes, así como la clasificación que otorga cada informante a dicho texto. En la columna de código se marcan en negrita aquellos textos que se repiten en la clasificación de algún otro informante.

Informante	Código	IDT $\bar{x} = 0,217$	IR $\bar{x} = 43,668$	Subemisor	Subdestinatario	Clasificado como			
						Estudiante	Us. M	H. O. i.	Trad.
Estudiante	DER ART08	0,210	109,330	Especialista investigador	Especialista investigador	TSE	TSE	TE	TE
	DER ART11	0,389	38,533	Especialista investigador	Instruido aprendiz de especialista	TSE	TE	TE	TE
	DER ART15	0,114	105,558	Especialista no investigador	Instruido semiespecialista	TSE	TE	TE	TE
	INF ART02	0,360	130,564	Especialista investigador	Instruido semiespecialista	TSE	TE	TE	TE
	INF ART06	0,123	130,255	Instruido aprendiz de especialista	Instruido lego formado	TSE	TE	TE	TE
	INF ART14	0,278	79,037	Especialista investigador	Instruido semiespecialista	TSE	TSE	TE	TE

Informante	Código	IDT	IR	Subemisor	Subdestinatario	Clasificado como			
		$\bar{x} = 0,217$	$\bar{x} = 43,668$			Estudiante	Us. M	H. O. i.	Trad.
	MED ART08	0,220	13,657	Especialista investigador	Especialista investigador	TSE	TE	TE	TE
	MED ART09	0,192	219,098	Especialista investigador	Instruido aprendiz de especialista	TSE	TE	TSE	TE
Usuario medio	DER ART04	0,071	339,209	Especialista no investigador	Instruido lego formado	TNE	TSE	TSE	TSE
	DER ART08	0,173	132,857	Especialista investigador	Instruido aprendiz de especialista	TSE	TSE	TE	TE
	INF ART14	0,173	127,530	Especialista investigador	Especialista no investigador	TSE	TSE	TE	TE
Investigador	DER ART04	0,302	79,489	Especialista investigador	Instruido aprendiz de especialista	TNE	TSE	TSE	TSE
	MED ART09	0,192	219,098	Especialista investigador	Instruido aprendiz de especialista	TSE	TE	TSE	TE
Traductor	DER ART04	0,071	339,209	Especialista no investigador	Instruido lego formado	TNE	TSE	TSE	TSE

Tabla 130: Características de artículos en revista científica semiespecializados según informantes

Con el objetivo de justificar las causas que conllevan la diferencia de clasificación para los textos enumerados anteriormente, se ha considerado relevante recurrir a los valores de los diferentes IDT, puesto que la relevancia de las FLE délficas según los estudiantes hace considerar la posibilidad de que la clasificación como texto semiespecializado se deba a un aumento del IDT de términos délficos. La Tabla 131 no justifica nuestro argumento, puesto que solo dos textos superan ampliamente la media para IDT de términos délficos y no es posible afirmar que el resto de textos cuente con un conjunto de formas délficas relevante.

	IDT a)	IDT b)	IDT c)
	$\bar{x} = 0,100$	$\bar{x} = 0,062$	$\bar{x} = 0,055$
DER ART11	0,376	0,005	0,008
DER ART15	0,044	0,044	0,026
INF ART02	0,034	0,012	0,315
INF ART06	0,073	0,032	0,017
MED ART08	0,195	0,016	0,008

Tabla 131: Tipos de IDT para artículos en revista científica semiespecializados exclusivos del grupo de estudiantes

Desde nuestro punto de vista, en el caso del texto semiespecializado, en el estudiante prima la consideración del perfil cognitivo-comunicativo del destinatario, ya que los textos semiespecializados que se relacionan en la Tabla 130 cuentan con un subdestinatario instruido. Obviamente, la consideración del subdestinatario instruido queda justificada por las características discursivas de los textos. De hecho, el MED ART08 cumple casi todas las características para ser considerado texto especializado, como son un IDT elevado, un

emisor especialista y un destinatario especialista. Tan solo la reformulación es reducida en dicho texto. En este sentido, las características prominentes que el grupo de estudiantes identifica cuando opera con la singularización quedan vinculadas con el subdestinatario instruido. Además, esta clasificación resulta de la combinación de varias variables (siendo el subdestinatario instruido una pieza clave de la clasificación, como también se comprueba en el apartado IV. 3.3., p. 428), puesto que, como hemos justificado, el IDT de formas delficas es insuficiente para la clasificación del texto semiespecializado según los estudiantes.

Por último, en la clasificación de tipos textuales dentro del texto no especializado, las similitudes se mantienen. De nuevo, el grupo de estudiantes es el que muestra la principal divergencia con respecto a los otros tres grupos, ya que las noticias cuentan con una media de 7 textos menos que en el resto de grupos, y las revistas o webs de divulgación especializada cuentan con una media de 6 textos menos (*vid. infra* Gráfico 128 a Gráfico 131).

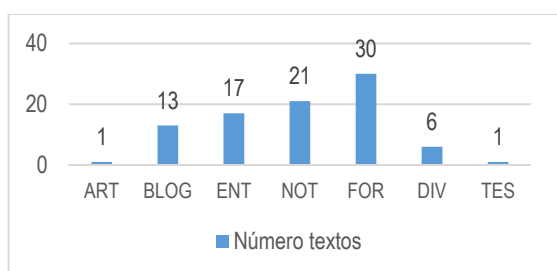


Gráfico 128: Número de textos no especializados según grupo de informantes de estudiantes

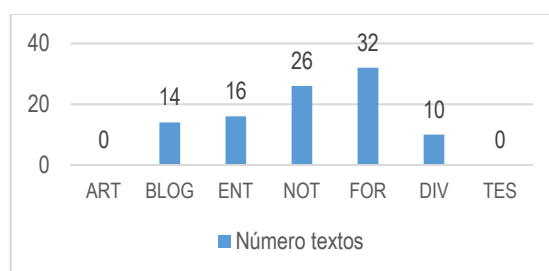


Gráfico 129: Número de textos no especializados según grupo de informantes de usuarios medios

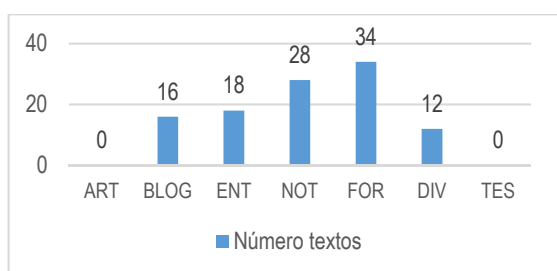


Gráfico 130: Número de textos no especializados según investigador

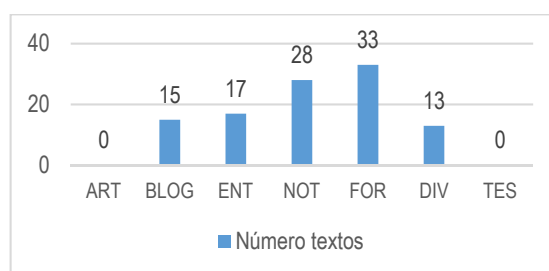


Gráfico 131: Número de textos no especializados según grupo de informantes de traductores

Entre los textos no especializados no hay grandes diferencias, de hecho, la ausencia de artículos en revista científica o tesis doctorales entre los textos no especializados en todos los informantes (a nuestro juicio, los seleccionados por los estudiantes no son relevantes, puesto que solo se trata de un texto) manifiesta la clara relación que establecen los usuarios entre el máximo grado de especialización textual y los dos tipos textuales relacionados con la transmisión de información científica. En otras palabras, existen rasgos

prominentes en estos dos tipos textuales que se perciben como representativos de la comunicación especializada.

Uno de los elementos de los que podrían servirse los informantes para asociar estos dos tipos textuales al grado máximo de especialización podría ser la superestructura del texto, es decir, la organización y distribución formal de la información. A pesar de que los textos han sido desprovistos de su formato original y facilitados en texto plano a los informantes, las tesis doctorales y los artículos en revistas científicas sí dejan entrever una estructura formal ligeramente diferente, a veces en la densidad de formas acronímicas o la estructuración en bloques de apartados. No obstante, lo que más destaca en estos textos suele ser el uso de formas léxicas poco frecuentes, y no su estructuración formal.

En el apartado IV. 2.1. (*vid. supra* Gráfico 119, p. 359) se mostró la preferencia del grupo de estudiantes por los textos especializados y semiespecializados, ya que, no solo son el grupo que más relevancia tiene en estas clases, sino el que menos tiene para el texto no especializado. La diferencia para la noticia o los textos procedentes de revistas o webs de divulgación especializada queda justificada desde el punto de vista de la densidad terminológica con la que cuentan dichos textos, que inclina al estudiante a asignar dichos textos a la clase de textos semiespecializados.

2.2. Análisis del EMISOR y del SUBEMISOR

A pesar de que en la clasificación de los tipos de emisores los valores según los informantes son muy similares, existen tres particularidades (*vid. infra* Gráfico 132):

- a) El investigador opta por el emisor instruido antes que por el especialista.
- b) El estudiante opta por el especialista antes que por el instruido. Esta tendencia concuerda con la tendencia de clasificación de textos especializados por parte del estudiante (*vid. supra* Gráfico 119, apartado IV. 2.1., p. 359).
- c) El usuario medio opta por el especialista antes que por el instruido.

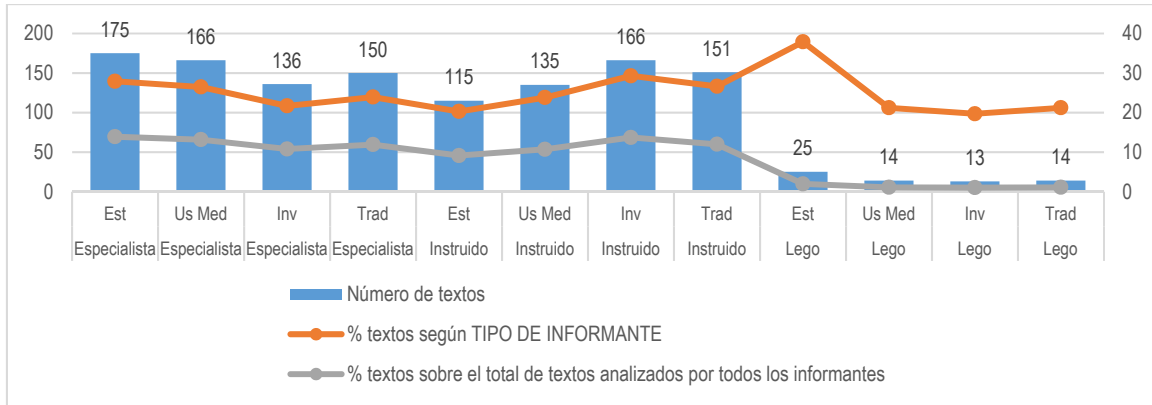


Gráfico 132: Datos de emisor según tipo de informante

En el caso de los subemisores (*vid. infra* Gráfico 133), las diferencias entre informantes son más notables, especialmente porque se observan ciertas tendencias de asignación de valores:

- a) El estudiante tiende muy ligeramente hacia el subemisor especialista investigador.
- b) El usuario medio, el investigador y el traductor tienden a optar por el instruido semiespecialista en lugar de por el especialista no investigador. Esta tendencia es más evidente en el caso del investigador.

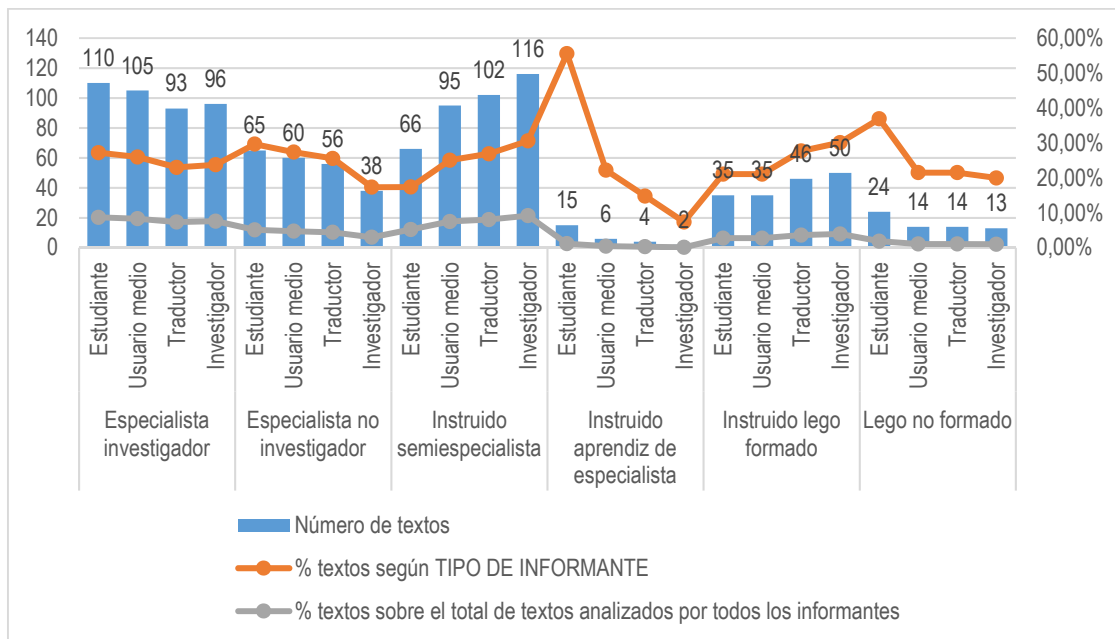


Gráfico 133: Datos de subemisor según tipo de informante

A la vista de los resultados de los subemisores más asignados a los textos (y atendiendo especialmente a una comparación entre los informantes), las categorías de subemisor más relevantes son las de especialista investigador, instruido semiespecialista e

instruido lego formado. Desde nuestro punto de vista, esta prevalencia de valores queda justificada por dos razones hipotéticas:

- a) Para los informantes es más fácil contar con tres valores para asignar, puesto que en ellos identifican un valor máximo, intermedio y mínimo.
- b) Para los informantes, estos tres valores de la categoría de subemisor constituyen las categorías prototípicas para la categoría de emisor, es decir: especialista investigador para la categoría de especialista; instruido semiespecialista para la categoría de instruido; e instruido lego formado para la categoría de lego. De hecho, aunque en nuestro modelo teórico el instruido lego formado es el último nivel de la categoría de instruido (y el lego no formado la única para la categoría de lego), es coherente con el criterio de los informantes que el destinatario cuente con un mínimo de conocimiento sobre un tema para poder llegar a comprenderlo (como cumple el instruido lego formado). Esta presuposición de conocimiento por parte del emisor o bien la dificultad de hallar un límite claro entre subcategorías podría justificar la ausencia de la subcategoría de aprendiz de especialista.

2.2.1. Vinculación con el TIPO DE TEXTO

En relación con la vinculación de los subemisores con los tipos de texto, se presentan a continuación siete bloques de gráficos que contienen los datos para cada subemisor según cada informante para los siete tipos de texto. La organización de este apartado atiende principalmente a la similitud de frecuencias según los informantes.

El subemisor para el tipo textual artículo en revista científica (Gráfico 134 a Gráfico 137) y tesis doctoral (Gráfico 138 a Gráfico 141) es casi unánime en los cuatro grupos de informantes: el especialista investigador es el más frecuente, seguido muy de lejos del especialista no investigador.

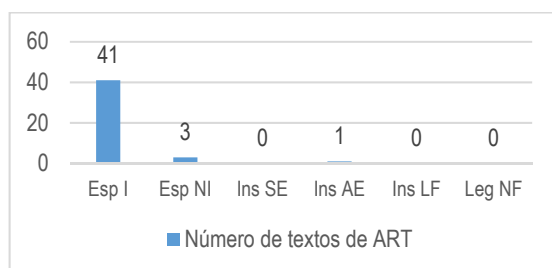


Gráfico 134: Número de textos de artículos en revista científica según subemisor según grupo de informantes de estudiantes

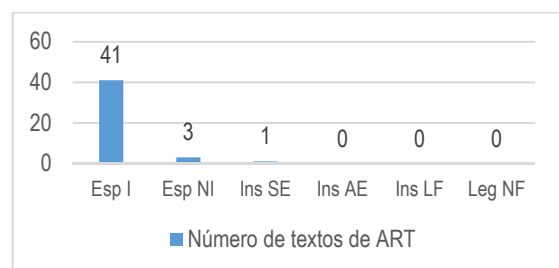


Gráfico 135: Número de textos de artículos en revista científica según subemisor según grupo de informantes de usuarios medios

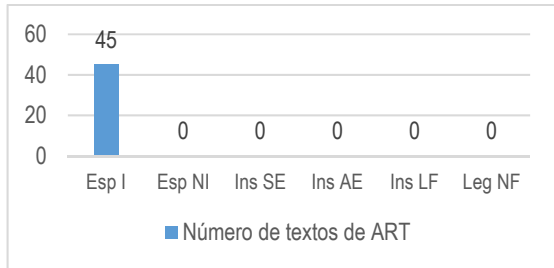


Gráfico 136: Número de textos de artículos en revista científica según subemisor según investigador

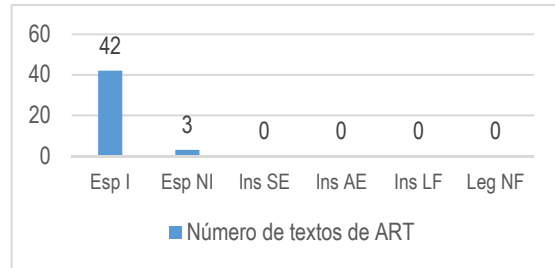


Gráfico 137: Número de textos de artículos en revista científica según subemisor según grupo de informantes de traductores

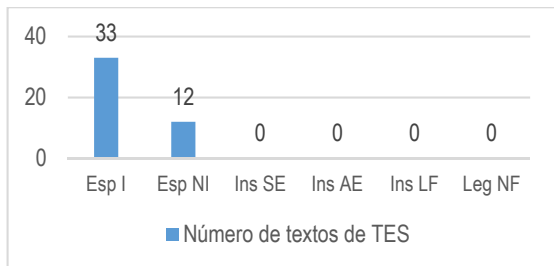


Gráfico 138: Número de textos de tesis doctorales según subemisor según grupo de informantes de estudiantes

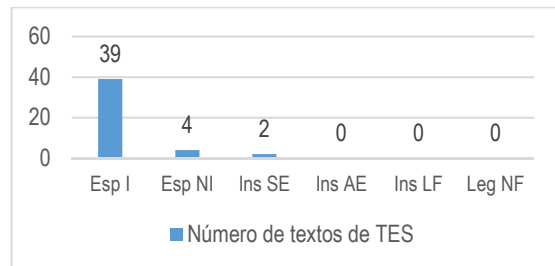


Gráfico 139: Número de textos de tesis doctorales según subemisor según grupo de informantes de usuarios medios

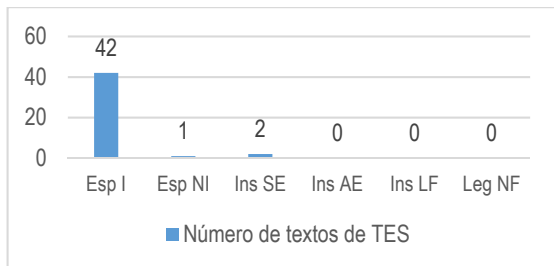


Gráfico 140: Número de textos de tesis doctorales según subemisor según investigador

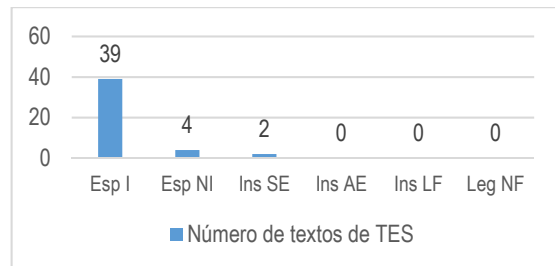


Gráfico 141: Número de textos de tesis doctorales según subemisor según grupo de informantes de traductores

Como contrapunto al máximo nivel de competencia cognitivo-comunicativa, la participación en foro también cuenta con un criterio homogéneo entre los informantes, puesto que se vincula estrechamente con el instruido lego formado y el lego no formado. La frecuencia se invierte para el estudiante.

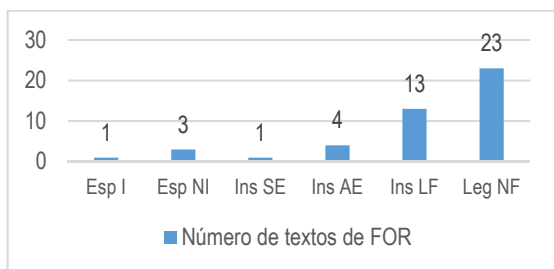


Gráfico 142: Número de textos de participaciones en foro según subemisor según grupo de informantes de estudiantes

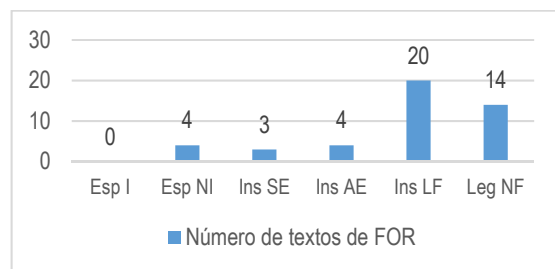


Gráfico 143: Número de textos de participaciones en foro según subemisor según grupo de informantes de usuarios medios

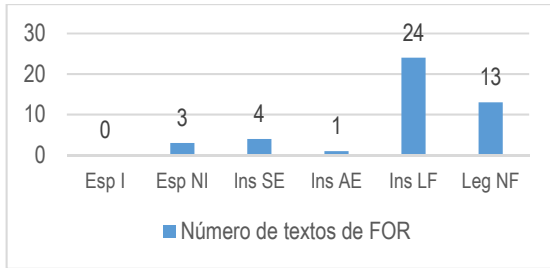


Gráfico 144: Número de textos de participaciones en foro según subemisor según investigador

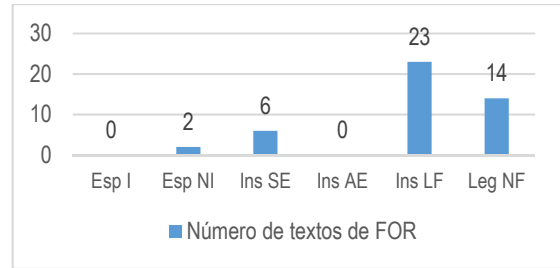


Gráfico 145: Número de textos de participaciones en foro según subemisor según grupo de informantes de traductores

Esta convergencia en los criterios se mantiene en la noticia (*vid. infra* Gráfico 146 a Gráfico 149). Tan solo el estudiante diverge de los otros tres informantes, puesto que la distribución de subemisores es más dispar. En el caso del usuario medio, investigador y traductor, no solo coinciden los tipos de emisores, sino que, específicamente en los dos últimos, coincide el número de textos con dichos valores.

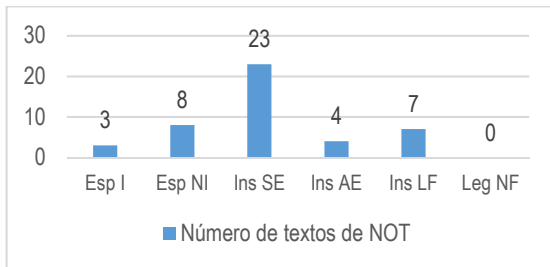


Gráfico 146: Número de textos de noticias según subemisor según grupo de informantes de estudiantes

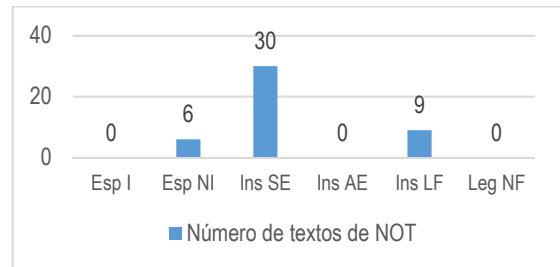


Gráfico 147: Número de textos de noticias según subemisor según grupo de informantes de usuarios medios

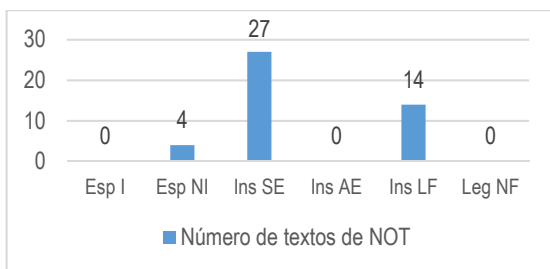


Gráfico 148: Número de textos de noticias según subemisor según investigador

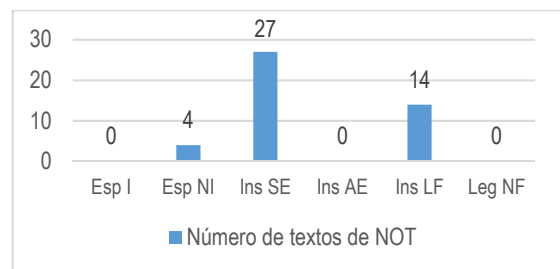


Gráfico 149: Número de textos de noticias según subemisor según grupo de informantes de traductores

Los tipos textuales restantes son más complejos, puesto que la asignación de valores para los subemisores por cada informante diverge ligeramente. En el caso del blog (*vid. infra* Gráfico 150 a Gráfico 153) y de la entrevista (*vid. infra* Gráfico 154 a Gráfico 157), es posible establecer una relación de similitud más clara entre el estudiante y el usuario medio, por una parte, y entre el investigador y el traductor, por otra. En el primer grupo los informantes optan por el especialista no investigador, mientras que en el segundo lo hacen por el instruido semiespecialista. Esta vinculación entre el estudiante y el usuario medio

contrasta ligeramente con lo defendido sobre la tendencia del usuario medio en cuanto al subemisor en el apartado IV. 2.2. (p. 367).

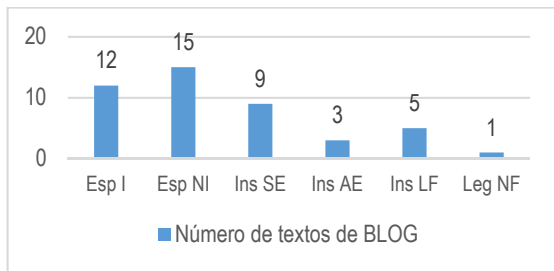


Gráfico 150: Número de textos de blogs según subemisor según grupo de informantes de estudiantes

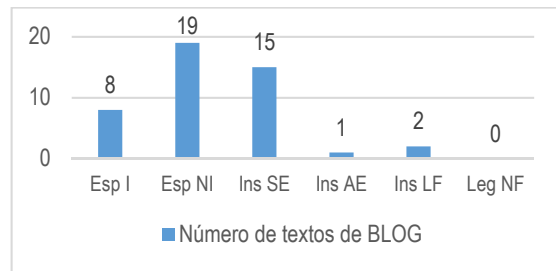


Gráfico 151: Número de textos de blogs según subemisor según grupo de informantes de usuarios medios

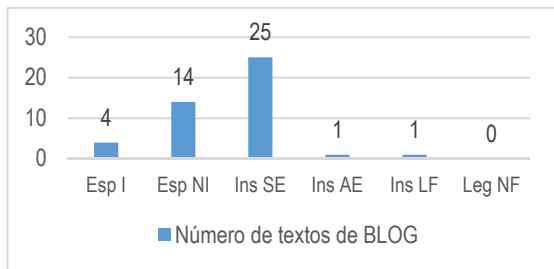


Gráfico 152: Número de textos de blogs según subemisor según investigador

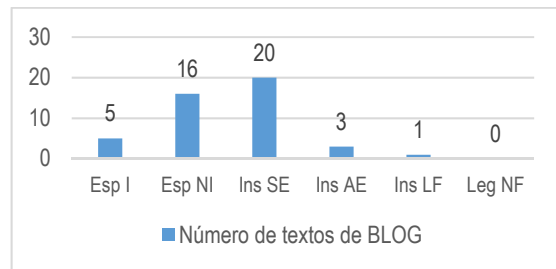


Gráfico 153: Número de textos de blogs según subemisor según grupo de informantes de traductores

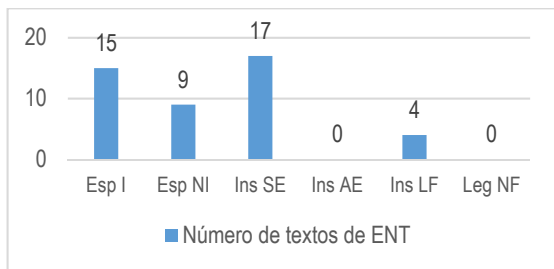


Gráfico 154: Número de textos de entrevistas según subemisor según grupo de informantes de estudiantes

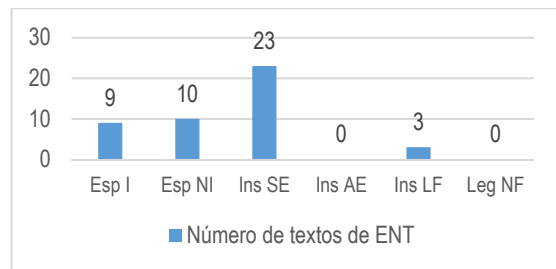


Gráfico 155: Número de textos de entrevistas según subemisor según grupo de informantes de usuarios medios

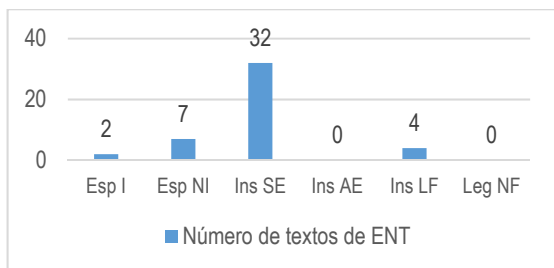


Gráfico 156: Número de textos de entrevistas según subemisor según investigador

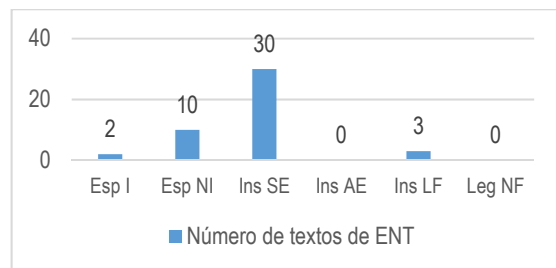


Gráfico 157: Número de textos de entrevistas según subemisor según grupo de informantes de traductores

Las diferencias más evidentes se encuentran en la revista o web de divulgación especializada (*vid. infra* Gráfico 158 a Gráfico 161). A pesar de la relevancia del subemisor instruido semiespecialista (y del especialista no investigador de acuerdo con el grupo de

estudiantes y de traductores) y de las similitudes entre estos dos últimos grupos, el investigador muestra una clara preferencia por el instruido semiespecialista.

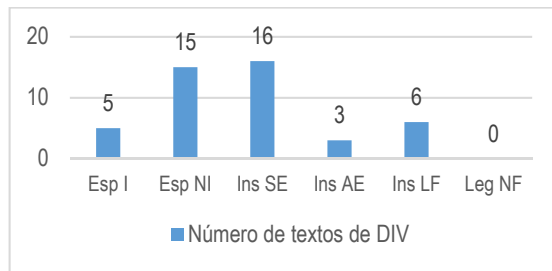


Gráfico 158: Número de textos de revistas o web de divulgación especializada según subemisor según grupo de informantes de estudiantes

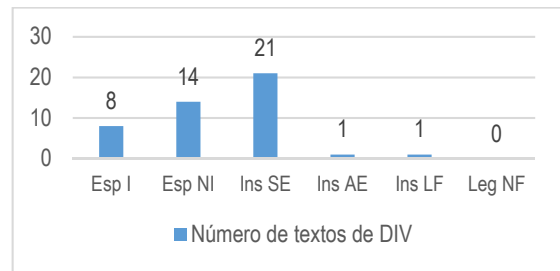


Gráfico 159: Número de textos de revistas o web de divulgación especializada según subemisor según grupo de informantes de usuarios medios

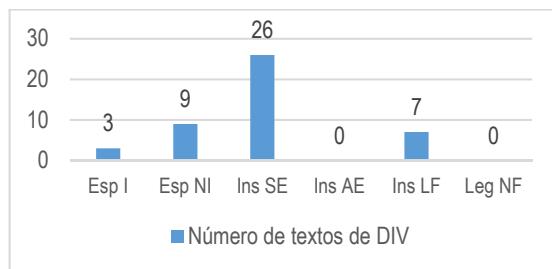


Gráfico 160: Número de textos de revistas o web de divulgación especializada según subemisor según investigador

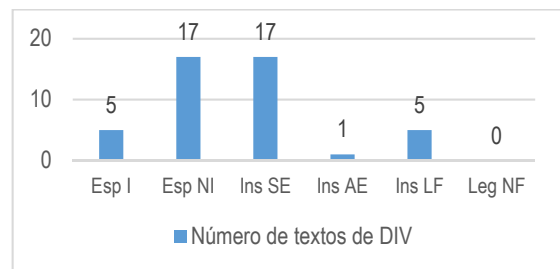


Gráfico 161: Número de textos de revistas o web de divulgación especializada según subemisor según grupo de informantes de traductores

2.3. Análisis del DESTINATARIO y del SUBDESTINATARIO

Las similitudes mencionadas anteriormente se mantienen en la categoría de subdestinatario. Asimismo, las similitudes entre informantes (y las divergencias entre grupos de ellos) también se advierten en el Gráfico 162. En el caso del especialista no existen diferencias relevantes. En el caso del instruido y el lego, a pesar de que las diferencias son reducidas, se advierte una relación pareja entre estudiante y usuario medio (que tienden a favor del destinatario instruido) y otra relación pareja entre el investigador y el traductor, que optan frecuentemente por el destinatario instruido, si bien destacan en el lego por encima de los otros dos grupos de informantes.

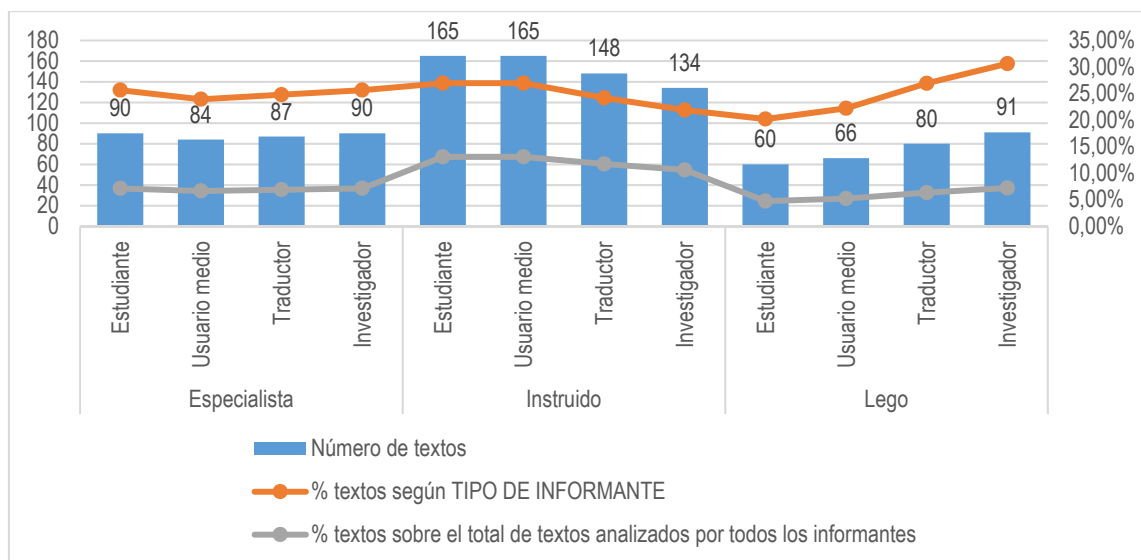


Gráfico 162: Datos de destinatario según tipo de informante

En cuanto a los datos de subdestinatario según informantes (*vid. infra* Gráfico 163), se observan similitudes entre todos los informantes en el caso del instruido lego formado, el más frecuente de todos los subdestinatarios. En el Gráfico 163, se distinguen los bloques mencionados con anterioridad: por una parte, las similitudes observadas en el investigador, y, por otra, las similitudes identificadas entre estudiante y usuario medio. Estos grupos cuentan con más similitudes que diferencias tanto al compararse en cada valor del subdestinatario como en la distribución de frecuencias para cada valor. Sí debe resaltarse que el grupo de estudiantes y de usuarios medios diverge más en el número de textos que cuenta con especialista investigador y especialista no investigador.

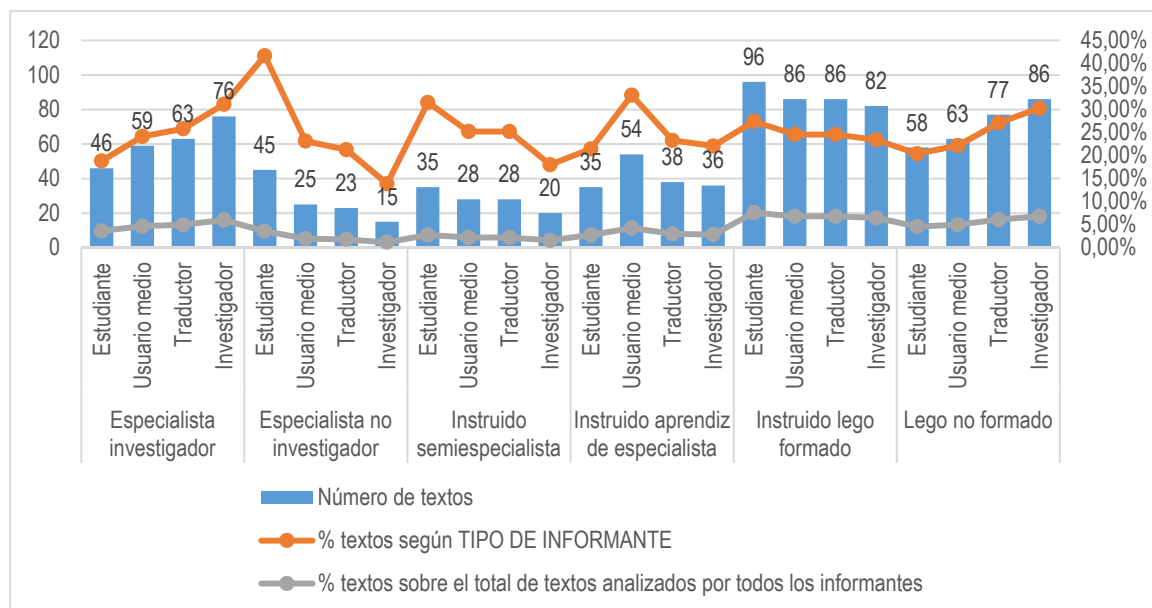


Gráfico 163: Datos de subdestinatario según tipo de informante

2.3.1. Vinculación con el TIPO DE TEXTO

Las relaciones entre subdestinatario y tipo de texto muestran que la percepción de los informantes deja de ser tan homogénea como en el caso del subemisor. Es cierto que existe un criterio homogéneo entre los diferentes informantes para los diversos tipos textuales, puesto que en todos los casos suelen coincidir tanto en los tipos de subdestinatarios con el máximo número de textos asignados como en la distribución del resto de subdestinatarios. No suelen coincidir, no obstante, en los valores numéricos con los que cuentan los subdestinatarios menos frecuentes.

En el caso del artículo en revista científica (*vid. infra* Gráfico 164 a Gráfico 167) y de la tesis doctoral (*vid. infra* Gráfico 168 a Gráfico 171) la dispersión de valores de subdestinatario va desde el investigador, más restringido al especialista investigador, hacia el estudiante, que asigna más textos al resto de subdestinatarios.

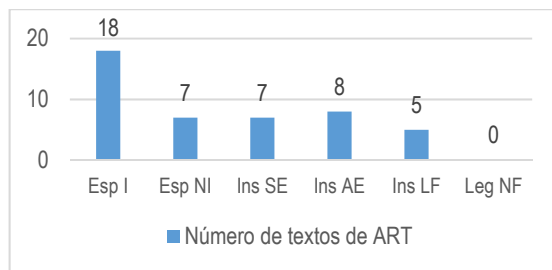


Gráfico 164: Número de textos de artículos en revista científica según subdestinatario según grupo de informantes de estudiantes

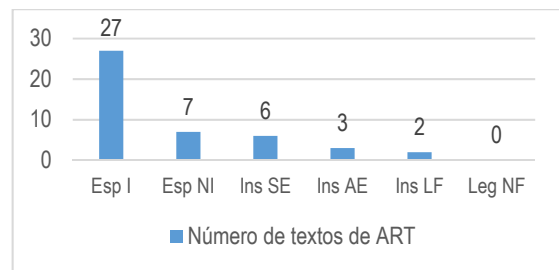


Gráfico 165: Número de textos de artículos en revista científica según subdestinatario según grupo de informantes de usuarios medios

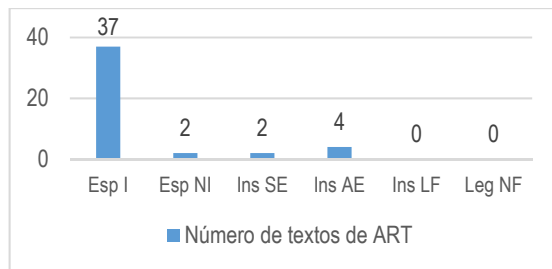


Gráfico 166: Número de textos de artículos en revista científica según subdestinatario según investigador

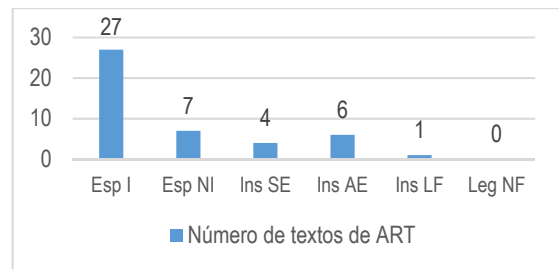


Gráfico 167: Número de textos de artículos en revista científica según subdestinatario según grupo de informantes de traductores

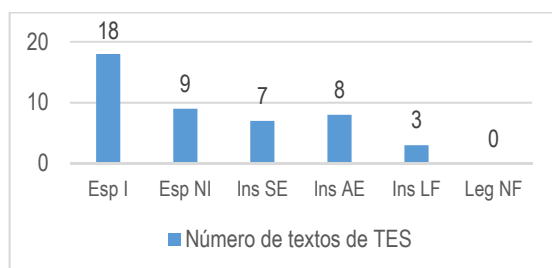


Gráfico 168: Número de textos de tesis doctorales según subdestinatario según grupo de informantes de estudiantes

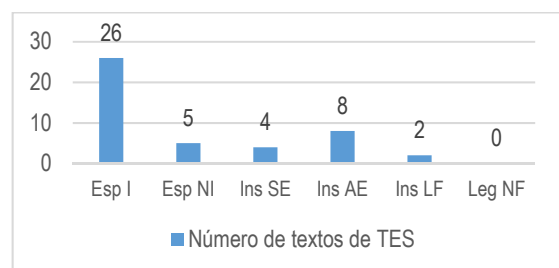


Gráfico 169: Número de textos de tesis doctorales según subdestinatario según grupo de informantes de usuarios medios

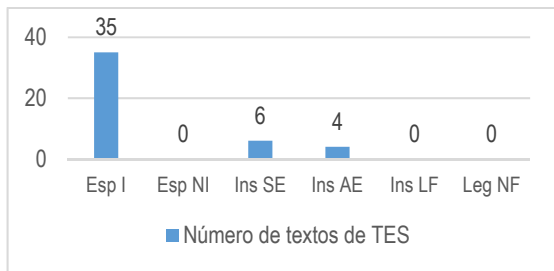


Gráfico 170: Número de textos de tesis doctorales según subdestinatario según investigador

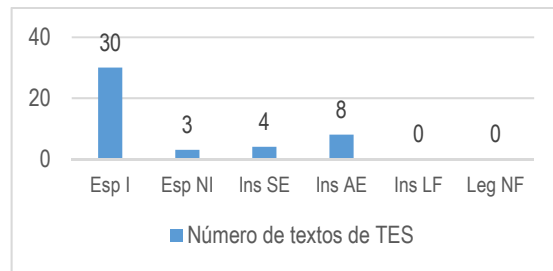


Gráfico 171: Número de textos de tesis doctorales según subdestinatario según grupo de informantes de traductores

Tras el investigador, el siguiente informante que restringe más el tipo de subdestinatario para el artículo en revista científica y la tesis doctoral es el traductor, que incorpora más textos para el instruido aprendiz de especialista y el especialista no investigador. Esta dispersión entre los valores de los subdestinatarios aumenta en el usuario medio y llega a su máximo nivel en el estudiante. En los casos más diferenciados, el estudiante y el investigador, los valores para subdestinatarios en artículo en revista científica (*vid. supra* Gráfico 164 y Gráfico 166) están íntimamente relacionados con sus respectivos valores en tesis doctoral (*vid. supra* Gráfico 168 y Gráfico 170).

En relación con la participación en foro (*vid. infra* Gráfico 172 a Gráfico 175), la distribución en todos los informantes es similar, si bien los datos proporcionados por el grupo de estudiantes divergen en el instruido lego formado a favor del especialista investigador (Gráfico 172).

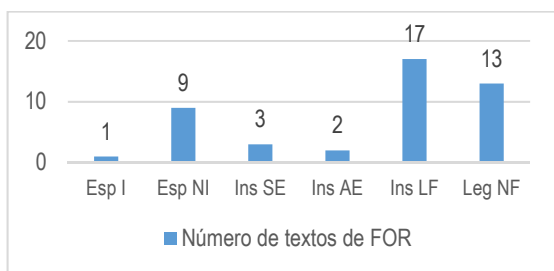


Gráfico 172: Número de textos de participaciones en foro según subdestinatario según grupo de informantes de estudiantes

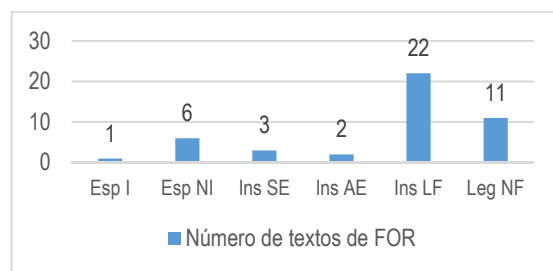


Gráfico 173: Número de textos de participaciones en foro según subdestinatario según grupo de informantes de usuarios medios

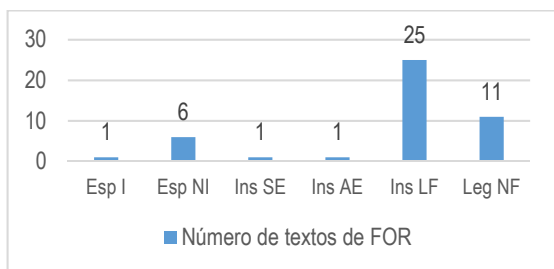


Gráfico 174: Número de textos de participaciones en foro según subdestinatario según investigador

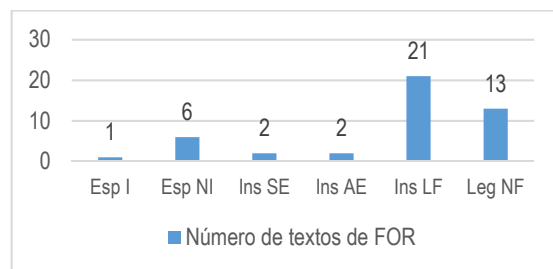


Gráfico 175: Número de textos de participaciones en foro según subdestinatario según grupo de informantes de traductores

Por otra parte, la coincidencia entre informantes también se mantiene para el tipo textual de noticia (*vid. infra* Gráfico 176 a Gráfico 179), siendo, de nuevo, el caso del grupo de estudiantes el que más diverge.

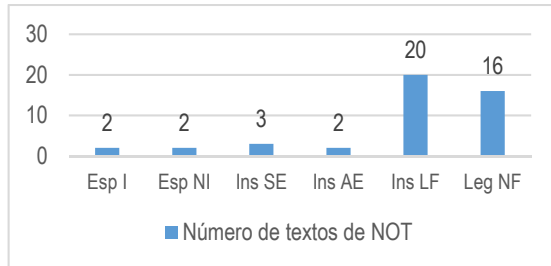


Gráfico 176: Número de textos de noticias según subdestinatario según grupo de informantes de estudiantes

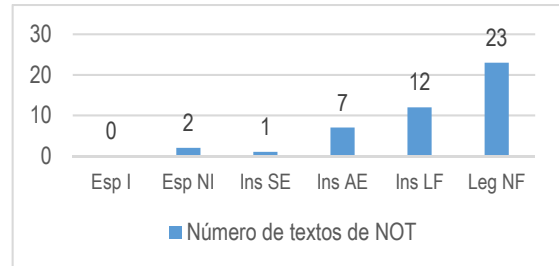


Gráfico 177: Número de textos de noticias según subdestinatario según grupo de informantes de usuarios medios

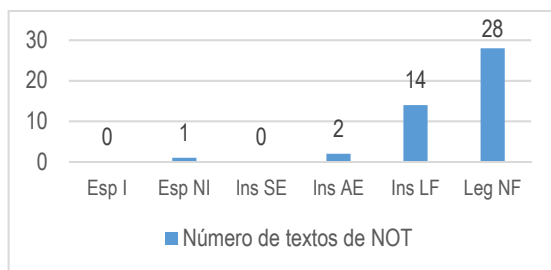


Gráfico 178: Número de textos de noticias según subdestinatario según investigador

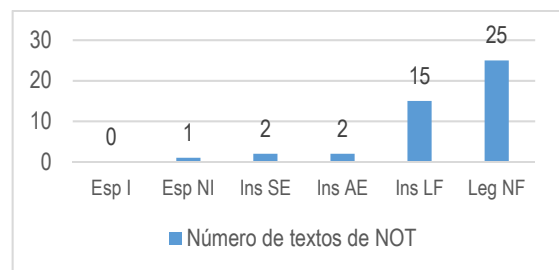


Gráfico 179: Número de textos de noticias según subdestinatario según grupo de informantes de traductores

Entre los tipos textuales con mayor número de valores para los subdestinatarios se encuentran la entrevista (*vid. infra* Gráfico 180 a Gráfico 183) y la revista o web de divulgación especializada (*vid. infra* Gráfico 184 a Gráfico 187). En estos gráficos se observa que las estrechas similitudes entre investigador y traductor en los dos tipos textuales se reiteran. En el caso del estudiante y del usuario medio, las similitudes solo se convergen en los subdestinatarios más frecuentes, es decir en el instruido lego formado y el lego no formado.

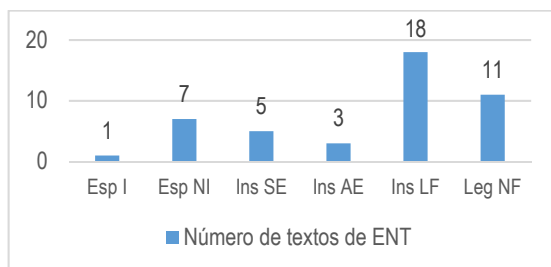


Gráfico 180: Número de textos de entrevistas según subdestinatario según grupo de informantes de estudiantes

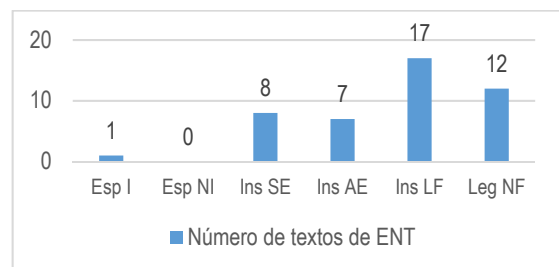


Gráfico 181: Número de textos de entrevistas según subdestinatario según grupo de informantes de usuarios medios

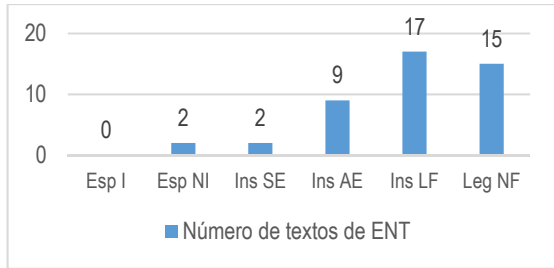


Gráfico 182: Número de textos de entrevistas según subdestinatario según investigador

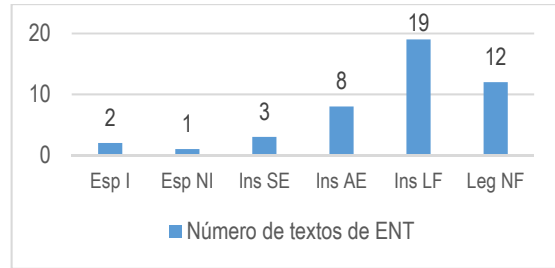


Gráfico 183: Número de textos de entrevistas según subdestinatario según grupo de informantes de traductores

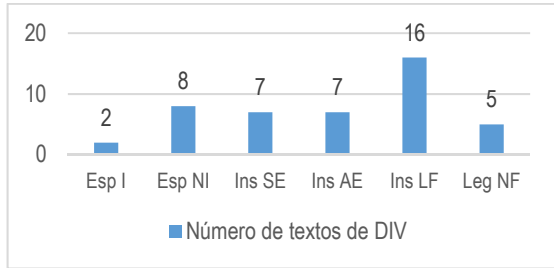


Gráfico 184: Número de textos de revistas o web de divulgación especializada según subdestinatario según grupo de informantes de estudiantes

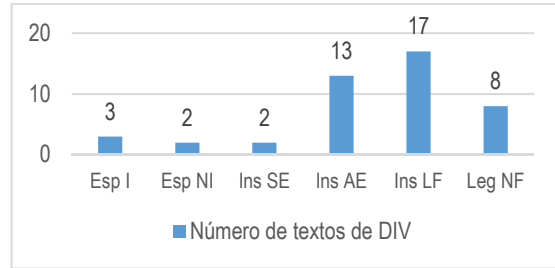


Gráfico 185: Número de textos de revistas o web de divulgación especializada según subdestinatario según grupo de informantes de usuarios medios

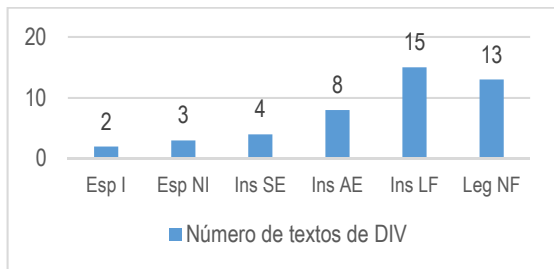


Gráfico 186: Número de textos de revistas o web de divulgación especializada según subdestinatario según investigador

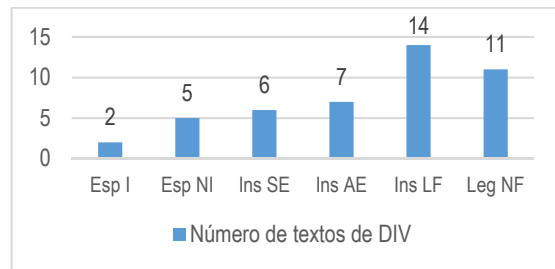


Gráfico 187: Número de textos de revistas o web de divulgación especializada según subdestinatario según grupo de informantes de traductores

Por último, el blog (*vid. infra* Gráfico 188 a Gráfico 191) constituye el caso más dispar entre los informantes, puesto que no suele coincidir ni la distribución de los valores ni los datos cuantitativos atribuidos a los mismos. Tan solo es posible defender que el protagonismo del subdestinatario en los cuatro grupos de informantes se halla en los subdestinatarios con menor competencia cognitivo-comunicativa.

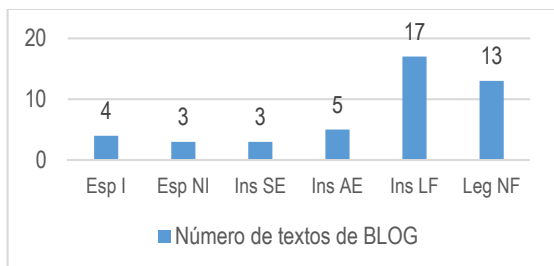


Gráfico 188: Número de textos de blogs según subdestinatario según grupo de informantes de estudiantes

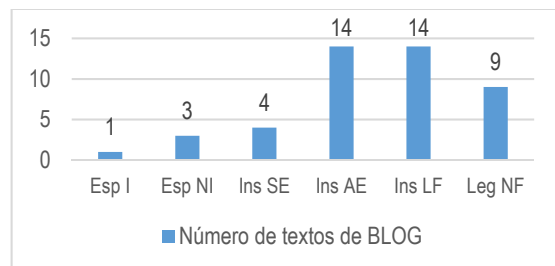


Gráfico 189: Número de textos de blogs según subdestinatario según grupo de informantes de usuarios medios

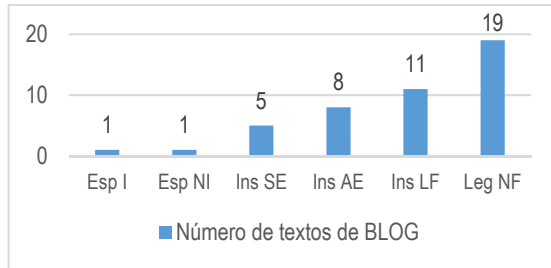


Gráfico 190: Número de textos de blogs según subdestinatario según investigador

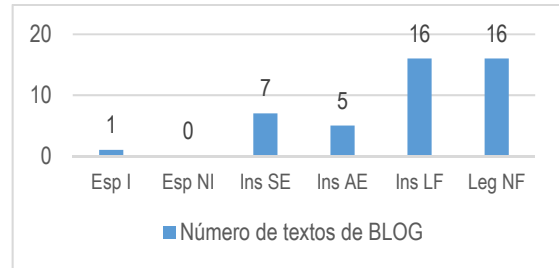


Gráfico 191: Número de textos de blogs según subdestinatario según grupo de informantes de traductores

2.4. Análisis de las FORMAS LÉXICAS ESPECIALIZADAS

Atendiendo al total de formas léxicas especializadas, las similitudes se restringen al investigador y al traductor, que muestran similitudes para los tres tipos de formas léxicas especializadas.

El estudiante es quien identifica mayor número de FLE délficas y el usuario medio el que menos (Gráfico 192). En la Tabla 132 y Tabla 133 se contrastan los datos que proporcionan estos dos grupos de informantes para un fragmento de un artículo en revista científica.

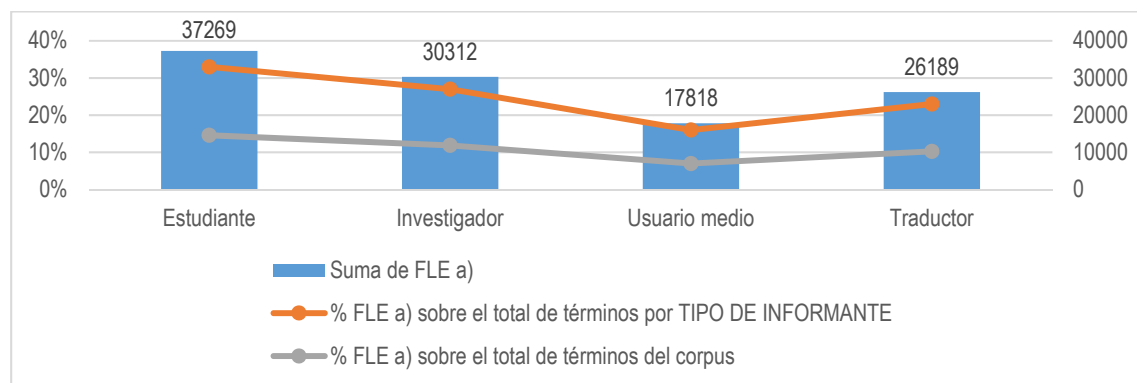


Gráfico 192: Datos de FLE a) según tipo de informante

Esta laguna legal nos hace plantearnos la posibilidad de que el régimen de relaciones personales entre abuelos y nietos tenga tal extensión, que pueda suponer una injerencia en el ejercicio de la patria potestad por los progenitores del menor. Y precisamente este es el problema planteado en la Sentencia del Tribunal Constitucional a la que antes se ha aludido, en la que se otorga a los abuelos maternos de dos menores de edad, un régimen de relaciones personales muy similar al que se concede al progenitor no custodio en los casos de crisis matrimonial.

DER ART11 (Estudiante)

Tabla 132: Ejemplo de texto con FLE a) según grupo de estudiantes

Esta laguna legal nos hace plantearnos la posibilidad de que el régimen de relaciones personales entre abuelos y nietos tenga tal extensión, que pueda suponer una injerencia en el ejercicio de la patria potestad por los progenitores del menor. Y precisamente este es el problema planteado en la Sentencia del Tribunal Constitucional a la que antes se ha aludido, en la que se otorga a los abuelos maternos de dos menores de

edad, un régimen de relaciones personales muy similar al que se concede al progenitor no custodio en los casos de crisis matrimonial.

DER ART11 (Usuario medio)

Tabla 133: Ejemplo de texto con FLE a) según grupo de usuarios medios

En dichas tablas se observa cómo las FLE semicrípticas sí se mantienen en los dos fragmentos pero no el número de formas léxicas especializadas frecuentes en el discurso común (i. e. las formas délficas) que identifican los usuarios medios. A nuestro juicio, el usuario medio no identifica que dichas formas léxicas sean especializadas, es decir, no considera que se active ningún valor especializado en dicho campo temático, sino que observa que se trata de formas léxicas no especializadas.

En el caso de las FLE semicrípticas (Gráfico 193), los datos observados anteriormente se invierten, puesto que el usuario medio es el que identifica mayor número de FLE semicrípticas y el estudiante el que menos.

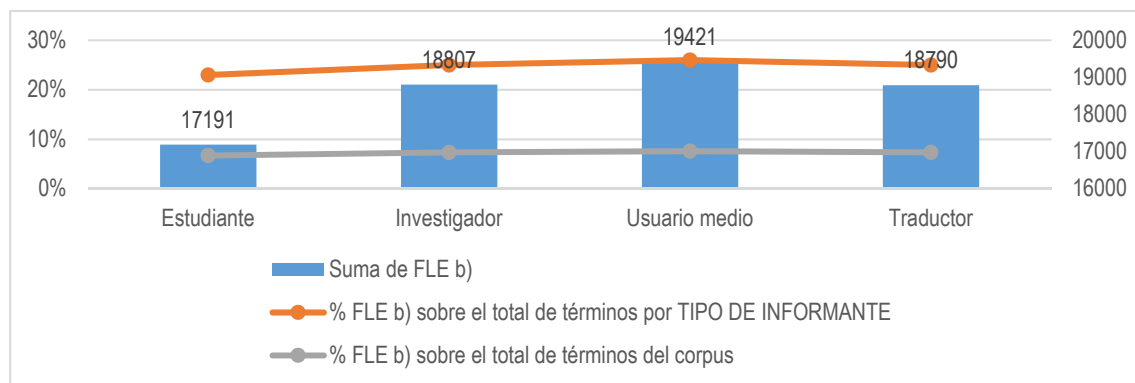


Gráfico 193: Datos de FLE b) según tipo de informante

Como se muestra en la Tabla 134 y en la Tabla 135, es evidente que el usuario medio clasifica las FLE délficas según el estudiante como FLE semicrípticas (por ejemplo, *incidencia*, *cultivo*, *amoxicilina* o *exantema típico*). Al igual sucede con algunas FLE b) según el estudiante, que pasan a ser FLE crípticas de acuerdo con el usuario medio, como *hiperemia*, *petequias*, *exudado amigdalár* y *recidivas*.

Resultados:

se obtuvieron 91 casos, resultando una incidencia de 3,2%, de los que 76 fueron confirmados microbiológicamente con test rápido o cultivo. La edad media fue 4,15 años. Los principales motivos de consulta fueron "fiebre y dolor de garganta" y "fiebre y erupción cutánea". Las alteraciones faringéas más frecuentes fueron la hiperemia y petequias en paladar, y en pocos pacientes se encontró exudado amigdalár. Casi un 40% de pacientes tenían síntomas catarrales, 71 pacientes presentaban un exantema típico, y 20 uno atípico. La mayoría se trató con amoxicilina o penicilina durante diez días; 15 pacientes tuvieron recidivas.

Conclusiones:

de los datos obtenidos destacan el gran número de casos, la presencia de síntomas catarrales y la poca frecuencia de exudado amigdalár. Fue llamativa la variabilidad de los exantemas con hallazgos como eritrodermia extensa, urticaria, exantema macular, petequias en localizaciones atípicas y edema facial y de miembros. El test rápido en Atención Primaria permite, por su utilidad, el diagnóstico de casos dudosos.

MED ART15 (Estudiante)

Tabla 134: Ejemplo de texto con FLE b) según grupo de estudiantes

Resultados:

se obtuvieron 91 casos, resultando una incidencia de 3,2%, de los que 76 fueron confirmados microbiológicamente con test rápido o cultivo. La edad media fue 4,15 años. Los principales motivos de consulta fueron "fiebre y dolor de garganta" y "fiebre y erupción cutánea". Las alteraciones faringéas más frecuentes fueron la hiperemia y petequias en paladar, y en pocos pacientes se encontró exudado amigdalár. Casi un 40% de pacientes tenían síntomas catarrales, 71 pacientes presentaban un exantema típico, y 20 uno atípico. La mayoría se trató con amoxicilina o penicilina durante diez días; 15 pacientes tuvieron recidivas.

Conclusiones:

de los datos obtenidos destacan el gran número de casos, la presencia de síntomas catarrales y la poca frecuencia de exudado amigdalár. Fue llamativa la variabilidad de los exantemas con hallazgos como eritrodermia extensa, urticaria, exantema macular, petequias en localizaciones atípicas y edema facial y de miembros. El test rápido en Atención Primaria permite, por su utilidad, el diagnóstico de casos dudosos.

MED ART15 (Usuario medio)

Tabla 135: Ejemplo de texto con FLE b) según grupo de usuarios medios

A pesar de estas diferencias, existen otros textos en los que las similitudes son más evidentes (*vid. infra* Tabla 136 y Tabla 137). En estos casos, los dos ejemplos muestran una clasificación terminológica más homogénea, posiblemente relacionada con la frecuencia de este tema o de las unidades de representación léxica en la vida cotidiana de los informantes.

Si refiere molestias músculo esqueléticas en los últimos doce meses especificando la localización (Columna cervical, hombros, codos, manos, columna dorsal, columna lumbar, caderas, rodillas, tobillos/pies).

Si refiere molestias músculo esqueléticas en los últimos siete meses especificando la localización (Columna cervical, hombros, codos, manos, columna dorsal, columna lumbar, caderas, rodillas, tobillos/pies).

Si durante los últimos doce meses previos al examen de salud ha cursado con incapacidad laboral transitoria derivada de padecer molestias músculo esqueléticas (Columna cervical, hombros, codos, manos, columna dorsal, columna lumbar, caderas, rodillas, tobillos/pies).

MED ART13 (Estudiante)

Tabla 136: Ejemplo de texto con FLE b) según grupo de estudiantes

Si refiere molestias músculo esqueléticas en los últimos doce meses especificando la localización (Columna cervical, hombros, codos, manos, columna dorsal, columna lumbar, caderas, rodillas, tobillos/pies).

Si refiere molestias músculo esqueléticas en los últimos siete meses especificando la localización (Columna cervical, hombros, codos, manos, columna dorsal, columna lumbar, caderas, rodillas, tobillos/pies).

Si durante los últimos doce meses previos al examen de salud ha cursado con incapacidad laboral transitoria derivada de padecer molestias músculo esqueléticas (Columna cervical, hombros, codos, manos, columna dorsal, columna lumbar, caderas, rodillas, tobillos/pies).

MED ART13 (Usuario medio)

Tabla 137: Ejemplo de texto con FLE b) según grupo de usuarios medios

En último lugar, en relación con las FLE crípticas, el usuario medio es el que identifica mayor número de estas unidades (Gráfico 194). En la Tabla 138 y la Tabla 139 se observa cómo las formas marcadas como FLE semicrípticas según el grupo de estudiantes son clasificadas como FLE crípticas por el usuario medio.

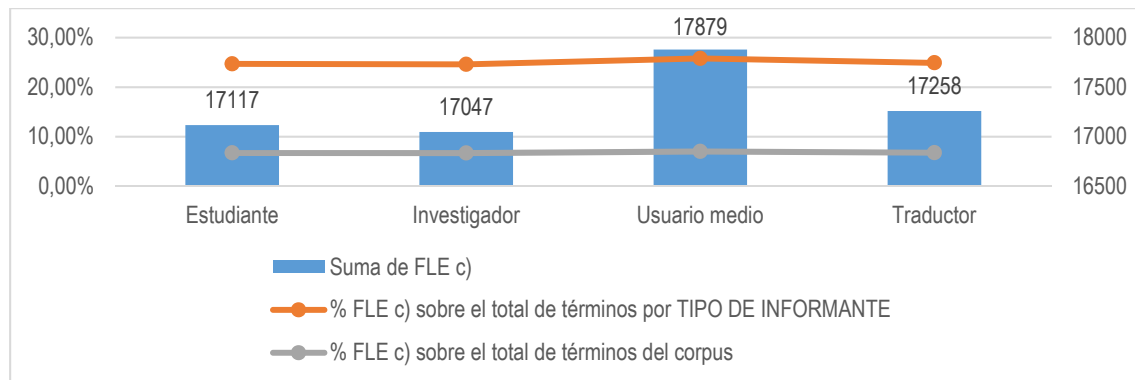


Gráfico 194: Datos de FLE c) según tipo de informante

Al igual que en la población general, en los pacientes en HD, la hiponatremia también se asocia a mayor riesgo de muerte⁷⁻¹². Esos estudios encuentran que las natremias prehemodiálisis bajas se asocian a diabetes, enfermedades neurológicas y psiquiátricas, mayor ganancia de peso interdiálisis (GPID) y mayor ultrafiltración (UF). La relación de las natremias con la mortalidad se mantiene teniendo en cuenta la evolución de las natremias en el tiempo⁷.

En pacientes en HD, no está claro si la relación hiponatremia/mortalidad, es causa-efecto o se explica por algún cofactor asociado. Está claramente demostrada la asociación entre hiponatremia y el deterioro de las funciones cerebrales^{6,13}. De hecho, la hiponatremia puede actuar como un tóxico cerebral directo, provocando encefalopatía. Al mismo tiempo, existen situaciones en las que la asociación hiponatremia/mortalidad es aún mayor, como ocurre en la cirrosis o en la insuficiencia cardíaca y que también son más prevalentes en los pacientes en HD¹⁴. De todas formas, la hiponatremia en HD no tiene por qué ser igual que a la que aparece en la población general.

MED ART07 (Estudiante)

Tabla 138: Ejemplo de texto con FLE c) según grupo de estudiantes

Al igual que en la población general, en los pacientes en HD, la hiponatremia también se asocia a mayor riesgo de muerte⁷⁻¹². Esos estudios encuentran que las natremias prehemodiálisis bajas se asocian a diabetes, enfermedades neurológicas y psiquiátricas, mayor ganancia de peso interdiálisis (GPID) y mayor ultrafiltración (UF). La relación de las natremias con la mortalidad se mantiene teniendo en cuenta la evolución de las natremias en el tiempo⁷.

En pacientes en HD, no está claro si la relación hiponatremia/mortalidad, es causa-efecto o se explica por algún cofactor asociado. Está claramente demostrada la asociación entre hiponatremia y el deterioro de las funciones cerebrales^{6,13}. De hecho, la hiponatremia puede actuar como un tóxico cerebral directo, provocando encefalopatía. Al mismo tiempo, existen situaciones en las que la asociación hiponatremia/mortalidad es aún mayor, como ocurre en la cirrosis o en la insuficiencia cardíaca y que también son más prevalentes en los pacientes en HD¹⁴. De todas formas, la hiponatremia en HD no tiene por qué ser igual que a la que aparece en la población general.

MED ART07 (Usuario medio)

Tabla 139: Ejemplo de texto con FLE c) según grupo de usuarios medios

El traductor muestra un criterio similar al del investigador en los tres tipos de FLE, si bien las FLE a) son más reducidas, por lo cual puede considerarse que el grupo de traductores no identifica un valor especializado en muchas de las unidades que son frecuentes en el discurso común, al menos no de la misma forma que las identifica el investigador. Este fenómeno queda ilustrado en la Tabla 140 y la Tabla 141 (*vid. infra*).

Carácter vinculante

Esta resolución del Teac, en unificación de criterio, tiene carácter vinculante para toda la Administración tributaria, tanto del Estado como de las Comunidades y Ciudades con Estatuto de Autonomía.

Este carácter vinculante se extiende tanto a los órganos de aplicación de los tributos como a los que tienen función revisora, tanto los criterios que con carácter reiterado fija el Teac, como las resoluciones de este mismo Tribunal dictadas, como sucede en el presente caso, en la resolución de recursos de alzada extraordinarios en unificación de criterio.

Hasta ahora, tan sólo había una sentencia judicial sobre la materia, pero solo afectaba a quienes habían reclamado la deducción de sus gastos por la actividad económica. Es un fallo del Tribunal Superior de Justicia de Madrid, de 10 de marzo de 2015, que si bien reconoce la deducción de los suministros, rechaza la deducción de los gastos del vehículo al considerar que puede usarse para actividades privadas.

DER DIV15 (Investigador)

Tabla 140: Ejemplo de texto con FLE a) según investigador

Carácter vinculante

Esta resolución del Teac, en unificación de criterio, tiene carácter vinculante para toda la Administración tributaria, tanto del Estado como de las Comunidades y Ciudades con Estatuto de Autonomía.

Este carácter vinculante se extiende tanto a los órganos de aplicación de los tributos como a los que tienen función revisora, tanto los criterios que con carácter reiterado fija el Teac, como las resoluciones de este mismo Tribunal dictadas, como sucede en el presente caso, en la resolución de recursos de alzada extraordinarios en unificación de criterio.

Hasta ahora, tan sólo había una sentencia judicial sobre la materia, pero solo afectaba a quienes habían reclamado la deducción de sus gastos por la actividad económica. Es un fallo del Tribunal Superior de Justicia de Madrid, de 10 de marzo de 2015, que si bien reconoce la deducción de los suministros, rechaza la deducción de los gastos del vehículo al considerar que puede usarse para actividades privadas.

DER DIV15 (Traductor)

Tabla 141: Ejemplo de texto con FLE c) según grupo de traductores

En general, el usuario medio es el que identifica mayor número de FLE semicríticas y FLE críticas, que son las clasificadas como menos frecuentes en el discurso común. El contraste de estos valores con la ausencia de FLE délficas

(especialmente si se comparan con el grupo de estudiantes) hace tener en cuenta el factor de referencia que toma cada grupo para regular su percepción. En efecto, a nuestro juicio el usuario medio asimila el valor especializado de una unidad léxica a su transparencia semántica a través de la morfología: cuanto más opaca es una unidad léxica, mayor probabilidad existe de que el usuario medio clasifique dicha unidad como especializada. En contraste con el estudiante, a nuestro juicio, este encuentra con mayor frecuencia diferentes formas léxicas especializadas entre las unidades del discurso común, debido a su reciente instrucción en terminología y acercamiento a la noción de valor especializado como sentido situado. Así, en definitiva, puede defenderse que los discursos prototípicos que toman como referencia para medir la frecuencia de una unidad en el discurso común son diferentes en el estudiante y en el usuario medio, como demuestran las diferencias perceptivas de estos casos (*vid. supra* Tabla 134 y Tabla 135, por un lado, y Tabla 138 y Tabla 139, por otro).

2.4.1. Vinculación con el TIPO DE TEXTO

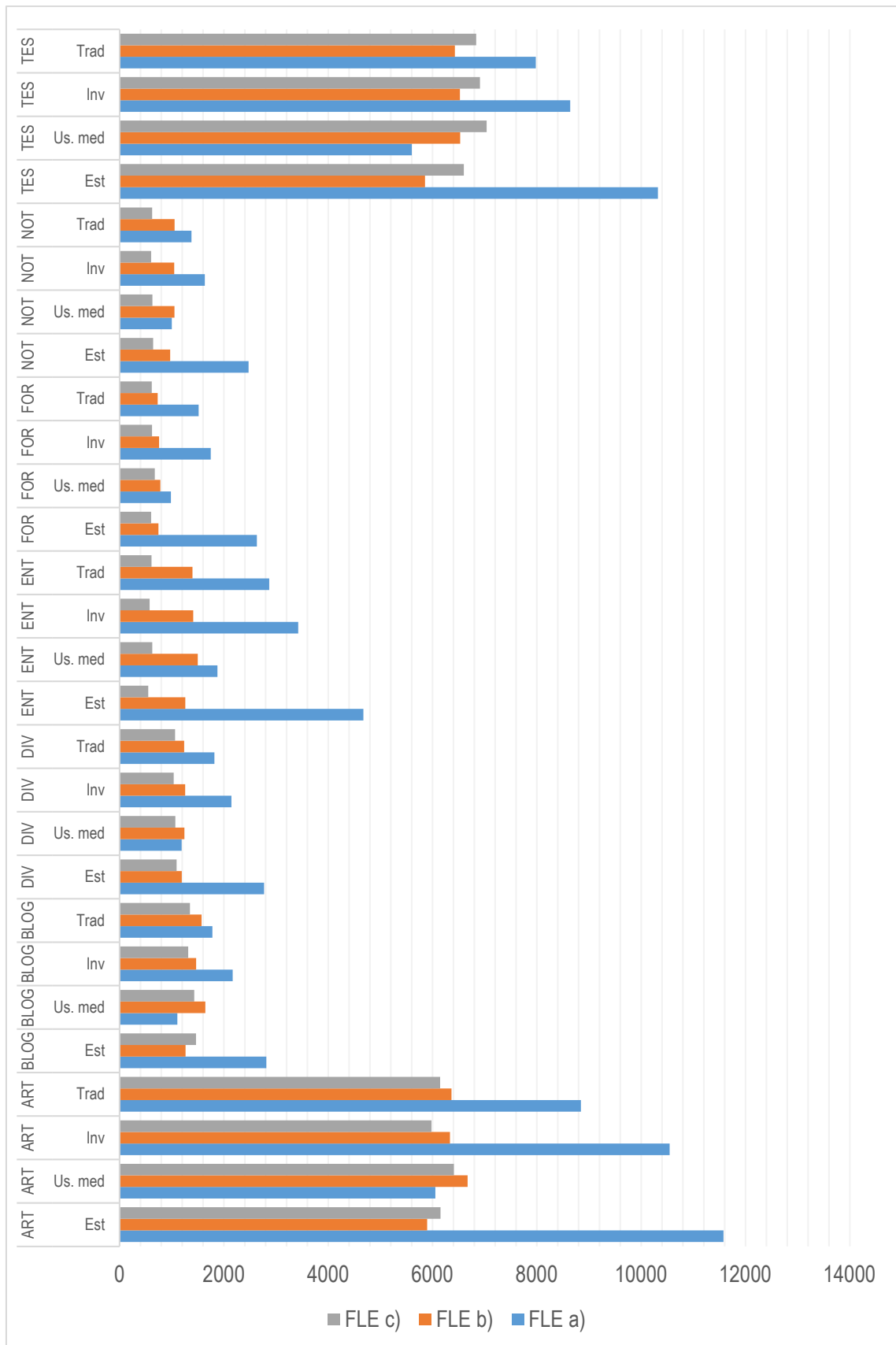


Gráfico 195: Número de tipos de FLE según tipo de texto según tipo de informante

El Gráfico 195 relaciona cada tipo de FLE con cada tipo de texto y cada grupo de informante. Más allá de las diferencias advertidas para usuario medio y estudiante en el apartado anterior (*vid. supra* IV. 2.4., p. 379), no es posible hallar diferencias relevantes entre informantes en cada tipo textual. La similitud en los valores numéricos coincide con nuestra hipótesis sobre la ausencia de diferencias perceptivas entre diferentes grupos sociales. Además, la relevancia de un tipo de FLE suele coincidir en todos los informantes para cada tipo textual. En efecto, en todos los tipos textuales la distribución de mayor a menor cantidad de FLE a) se organiza en este sentido: estudiantes > investigador > traductor > usuario medio. Este hecho podría permitir considerar que cada grupo de informantes percibe los tipos de formas léxicas especializadas de forma similar en cada tipo textual y que ningún grupo presenta una tendencia sobresaliente hacia un tipo textual en concreto.

2.5. Análisis del ÍNDICE DE DENSIDAD TERMINOLÓGICA

Los valores para el índice de la densidad terminológica no suelen divergir de forma marcada. De hecho, las diferencias que existen están vinculadas con las relaciones encontradas en el apartado IV. 2.4. En primer lugar, el usuario medio se encuentra por debajo de la media en el IDT a) porque se trata del grupo de informantes que identifica menos número de FLE a) (en contraste con el estudiante) (Gráfico 196). Por extensión, dada la relevancia de este tipo de FLE en el corpus, el usuario medio se mantiene por debajo de la media en el IDT a) (en contraste con el estudiante) (Gráfico 197). El caso inverso tiene lugar en el Gráfico 198, donde el usuario medio se halla por encima de la media del IDT b) (no obstante, la similitud entre los cuatro grupos de informantes es clara). Por último, el caso del IDT c) (Gráfico 199) mantiene una diferencia milesimal y queda justificada por la tendencia del investigador por las FLE b) en lugar de por las FLE c).

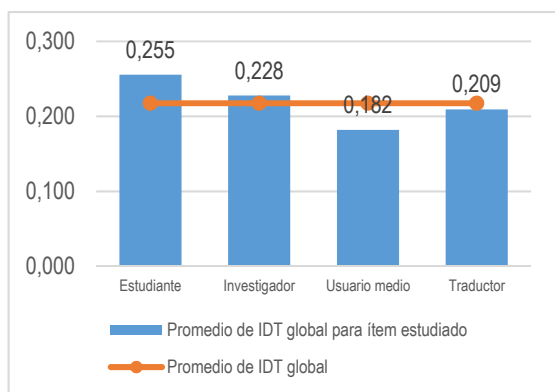


Gráfico 196: Promedio de IDT global según tipo de informante y diferencias con promedio de IDT global

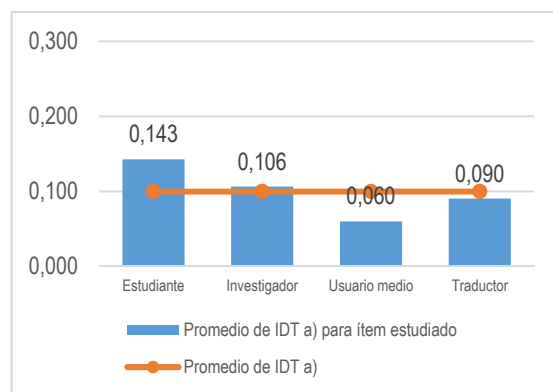


Gráfico 197: Promedio de IDT a) según tipo de informante y diferencias con promedio de IDT a)

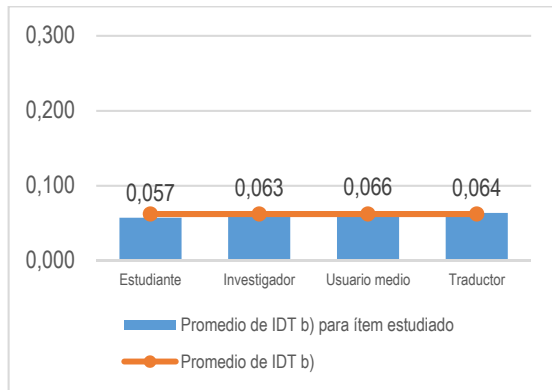


Gráfico 198: Promedio de IDT b) según tipo de informante y diferencias con promedio de IDT b)

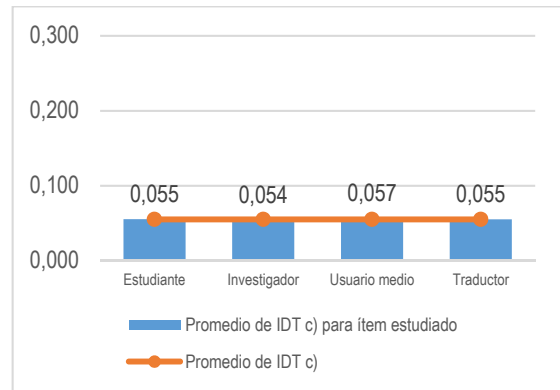


Gráfico 199: Promedio de IDT c) según tipo de informante y diferencias con promedio de IDT c)

2.5.1. Vinculación con el TEMA

En el Gráfico 200, se identifican las mismas (y escasas) diferencias entre temas observadas en el Gráfico 142 (*vid. supra* apartado IV. 1.4.1., p. 308). Además, los diferentes grupos de informantes comparten un criterio común, especialmente en IDT de términos semicríticos e IDT de términos críticos, puesto que en los tres temas los cuatro grupos de informantes muestran valores muy similares, a excepción de los identificados previamente (*vid. supra* apartado IV. 2.5.).

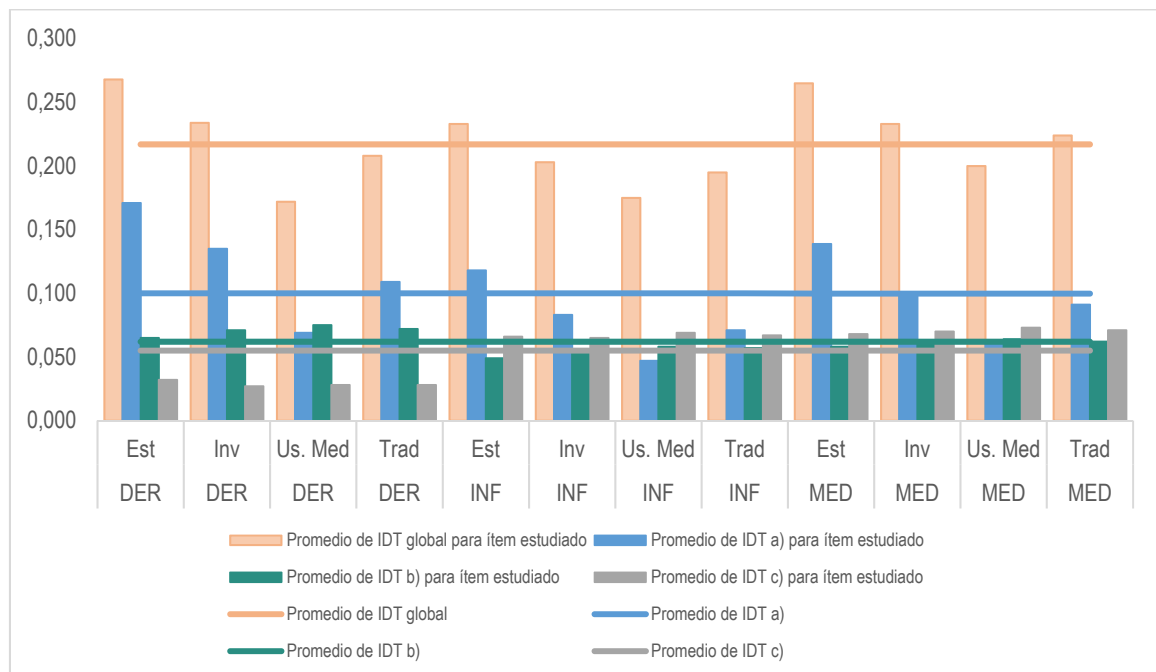


Gráfico 200: Promedio de tipos de IDT según informantes según temas y diferencia con promedios de IDT

2.5.2. Vinculación con el TIPO DE TEXTO

La prevalencia del IDT de términos délficos por parte del grupo de informantes de estudiante no se vincula estrictamente con algún tipo textual concreto, sino que se trata de una tendencia generalizada: para los estudiantes, la percepción del valor especializado de las formas léxicas no parece depender estrictamente del tipo textual en el que se inscribe el contenido, sino, más bien del destinatario del texto. No obstante, a la luz de los datos para el IDT de términos délficos en todos los informantes (a excepción del usuario medio, como ya se mencionó previamente), se observa que los tipos textuales clasificados principalmente como especializados (es decir, el artículo en revista científica y la tesis doctoral) muestran valores equilibrados (*vid. infra* Gráfico 202 y Gráfico 206). Así, sería posible determinar que los tipos textuales clasificados en el máximo grado de especialización condicionan la percepción del valor especializado de las formas léxicas de un texto.

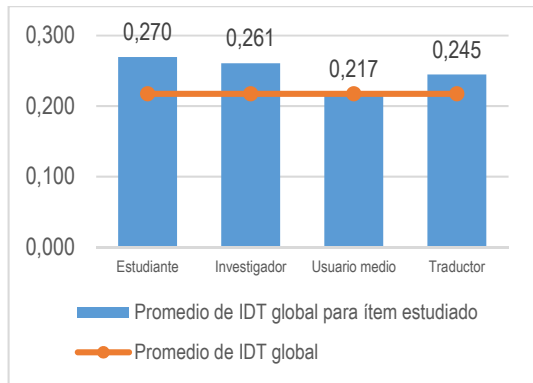


Gráfico 201: Promedio de IDT global para artículo en revista científica según informantes y diferencia con promedio de IDT global

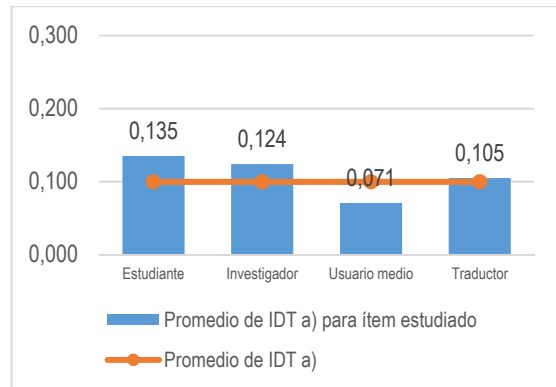


Gráfico 202: Promedio de IDT a) para artículo en revista científica según informantes y diferencia con promedio de IDT a)

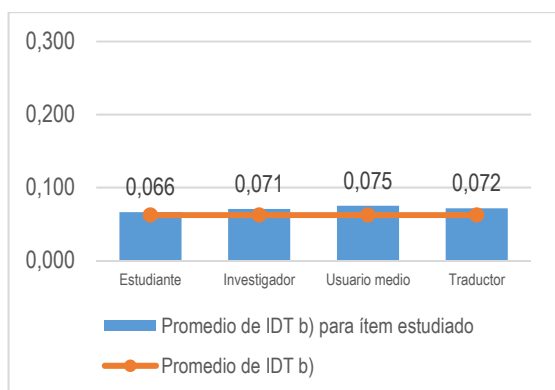


Gráfico 203: Promedio de IDT b) para artículo en revista científica según informantes y diferencia con promedio de IDT b)

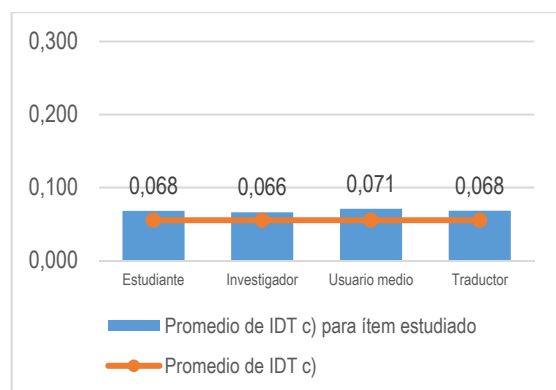


Gráfico 204: Promedio de IDT c) para artículo en revista científica según informantes y diferencia con promedio de IDT c)

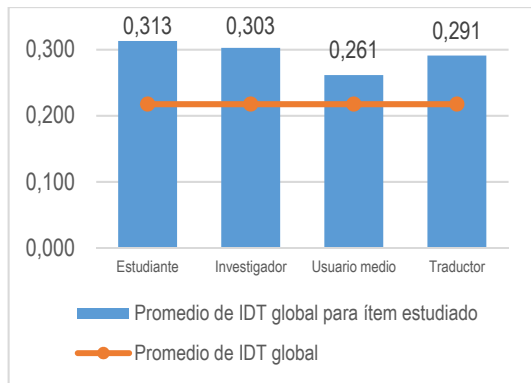


Gráfico 205: Promedio de IDT global para tesis doctoral según informantes y diferencia con promedio de IDT global

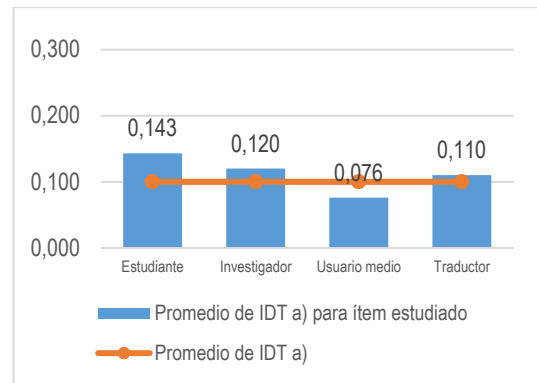


Gráfico 206: Promedio de IDT a) para tesis doctoral según informantes y diferencia con promedio de IDT a)

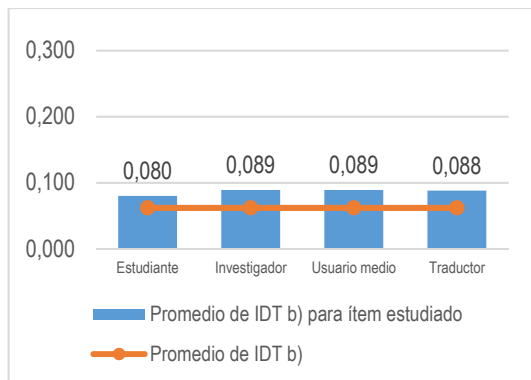


Gráfico 207: Promedio de IDT b) para tesis doctoral según informantes y diferencia con promedio de IDT b)

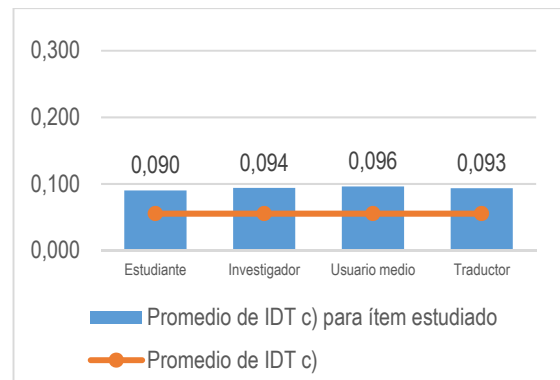


Gráfico 208: Promedio de IDT c) para tesis doctoral según informantes y diferencia con promedio de IDT c)

En el resto de casos, las diferencias suelen corresponder con las analizadas previamente (*vid. supra* apartado IV. 2.5.). Los informantes suelen diferir en el IDT de términos délficos.

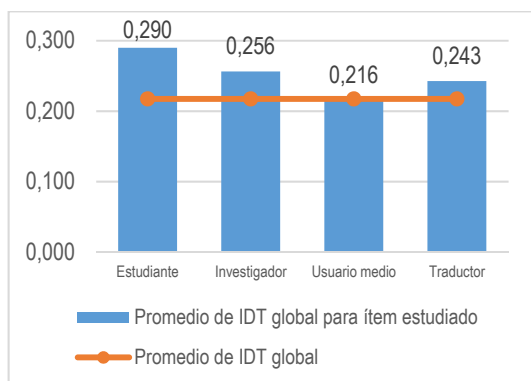


Gráfico 209: Promedio de IDT global para blog según informantes y diferencia con promedio de IDT global

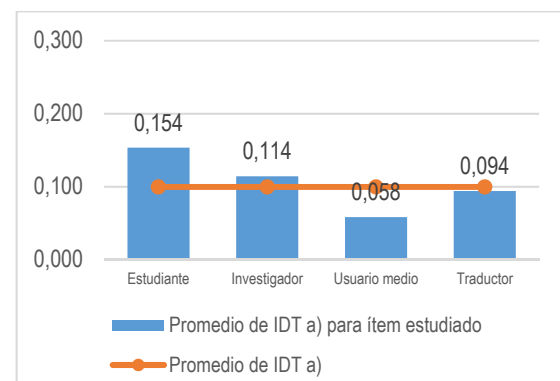


Gráfico 210: Promedio de IDT a) para blog según informantes y diferencia con promedio de IDT a)

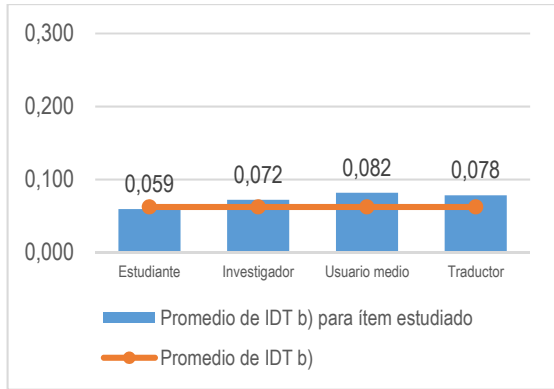


Gráfico 211: Promedio de IDT b) para blog según informantes y diferencia con promedio de IDT b)

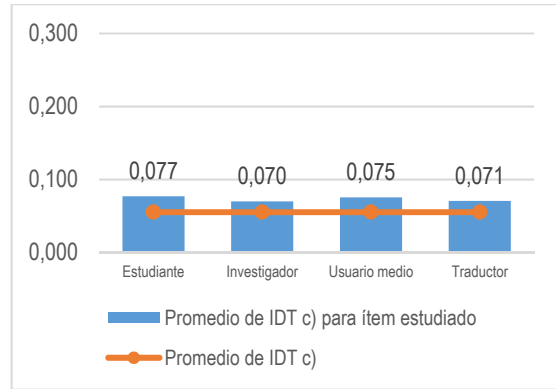


Gráfico 212: Promedio de IDT c) para blog según informantes y diferencia con promedio de IDT c)

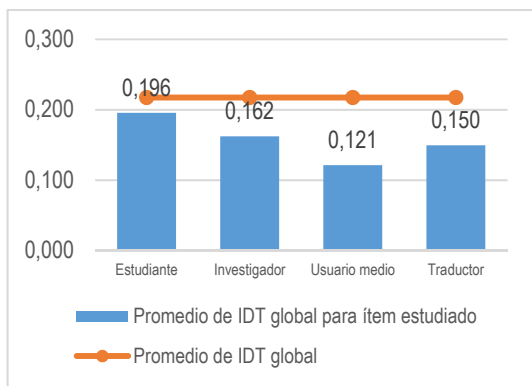


Gráfico 213: Promedio de IDT global para entrevista según informantes y diferencia con promedio de IDT global

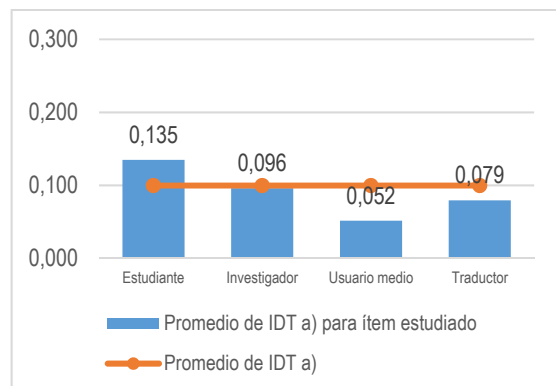


Gráfico 214: Promedio de IDT a) para entrevista según informantes y diferencia con promedio de IDT a)

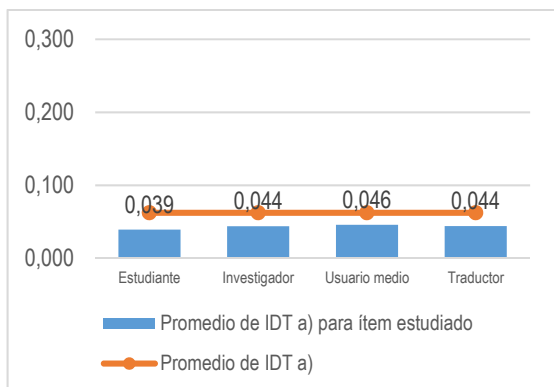


Gráfico 215: Promedio de IDT b) para entrevista según informantes y diferencia con promedio de IDT b)

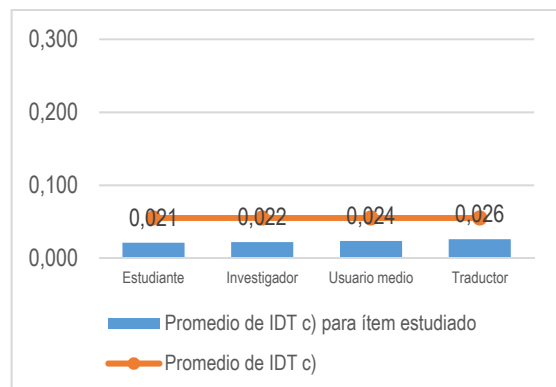


Gráfico 216: Promedio de IDT c) para entrevista según informantes y diferencia con promedio de IDT c)

Como se comprueba en los casos para el blog y para la entrevista (también en los tipos textuales siguientes) el esquema de distribución siempre es el mismo: divergencias notables en el IDT de formas délficas y un criterio común en el IDT de formas semicrípticas y crípticas. Así, puede defenderse que el rasgo singular o prominente de estos textos es la opacidad semántica o falta de frecuencia en el discurso común.

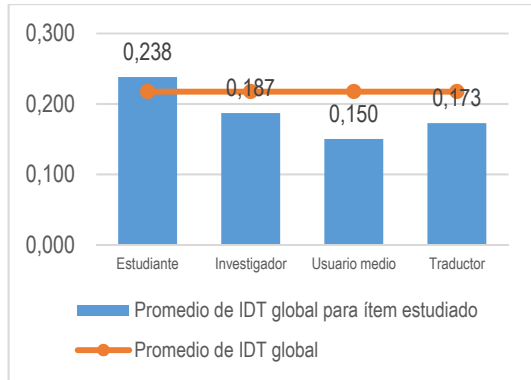


Gráfico 217: Promedio de IDT global para noticia según informantes y diferencia con promedio de IDT global

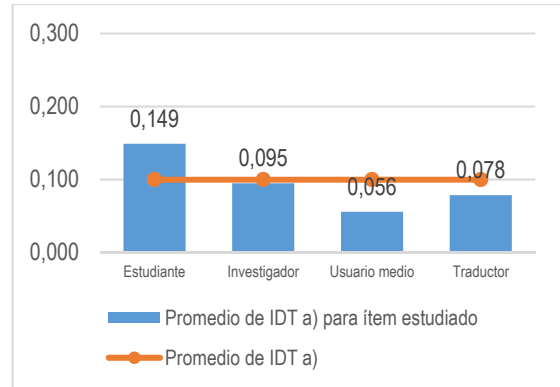


Gráfico 218: Promedio de IDT a) para noticia según informantes y diferencia con promedio de IDT a)

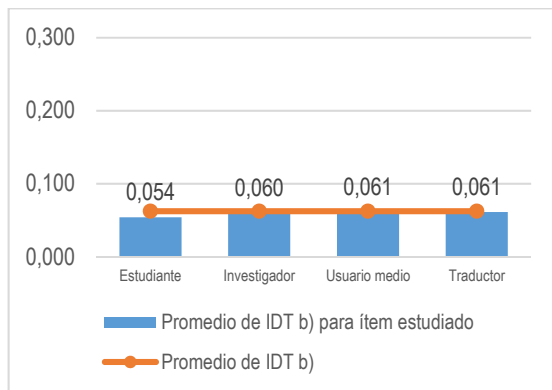


Gráfico 219: Promedio de IDT b) para noticia según informantes y diferencia con promedio de IDT b)

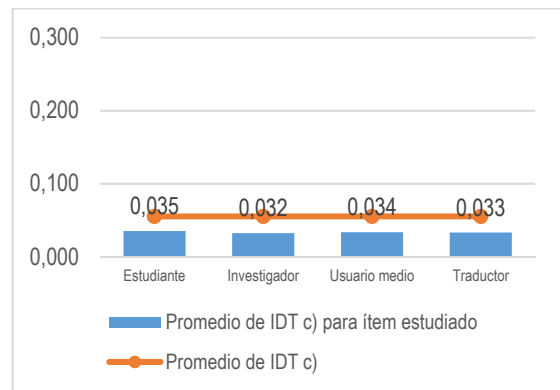


Gráfico 220: Promedio de IDT c) para noticia según informantes y diferencia con promedio de IDT c)

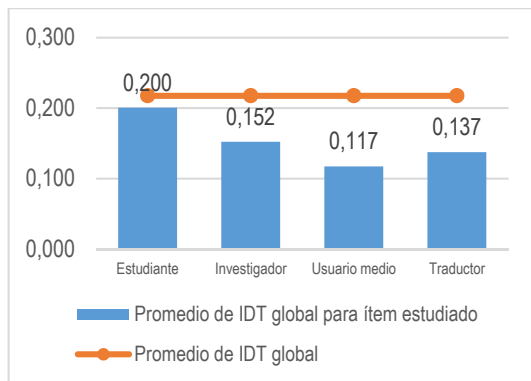


Gráfico 221: Promedio de IDT global para participación en foro según informantes y diferencia con promedio de IDT global

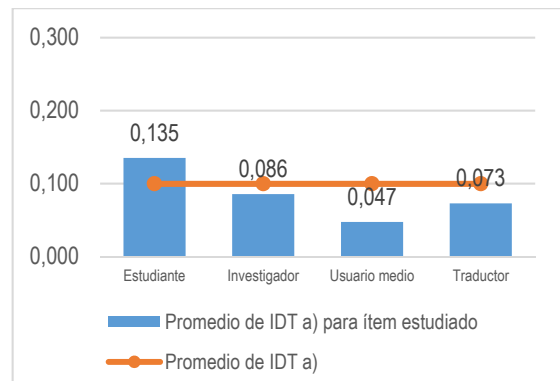


Gráfico 222: Promedio de IDT a) para participación en foro según informantes y diferencia con promedio de IDT a)

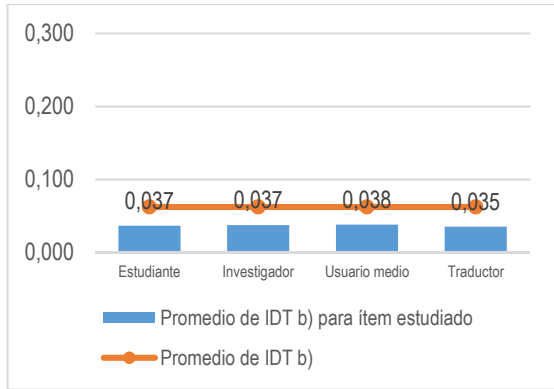


Gráfico 223: Promedio de IDT b) para participación en foro según informantes y diferencia con promedio de IDT b)

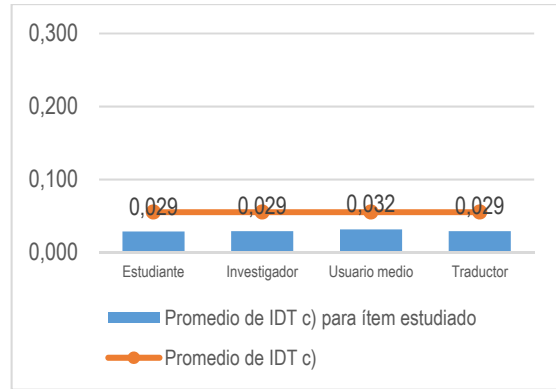


Gráfico 224: Promedio de IDT c) para participación en foro según informantes y diferencia con promedio de IDT c)

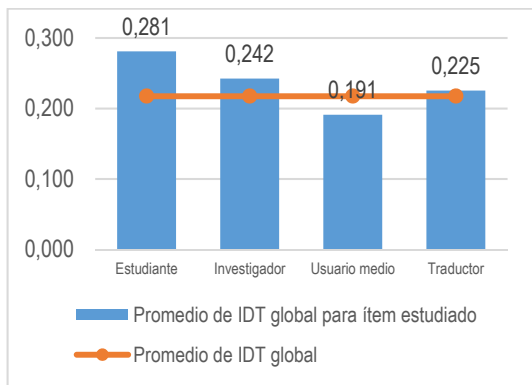


Gráfico 225: Promedio de IDT global para revista o web de divulgación especializada según informantes y diferencia con promedio de IDT global

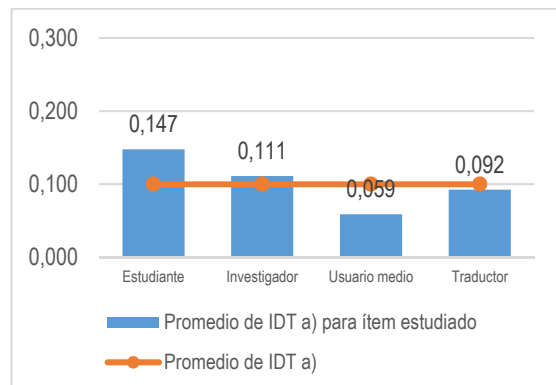


Gráfico 226: Promedio de IDT a) para revista o web de divulgación especializada según informantes y diferencia con promedio de IDT a)

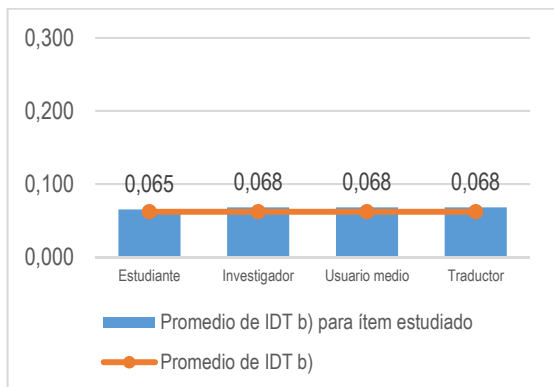


Gráfico 227: Promedio de IDT b) para revista o web de divulgación especializada según informantes y diferencia con promedio de IDT b)

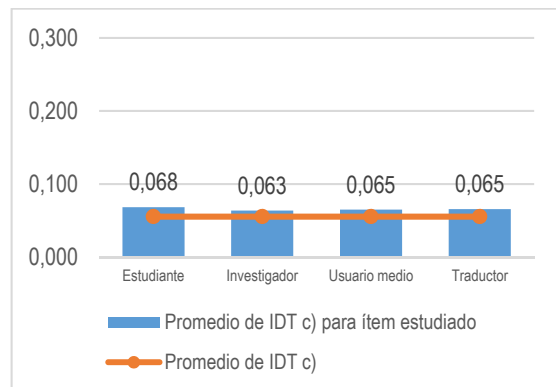


Gráfico 228: Promedio de IDT c) para revista o web de divulgación especializada según informantes y diferencia con promedio de IDT c)

Las similitudes observadas entre los cuatro grupos de informantes en los siete tipos textuales para la densidad terminológica de formas semicrípticas y formas crípticas permite defender que la percepción de las formas léxicas consideradas especializadas es compartida cuando estos términos son poco frecuentes en el discurso común (es decir, cuando la

prominencia de estos rasgos es estrechamente asociada a la comunicación especializada). Es decir, en efecto, la consideración de los grupos de informantes para la activación del valor especializado en las formas léxicas tiene relación con su falta de reconocimiento como formas léxicas comunes o frecuentes, ya sea en su denominación, ya sea en la comprensión de su sentido situado.

2.5.3. Vinculación con la CLASE

La clase tampoco parece incidir significativamente en la percepción de todos los informantes aunque sí se puede describir una relación entre el texto semiespecializado y no especializado y el IDT de términos délficos en el grupo de estudiantes (*vid. infra* Gráfico 229).

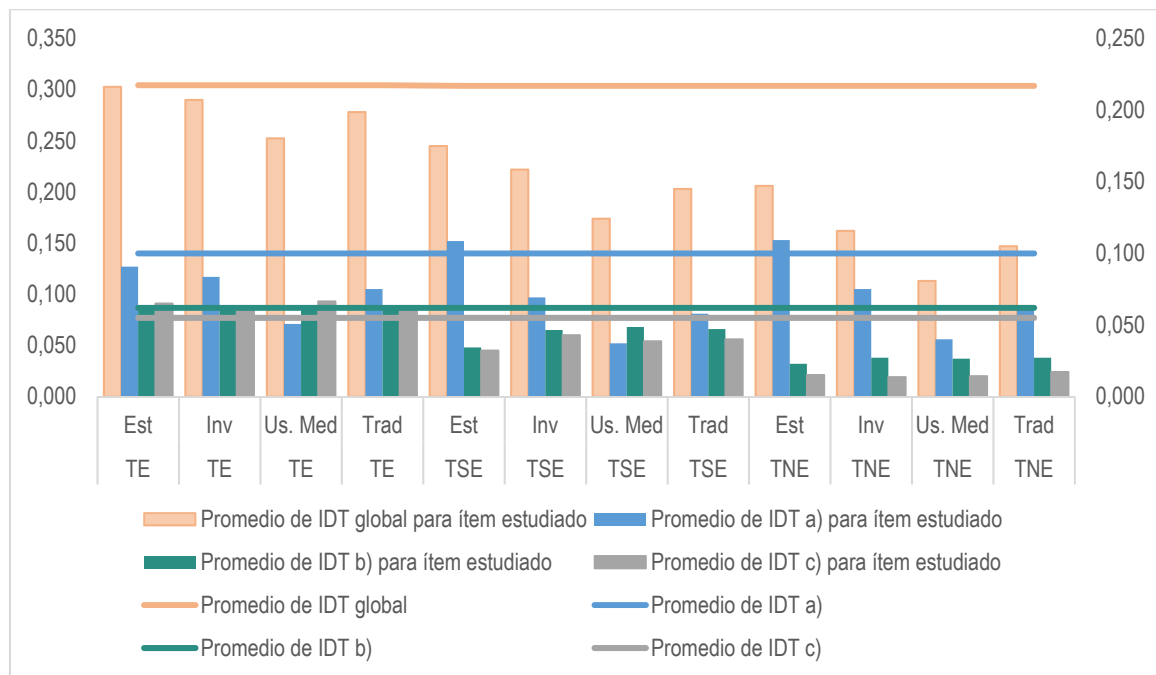


Gráfico 229: Promedio de tipos de IDT según informantes según clases y diferencia con promedios de IDT

En primer lugar, en el texto especializado todos los informantes muestran valores similares. La principal diferencia se halla a favor del IDT de términos délficos, tanto para el grupo de estudiantes como para el investigador.

En el caso del texto semiespecializado y no especializado, destacan los valores de IDT de términos délficos según el grupo de estudiantes. Esto se debe, a nuestro juicio, a que el grupo de estudiantes no considera el criterio de la prominencia formal como determinante de la activación del valor especializado de las unidades léxicas, lo que le permite identificar o considerar mayor número de términos délficos, especialmente en

aquellos textos en los que no destacan los términos crípticos o semicrípticos, como podrían ser el texto semiespecializado y el texto no especializado.

2.5.4. Vinculación con el EMISOR y el SUBEMISOR

Los grupos de informantes tampoco muestran diferencias relevantes en los diferentes IDT para los diversos tipos de subemisores. Suele despuntar el usuario medio, que siempre muestra diferencias con respecto a los otros informantes, ya que suele encontrarse por debajo del nivel que les corresponde para IDT global y para IDT de términos delficos. El IDT de términos semicrípticos y el IDT de términos crípticos siempre muestran valores equilibrados en los cuatro grupos de informantes en los seis subemisores, como se comentó en secciones anteriores.

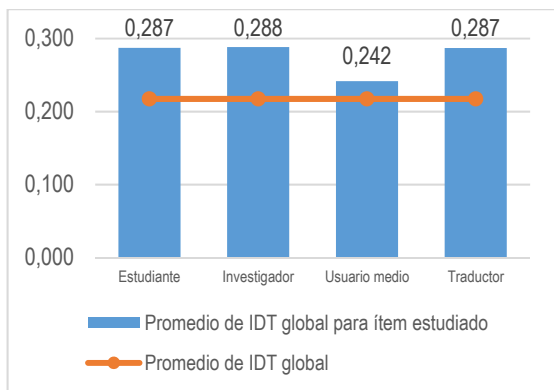


Gráfico 230: Promedio de IDT global para subemisor especialista investigador según informantes y diferencia con promedio de IDT global

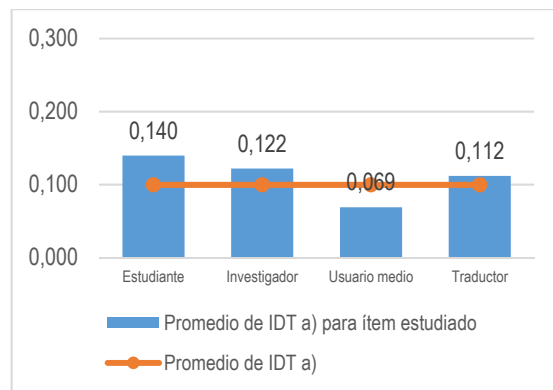


Gráfico 231: Promedio de IDT a) para subemisor especialista investigador según informantes y diferencia con promedio de IDT a)

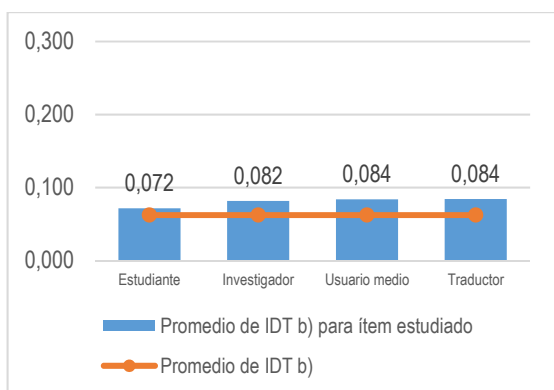


Gráfico 232: Promedio de IDT b) para subemisor especialista investigador según informantes y diferencia con promedio de IDT b)

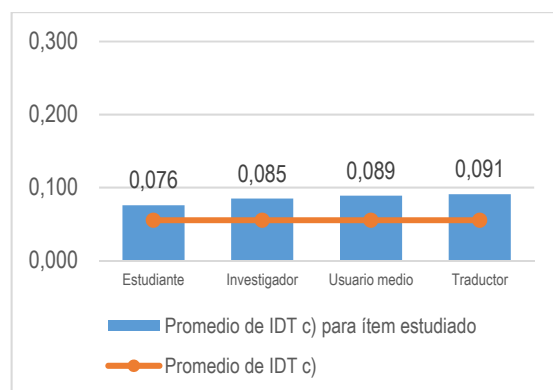


Gráfico 233: Promedio de IDT c) para subemisor especialista investigador según informantes y diferencia con promedio de IDT c)

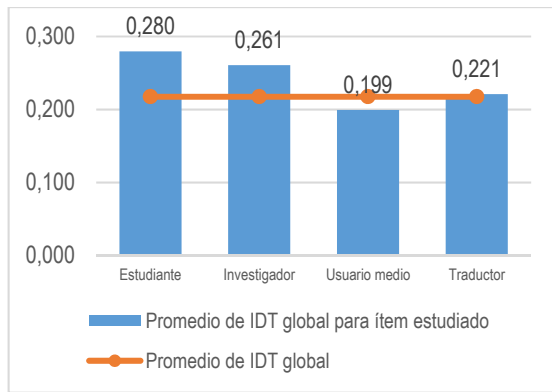


Gráfico 234: Promedio de IDT global para subemisor especialista no investigador según informantes y diferencia con promedio de IDT global

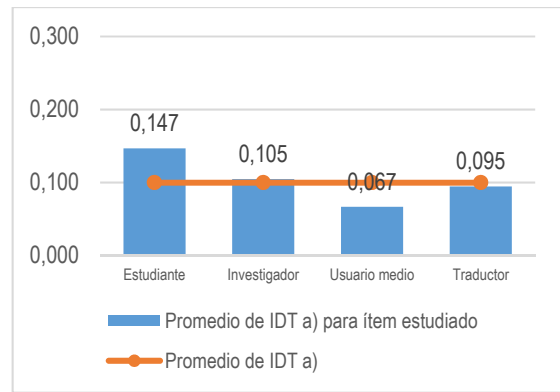


Gráfico 235: Promedio de IDT a) para subemisor especialista no investigador según informantes y diferencia con promedio de IDT a)

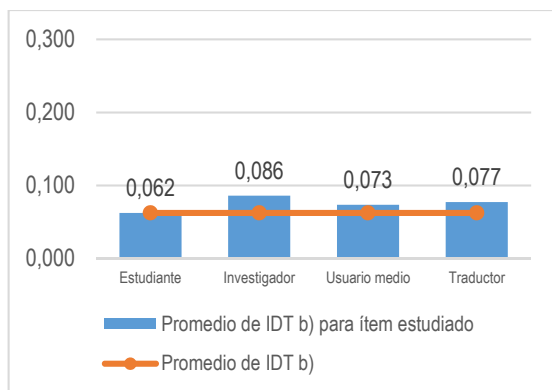


Gráfico 236: Promedio de IDT b) para subemisor especialista no investigador según informantes y diferencia con promedio de IDT b)

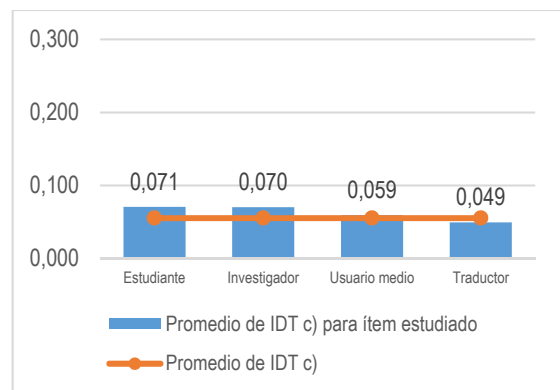


Gráfico 237: Promedio de IDT c) para subemisor especialista no investigador según informantes y diferencia con promedio de IDT c)

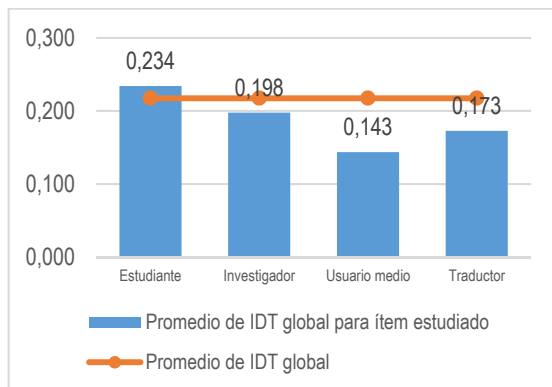


Gráfico 238: Promedio de IDT global para subemisor instruido semiespecialista según informantes y diferencia con promedio de IDT global

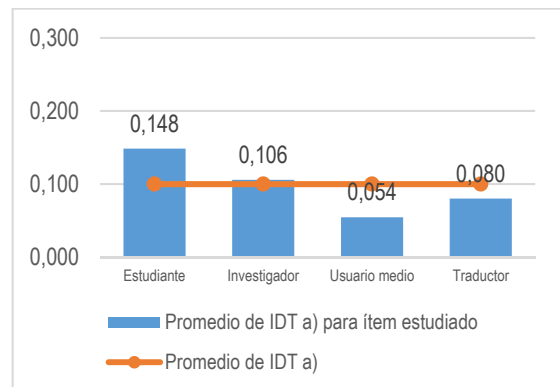


Gráfico 239: Promedio de IDT a) para subemisor instruido semiespecialista según informantes y diferencia con promedio de IDT a)

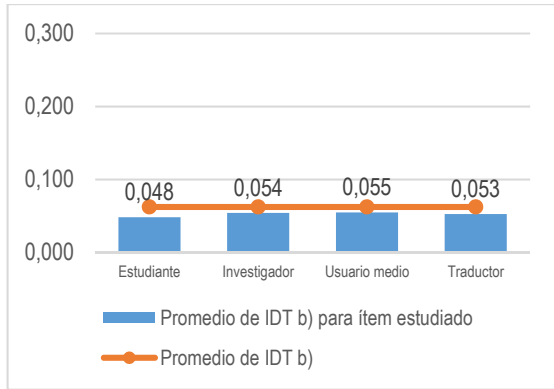


Gráfico 240: Promedio de IDT b) para subemisor instruido semiespecialista según informantes y diferencia con promedio de IDT b)

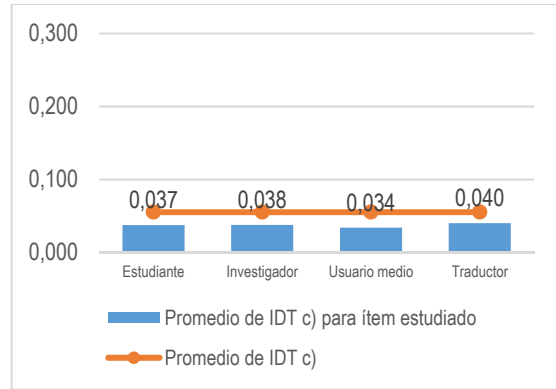


Gráfico 241: Promedio de IDT c) para subemisor instruido semiespecialista según informantes y diferencia con promedio de IDT c)

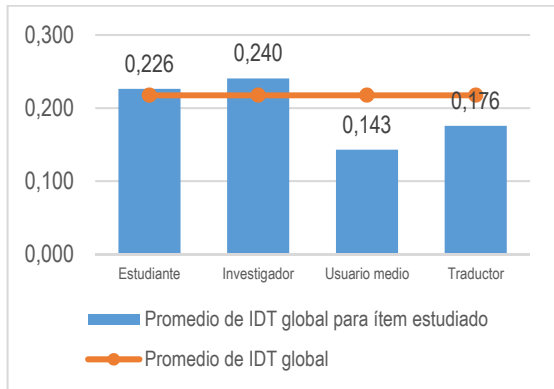


Gráfico 242: Promedio de IDT global para subemisor instruido aprendiz de especialista según informantes y diferencia con promedio de IDT global

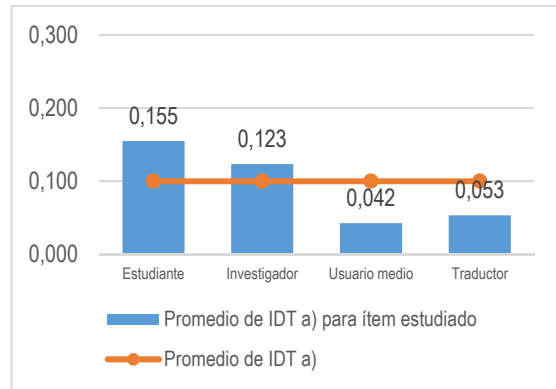


Gráfico 243: Promedio de IDT a) para subemisor instruido aprendiz de especialista según informantes y diferencia con promedio de IDT a)

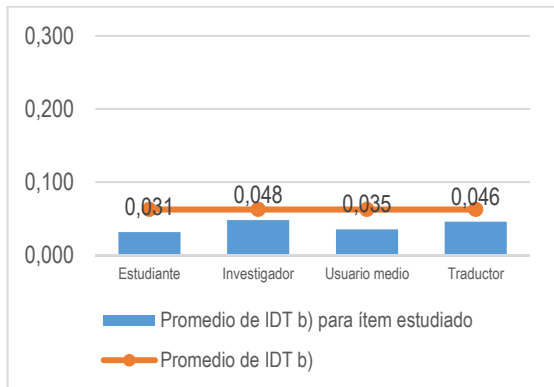


Gráfico 244: Promedio de IDT b) para subemisor instruido aprendiz de especialista según informantes y diferencia con promedio de IDT b)

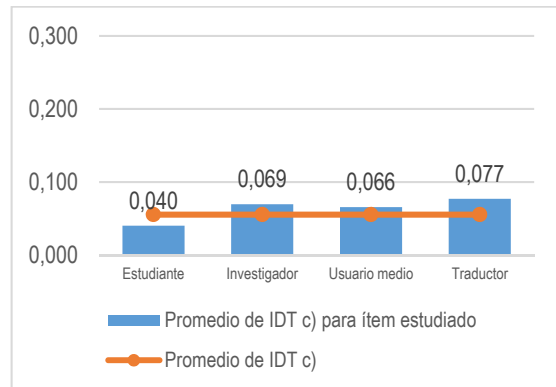


Gráfico 245: Promedio de IDT c) para subemisor instruido aprendiz de especialista según informantes y diferencia con promedio de IDT c)

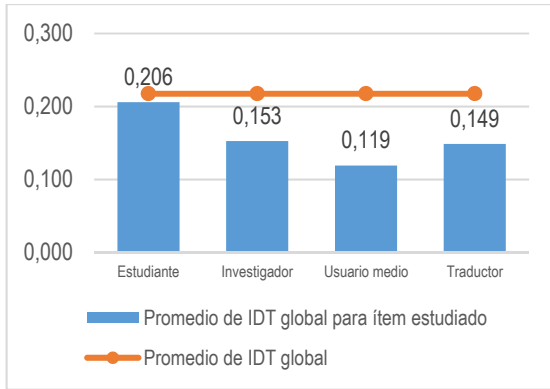


Gráfico 246: Promedio de IDT global para subemisor instruido lego formado según informantes y diferencia con promedio de IDT global

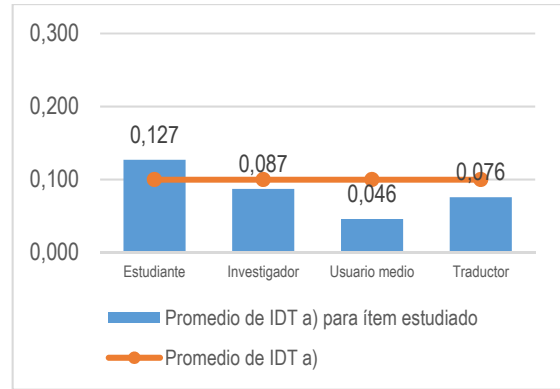


Gráfico 247: Promedio de IDT a) para subemisor instruido lego formado según informantes y diferencia con promedio de IDT a)

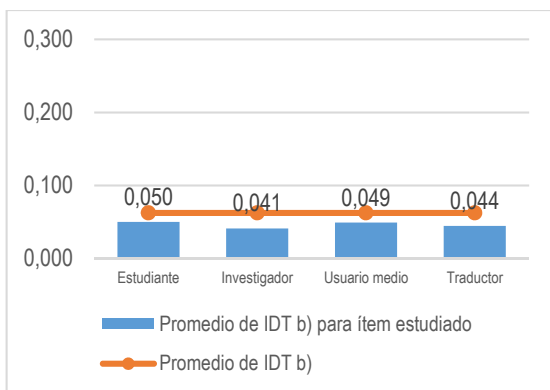


Gráfico 248: Promedio de IDT b) para subemisor instruido lego formado según informantes y diferencia con promedio de IDT b)

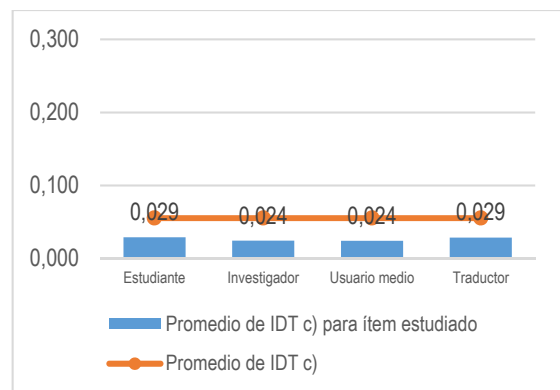


Gráfico 249: Promedio de IDT c) para subemisor instruido lego formado según informantes y diferencia con promedio de IDT c)

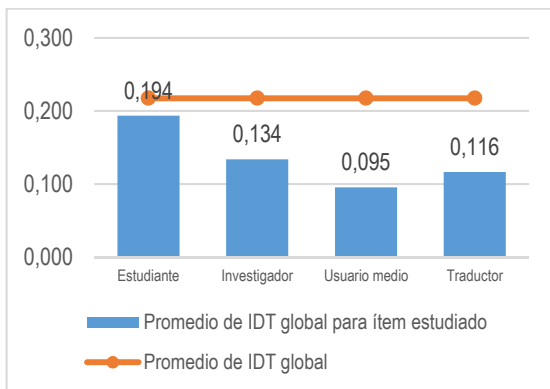


Gráfico 250: Promedio de IDT global para subemisor lego no formado según informantes y diferencia con promedio de IDT global

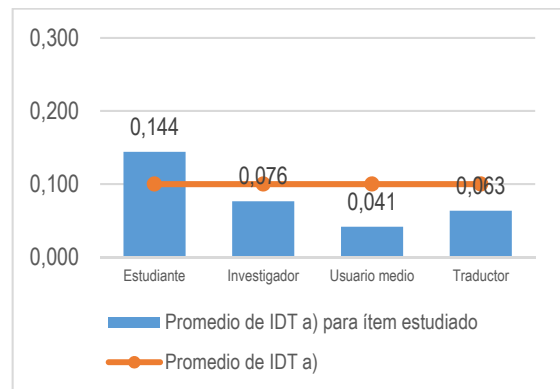


Gráfico 251: Promedio de IDT a) para subemisor lego no formado según informantes y diferencia con promedio de IDT a)

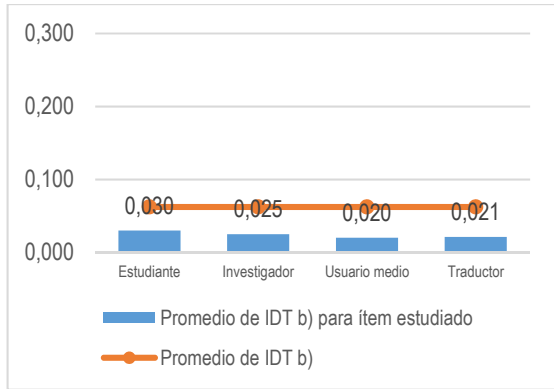


Gráfico 252: Promedio de IDT b) para subemisor lego no formado según informantes y diferencia con promedio de IDT b)

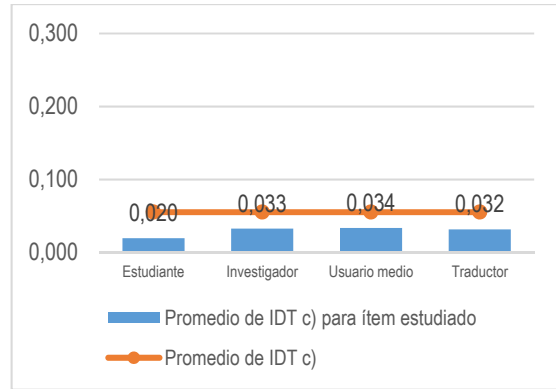


Gráfico 253: Promedio de IDT c) para subemisor lego no formado según informantes y diferencia con promedio de IDT c)

El estudiante muestra unos valores inferiores en el IDT c) en los casos del subemisor aprendiz de especialista y lego no formado (Gráfico 245 y Gráfico 253). Esta diferencia puede estar relacionada con el hecho de que se tratan de dos perfiles poco frecuentes en las clasificaciones propuestas por el grupo de estudiantes, así como que algunos de los términos clasificados como crípticos por el resto de los grupos es asociado a la categoría de términos semicrípticos por los estudiantes (*vid. supra* IV. 2.4. p. 379).

2.5.5. Vinculación con el DESTINATARIO y SUBDESTINATARIO

El subdestinatario presenta las mismas características que el subemisor (*vid. supra* apartado IV. 2.5.4.): valores equilibrados para IDT b) e IDT c) entre informantes y un ligero descenso en los valores de IDT global e IDT a) para el grupo de usuarios medios.

Debe mencionarse que el caso del subdestinatario instruido semiespecialista (Gráfico 262 a Gráfico 265) muestra unos valores menos equilibrados para el IDT c) (aun así, las diferencias son poco representativas).

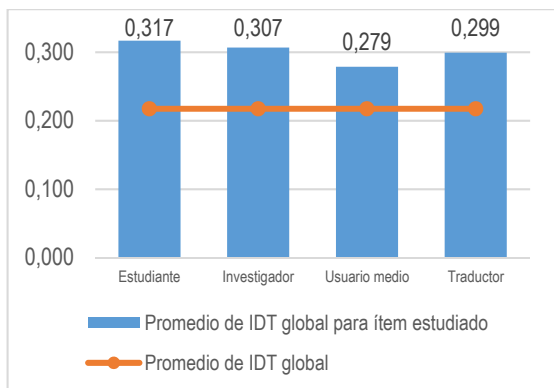


Gráfico 254: Promedio de IDT global para subdestinatario especialista investigador según informantes y diferencia con promedio de IDT global

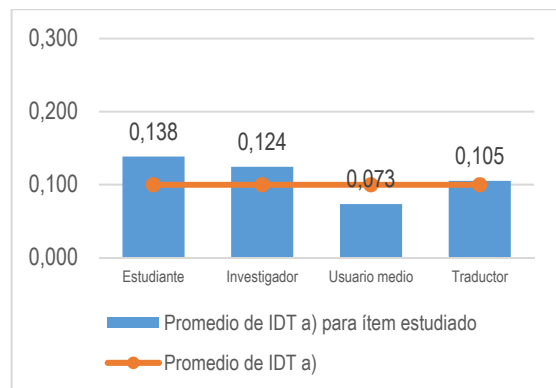


Gráfico 255: Promedio de IDT a) para subdestinatario especialista investigador según informantes y diferencia con promedio de IDT a)

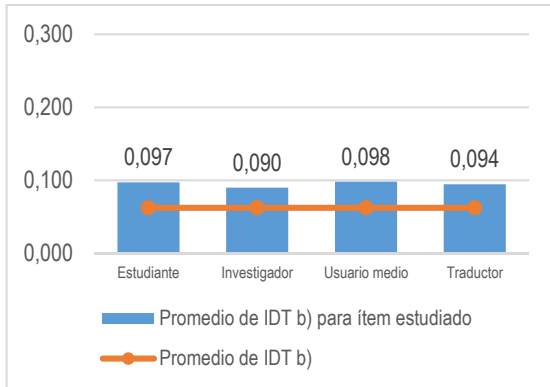


Gráfico 256: Promedio de IDT b) para subdestinatario especialista investigador según informantes y diferencia con promedio de IDT b)

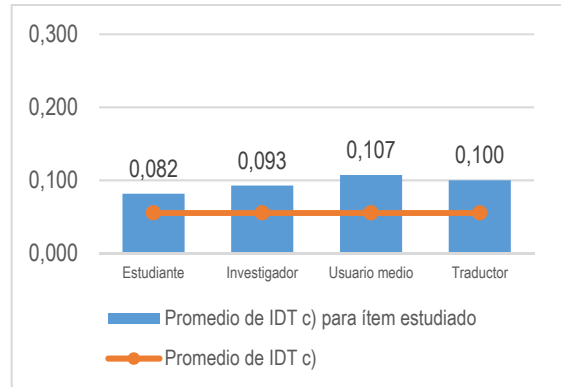


Gráfico 257: Promedio de IDT c) para subdestinatario especialista investigador según informantes y diferencia con promedio de IDT c)

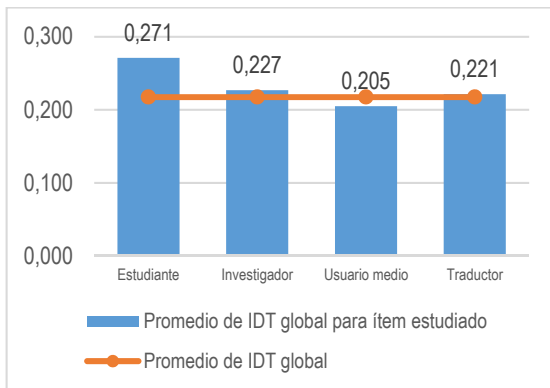


Gráfico 258: Promedio de IDT global para subdestinatario especialista no investigador según informantes y diferencia con promedio de IDT global

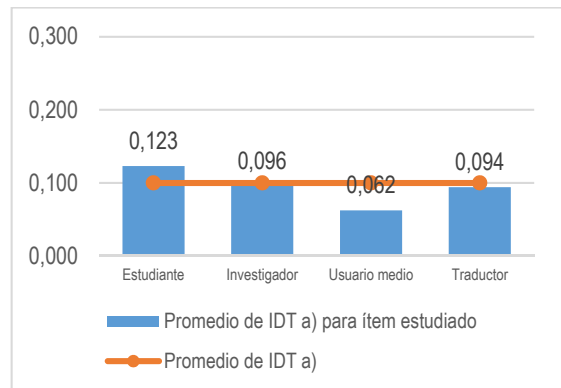


Gráfico 259: Promedio de IDT a) para subdestinatario especialista no investigador según informantes y diferencia con promedio de IDT a)

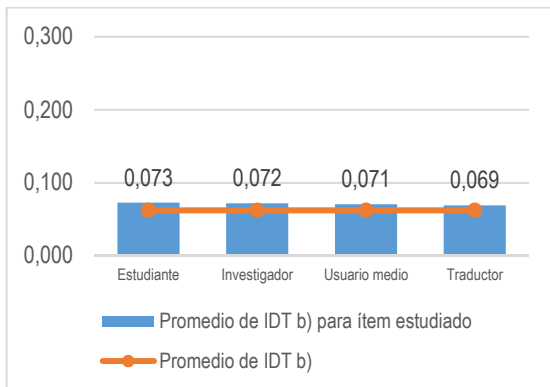


Gráfico 260: Promedio de IDT b) para subdestinatario especialista no investigador según informantes y diferencia con promedio de IDT b)

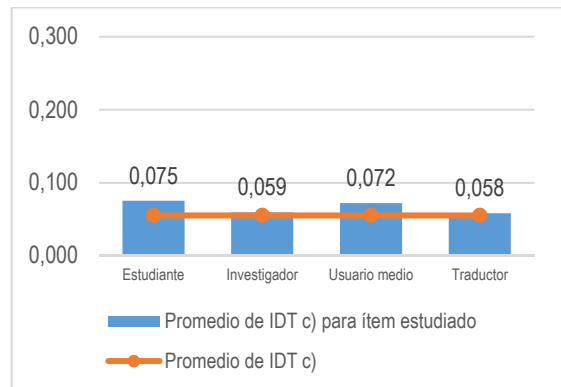


Gráfico 261: Promedio de IDT c) para subdestinatario especialista no investigador según informantes y diferencia con promedio de IDT c)

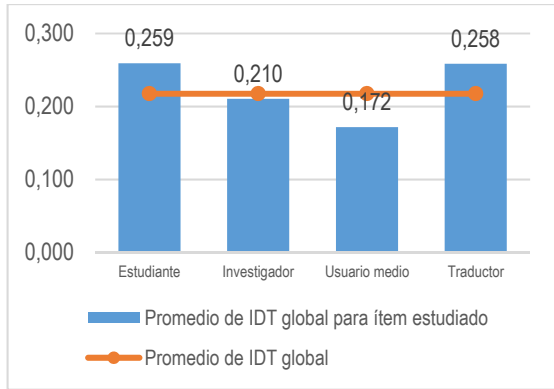


Gráfico 262: Promedio de IDT global para subdestinatario instruido semiespecialista según informantes y diferencia con promedio de IDT global

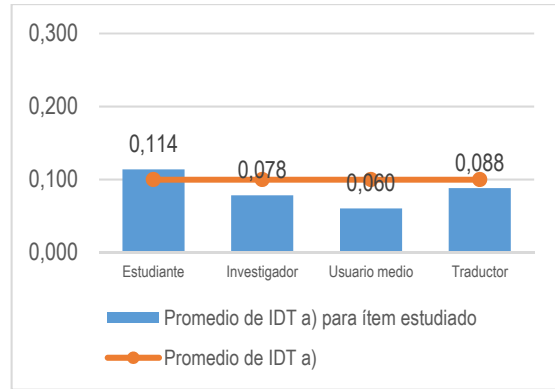


Gráfico 263: Promedio de IDT a) para subdestinatario instruido semiespecialista según informantes y diferencia con promedio de IDT a)

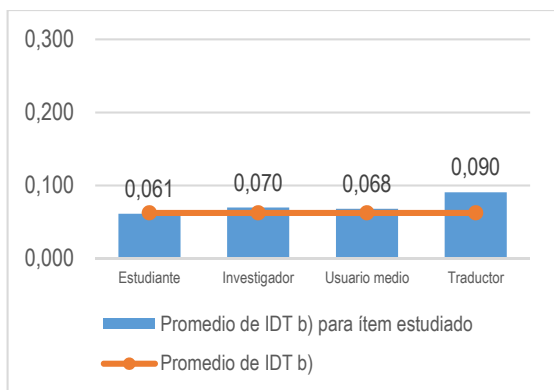


Gráfico 264: Promedio de IDT b) para subdestinatario instruido semiespecialista según informantes y diferencia con promedio de IDT b)

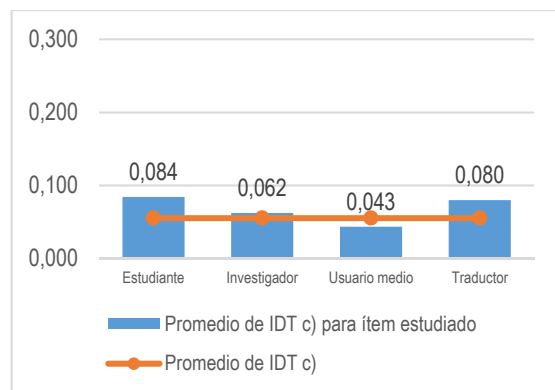


Gráfico 265: Promedio de IDT c) para subdestinatario instruido semiespecialista según informantes y diferencia con promedio de IDT c)

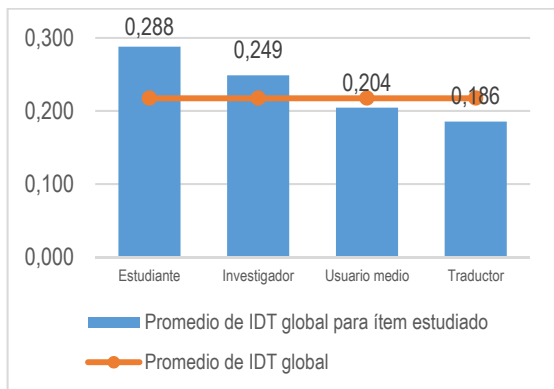


Gráfico 266: Promedio de IDT global para subdestinatario instruido aprendiz de especialista según informantes y diferencia con promedio de IDT global

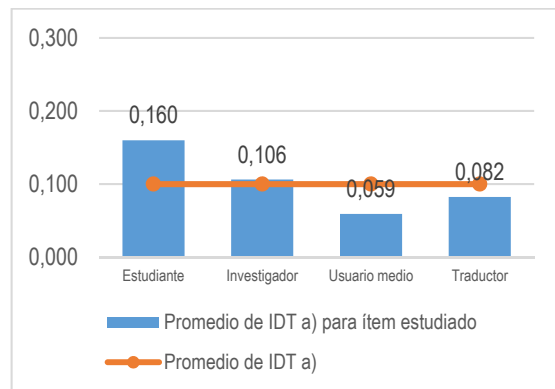


Gráfico 267: Promedio de IDT a) para subdestinatario instruido aprendiz de especialista según informantes y diferencia con promedio de IDT a)

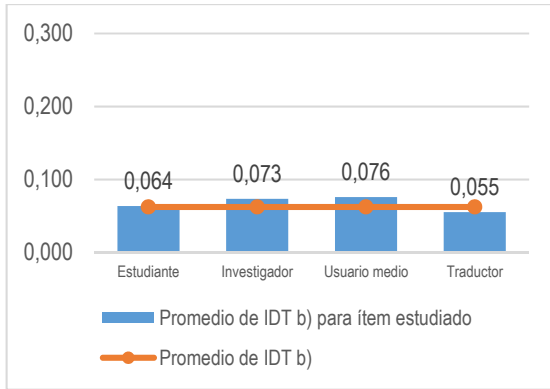


Gráfico 268: Promedio de IDT b) para subdestinatario instruido aprendiz de especialista según informantes y diferencia con promedio de IDT b)

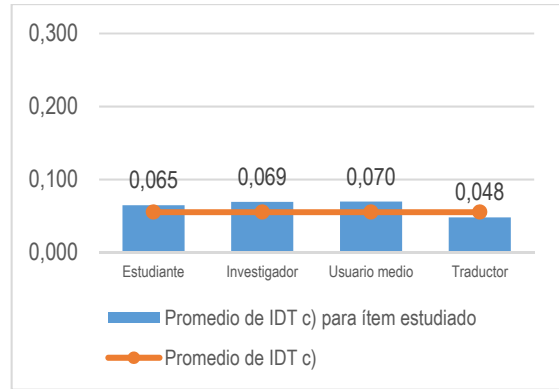


Gráfico 269: Promedio de IDT c) para subdestinatario instruido aprendiz de especialista según informantes y diferencia con promedio de IDT c)

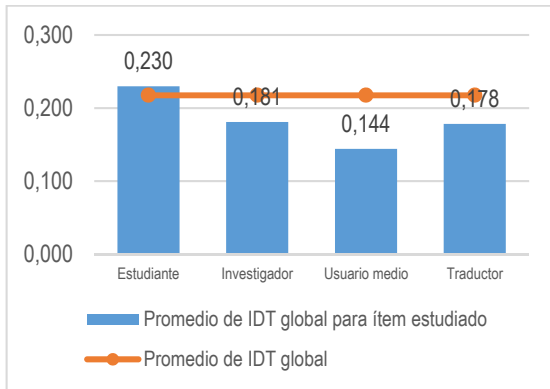


Gráfico 270: Promedio de IDT global para subdestinatario instruido lego formado según informantes y diferencia con promedio de IDT global

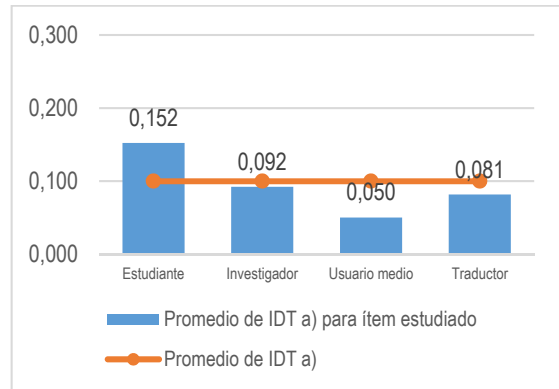


Gráfico 271: Promedio de IDT a) para subdestinatario instruido lego formado según informantes y diferencia con promedio de IDT a)

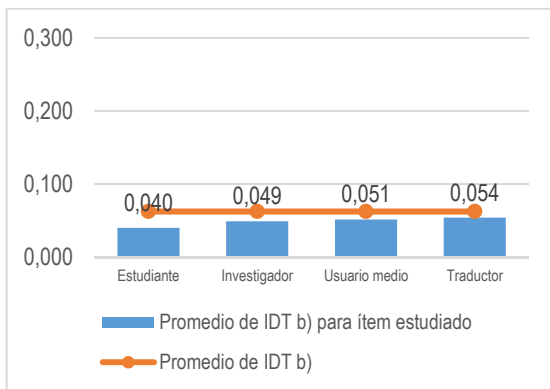


Gráfico 272: Promedio de IDT b) para subdestinatario instruido lego formado según informantes y diferencia con promedio de IDT b)

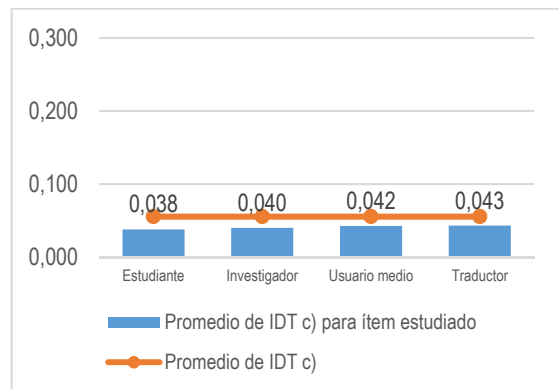


Gráfico 273: Promedio de IDT c) para subdestinatario instruido lego formado según informantes y diferencia con promedio de IDT c)

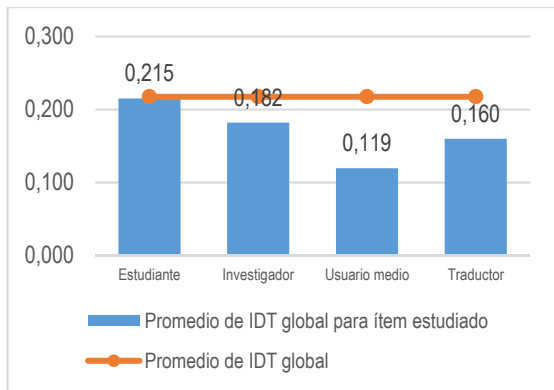


Gráfico 274: Promedio de IDT global para subdestinatario lego no formado según informantes y diferencia con promedio de IDT global

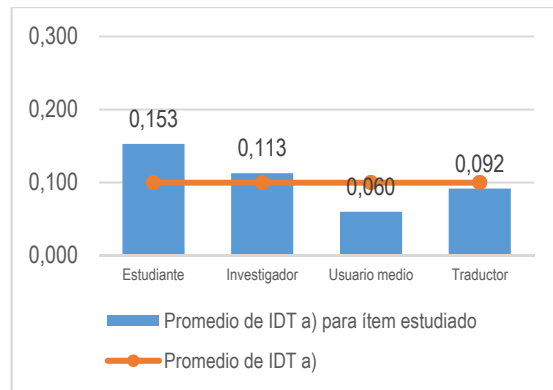


Gráfico 275: Promedio de IDT a) para subdestinatario lego no formado según informantes y diferencia con promedio de IDT a)

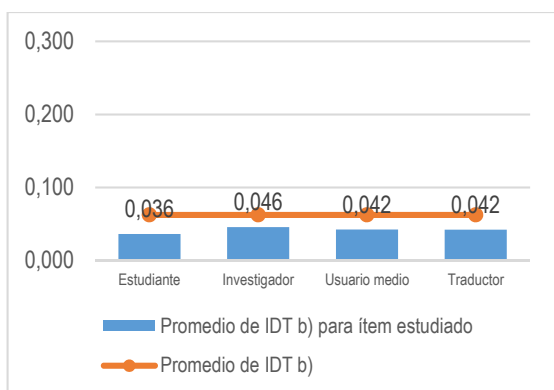


Gráfico 276: Promedio de IDT b) para subdestinatario lego no formado según informantes y diferencia con promedio de IDT b)

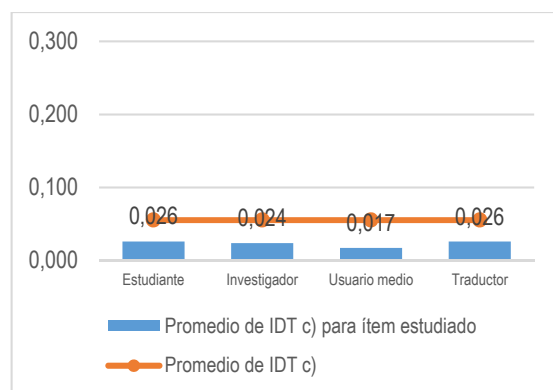


Gráfico 277: Promedio de IDT c) para subdestinatario lego no formado según informantes y diferencia con promedio de IDT c)

2.6. Conclusiones parciales sobre los INFORMANTES

A continuación, se relaciona una recapitulación de las interpretaciones más relevantes sobre los datos relacionados con la percepción de los informantes sobre las distintas variables de análisis de los textos del corpus.

- a) Existen más similitudes que divergencias en el criterio de asignación de valores de los informantes para cada atributo. En otras palabras, el percepto (el producto de la percepción, resultado de la fijación del fenómeno lingüístico que los hablantes captan) de los informantes comparte más características comunes que diferentes, por lo que se advierte cierto grado de convención o aceptación colectiva al analizar el grado de especialización de un texto.
- b) El grupo de estudiantes es el que presenta divergencias más notables, a veces junto con el usuario medio. Este par contrasta con las similitudes que suelen presentar siempre el investigador con el grupo de traductores.

- c) En relación con la clase, los datos determinan que las diferencias de los informantes para clasificar los textos según su grado de especialización no dependen estrictamente del IDT global o el IDT de términos crípticos, como podría suponerse, sino que atienden a la percepción sintética de cada informante sobre el conjunto de características discursivas del texto. Es decir, los informantes distinguen la clasificación de los fenómenos lingüísticos (las clases) atendiendo a la globalidad del objeto complejo (el texto). Si se vincula con el tipo textual, se advierte que el grupo de estudiantes es el que muestra más diferencias con los otros tres grupos: relaciona más estrechamente la aparición de terminología con la clase de texto especializado y determina que el subdestinatario instruido prima en su relación con el texto semiespecializado.
- d) En cuanto al emisor y al subemisor, los informantes muestran similitudes, pero también ciertas tendencias que permiten diferenciarlos. Por ejemplo, el grupo de estudiantes tiende ligeramente hacia el subemisor especialista investigador y el resto de informantes opta por el instruido semiespecialista en lugar de por el especialista no investigador. Las tendencias que muestran los informantes a clasificar los subemisores en torno a tres categorías puede estar relacionada con la facilidad de vincular textos en torno a tres categorías prototípicas que podrían corresponder con los tres grados de especialización (al igual que la supercategoría de emisor). En cuanto a los tipos textuales, existe una tendencia convergente generalizada, tan solo excepcional en los tipos textuales de blog, entrevista o revista o web de divulgación especializada.
- e) A propósito del destinatario y subdestinatario, se conservan, en gran medida, los comentarios reservados para el emisor. Las similitudes son más claras en el caso del instruido lego formado, el más frecuente según todos los subdestinatarios. En el caso de los tipos textuales, la homogeneidad presente en el caso del subemisor desaparece en el subdestinatario: suelen coincidir en la distribución de los subdestinatarios, pero no en los valores numéricos que se les asigna a los subdestinatarios menos frecuentes.
- f) Las formas léxicas especializadas muestran diferencias principalmente entre estudiante y usuario medio. Son, respectivamente, el informante que más y que menos formas léxicas especializadas délficas identifica; y el que menos y más formas semicrípticas identifica. Así, puede determinarse que el factor que toma cada grupo de informantes para regular su percepción puede ser diferente al

llevar a cabo la observación analítica para las formas léxicas, si bien no necesariamente tiene que serlo al determinar la clase mediante el análisis sintético.

- g) Por último, en relación con la densidad terminológica, las diferencias que existen están relacionadas con las mencionadas para las formas léxicas especializadas.

2.7. Resumen de análisis correlacional descriptivo según los informantes

Como conclusión de esta parte, se han diseñado cuatro tablas (*vid. infra* Tabla 142 a Tabla 145) que sintetizan los valores más representativos para cada una de las variables de análisis según los cuatro grupos de informantes. En primer lugar, se relacionan las tablas y en segundo lugar la explicación de estas cuatro tablas. En estas tablas los valores numéricos muestran algunos colores que permiten diferenciar los valores máximos y mínimos (teniendo en cuenta algunas centésimas por encima o por debajo):

- a) Los colores para la fuente del texto comparan los valores para una misma variable entre los tres temas dentro de los datos proporcionados por un mismo informante. El color rojo muestra el valor mínimo observable para una variable entre los tres temas en un informante. Por el contrario, el color verde muestra el valor máximo.
- b) Los colores para la celda comparan los valores para una misma variable entre los cuatro grupos de informantes. El color rojo muestra el valor mínimo y el verde el valor máximo al comparar los datos facilitados por los cuatro grupos de informantes.
- c) Los valores para IR muestran sus valores escalados.

Tema y tipo texto	TTR	IDT				IR				Emisor	Subemisor	Destinatario	Subdestinatario	Clase
		Global	a	b	c	Global	a	b	c					
Derecho														
ART	63,493	0,266	0,162	0,075	0,029	0,074	0,093	0,015	0,110	Esp	Esp I	Esp / Ins	Esp I / Ins LF	TE
BLOG	63,935	0,329	0,207	0,082	0,040	0,018	0,023	0,003	0,027	Esp	Esp NI	Ins	Ins LF / Leg NF	TSE
ENT	67,467	0,274	0,164	0,064	0,045	0,019	0,020	0,005	0,040	Esp / Ins	Ins SE	Ins / Leg	Leg NF	TNE
NOT	67,471	0,213	0,151	0,050	0,012	0,035	0,049	0,006	0,044	Ins	Ins SE	Ins	Ins LF	TE / TNE
FOR	63,268	0,214	0,158	0,046	0,011	0,008	0,015	0,001	0,004	Leg	Leg NF	Esp / Ins	Esp NI / Ins LF / Leg NF	TNE
DIV	68,982	0,290	0,167	0,077	0,046	0,007	0,006	0,001	0,021	Ins	Esp NI / Ins SE	Ins	Ins LF	TSE
TES	63,908	0,294	0,191	0,061	0,042	0,079	0,107	0,009	0,126	Esp	Esp I	Esp	Esp I	TE
Informática														
ART	60,358	0,259	0,098	0,060	0,101	0,164	0,186	0,048	0,243	Esp	Esp I	Esp / Ins	Esp I	TE
BLOG	65,187	0,283	0,137	0,043	0,104	0,019	0,021	0,006	0,023	Esp / Ins	Ins SE / Esp NI	Ins	Ins LF	TSE / TE
ENT	67,468	0,282	0,138	0,060	0,084	0,036	0,055	0,005	0,031	Esp / Ins	Esp I / Ins SE	Ins	Ins LF	TSE
NOT	68,623	0,173	0,103	0,035	0,035	0,075	0,053	0,052	0,060	Ins	Ins SE	Ins	Ins LF	TNE
FOR	64,095	0,199	0,133	0,026	0,040	0,012	0,013	0,004	0,016	Ins	Ins LF	Ins	Ins LF	TSE / TNE
DIV	71,452	0,178	0,121	0,029	0,028	0,025	0,028	0,004	0,056	Ins	Esp NI / Ins SE	Ins	Ins LF	TSE / TNE
TES	60,069	0,258	0,097	0,090	0,071	0,104	0,126	0,021	0,181	Esp	Esp I	Esp / Ins	Esp I	TE
Medicina														
ART	54,421	0,284	0,145	0,064	0,075	0,077	0,092	0,019	0,118	Esp	Esp I	Esp	Esp I	TE
BLOG	71,555	0,257	0,118	0,053	0,087	0,032	0,046	0,007	0,031	Esp	Esp I	Leg / Ins	Leg NF	TNE / TSE
ENT	65,665	0,287	0,140	0,072	0,074	0,031	0,036	0,008	0,048	Esp / Ins	Ins SE / Esp I	Ins	Ins LF	TSE
NOT	66,432	0,201	0,151	0,034	0,017	0,048	0,057	0,012	0,077	Ins	Ins SE	Ins	Ins LF / Leg NF	TSE / TNE
FOR	72,094	0,188	0,115	0,038	0,036	0,046	0,073	0,006	0,039	Leg	Leg NF	Leg	Leg NF	TNE
DIV	74,037	0,247	0,160	0,056	0,031	0,024	0,027	0,007	0,038	Esp / Ins	Ins SE / Esp I / Esp NI	Ins	Esp NI / Ins LF / Ins SE	TSE
TES	56,649	0,387	0,142	0,088	0,157	0,085	0,092	0,011	0,218	Esp	Esp I	Esp	Esp I	TE

Tabla 142: Resumen global de valores más representativos para cada variable según tema y tipo de texto según grupo de estudiantes

En el caso del grupo de estudiantes contrasta la percepción de la densidad terminológica entre el derecho, por un lado, y la informática y la medicina, por otro. La densidad de formas délficas y semicrípticas es mayor para el derecho, mientras que la informática cuenta con valores que sobresalen en la informática. En el caso de la reformulación, la relación es inversa, puesto que destacan la medicina y la informática sobre el derecho prácticamente en todos los casos. En efecto, los valores de densidad terminológica de formas délficas y la reformulación en el derecho son, respectivamente, los valores máximos y mínimos de todo el corpus y tienen relación, a nuestro juicio con la vinculación entre formas denominativas comprensibles por los usuarios y la falta de necesidad de explicitar la información mediante reformulación. La informática o la medicina, por el contrario, con valores de IDT c) más elevados, cuentan con un índice de reformulación mayor para cada tipo de reformulación.

Tema y tipo texto	TTR	IDT				IR				Emisor	Subemisor	Destinatario	Subdestinatario	Clase
		Global	a	b	c	Global	a	b	c					
Derecho														
ART	63,493	0,183	0,073	0,081	0,029	0,066	0,083	0,013	0,101	Esp	Esp I	Esp	Esp I	TE
BLOG	63,935	0,225	0,079	0,116	0,031	0,017	0,022	0,003	0,027	Esp	Esp NI	Ins	Ins AE / Ins LF / Leg NF	TSE
ENT	67,467	0,147	0,049	0,060	0,037	0,019	0,020	0,005	0,042	Ins / Esp	Ins SE	Ins	Leg NF / Ins SE	TNE
NOT	67,471	0,140	0,067	0,060	0,014	0,039	0,055	0,006	0,044	Ins	Ins SE	Leg	Leg NF	TNE
FOR	63,268	0,110	0,054	0,049	0,007	0,009	0,016	0,001	0,006	Leg / Ins	Leg NF	Ins / Esp	Esp NI / Ins LF	TNE
DIV	68,982	0,189	0,068	0,082	0,039	0,007	0,006	0,001	0,020	Ins	Ins SE	Ins / Leg	Leg NF / Ins LF	TSE
TES	63,908	0,206	0,091	0,078	0,037	0,073	0,100	0,009	0,115	Esp	Esp I	Esp	Esp I	TE
Informática														
ART	60,358	0,227	0,050	0,074	0,102	0,144	0,163	0,042	0,213	Esp	Esp I	Esp	Esp I	TE
BLOG	65,187	0,213	0,040	0,063	0,109	0,023	0,026	0,008	0,026	Ins	Ins SE	Ins	Ins AE / Ins LF	TSE
ENT	67,468	0,212	0,064	0,065	0,083	0,037	0,057	0,005	0,029	Ins	Ins SE	Ins	Ins LF	TSE
NOT	68,623	0,105	0,026	0,040	0,039	0,108	0,067	0,082	0,088	Ins	Ins SE	Leg / Ins	Leg NF	TNE
FOR	64,095	0,129	0,053	0,034	0,043	0,015	0,015	0,006	0,019	Ins	Ins LF	Ins	Ins LF	TNE / TSE
DIV	71,452	0,111	0,046	0,036	0,028	0,029	0,034	0,004	0,060	Ins / Esp	Ins SE / Esp NI	Ins	Ins AE / Ins LF	TSE
TES	60,069	0,225	0,052	0,092	0,081	0,095	0,114	0,020	0,165	Esp	Esp I	Esp / Ins	Esp I	TE
Medicina														
ART	54,421	0,241	0,089	0,070	0,082	0,067	0,080	0,017	0,101	Esp	Esp I	Esp	Esp I	TE
BLOG	71,555	0,209	0,056	0,067	0,086	0,029	0,041	0,006	0,026	Esp	Esp NI / Esp I	Ins	Ins LF / Leg NF	TNE / TSE
ENT	65,665	0,214	0,062	0,078	0,074	0,030	0,034	0,008	0,049	Esp	Ins SE / Esp NI	Ins	Ins LF / Leg NF	TSE
NOT	66,432	0,119	0,062	0,039	0,018	0,052	0,062	0,012	0,083	Ins	Ins SE	Ins / Leg	Leg NF	TSE / TNE
FOR	72,094	0,112	0,036	0,032	0,045	0,050	0,076	0,006	0,051	Ins	Ins LF	Ins / Leg	Leg NF / Ins LF	TNE
DIV	74,037	0,150	0,053	0,064	0,034	0,029	0,030	0,011	0,040	Esp	Ins SE / Esp I	Ins	Ins AE / Ins LF	TSE
TES	56,649	0,353	0,085	0,097	0,170	0,073	0,079	0,009	0,188	Esp	Esp I	Esp	Esp I	TE

Tabla 143: Resumen global de valores más representativos para cada variable según tema y tipo de texto según grupo de usuarios medios

El esquema observado en el grupo de estudiantes se repite para el grupo de usuarios medios, a excepción de ciertos casos. La densidad de términos délficos es mayor en el derecho, en contraste con la de términos crípticos, más destacada en la informática. Esto tiene correlación con la reformulación poco empleada en el derecho y muy notable en la informática o la medicina. A pesar de las diferencias en la densidad terminológica de los diferentes tipos de IDT entre el grupo de estudiantes y el grupo de usuarios medios, cabe destacar el hecho de que la relación entre densidad terminológica y reformulación sea muy similar entre temas en los dos grupos de informantes, lo que permite defender la relación entre la opacidad semántica o la falta de frecuencia de las denominaciones como un criterio para el empleo de reformulación en los textos.

Tema y tipo texto	TTR	IDT				IR				Emisor	Subemisor	Destinatario	Subdestinatario	Clase
		Global	a	b	c	Global	a	b	c					
Derecho														
ART	63,493	0,255	0,153	0,074	0,028	0,125	0,139	0,034	0,212	Esp	Esp I	Esp	Esp I	TE
BLOG	63,935	0,290	0,171	0,095	0,024	0,028	0,033	0,007	0,041	Ins / Esp	Ins SE / Esp NI	Ins / Leg	Leg NF	TSE / TNE
ENT	67,467	0,232	0,129	0,065	0,038	0,038	0,040	0,011	0,071	Ins	Ins SE	Ins	Leg NF	TNE
NOT	67,471	0,189	0,123	0,054	0,012	0,049	0,070	0,008	0,057	Ins	Ins SE	Leg	Leg NF	TNE
FOR	63,268	0,156	0,102	0,047	0,007	0,015	0,025	0,001	0,009	Ins / Leg	Leg NF	Ins	Ins LF	TNE
DIV	68,982	0,233	0,110	0,084	0,039	0,011	0,009	0,002	0,037	Ins	Ins SE	Leg	Leg NF	TNE / TSE
TES	63,908	0,280	0,161	0,080	0,039	0,114	0,151	0,014	0,193	Esp	Esp I	Esp	Esp I	TE
Informática														
ART	60,358	0,255	0,092	0,072	0,091	0,170	0,189	0,052	0,250	Esp	Esp I	Esp	Esp I	TE
BLOG	65,187	0,238	0,080	0,056	0,102	0,038	0,044	0,014	0,041	Ins	Ins SE	Ins	Ins LF	TSE
ENT	67,468	0,245	0,101	0,063	0,082	0,046	0,071	0,007	0,036	Ins	Ins SE	Ins	Ins LF	TSE
NOT	68,623	0,131	0,057	0,036	0,038	0,208	0,139	0,152	0,166	Ins	Ins SE	Leg	Leg NF	TNE
FOR	64,095	0,166	0,095	0,033	0,039	0,023	0,022	0,011	0,025	Ins	Ins LF	Ins	Ins LF	TNE / TSE
DIV	71,452	0,140	0,080	0,035	0,025	0,038	0,046	0,006	0,074	Ins	Ins SE	Ins	Ins LF	TSE
TES	60,069	0,247	0,077	0,093	0,077	0,110	0,134	0,022	0,190	Esp	Esp I	Esp / Ins	Esp I	TE
Medicina														
ART	54,421	0,272	0,127	0,066	0,078	0,088	0,112	0,020	0,115	Esp	Esp I	Esp	Esp I	TE
BLOG	71,555	0,241	0,092	0,065	0,083	0,036	0,052	0,008	0,031	Ins	Ins SE	Leg	Leg NF	TNE / TSE
ENT	65,665	0,249	0,103	0,076	0,071	0,036	0,042	0,010	0,054	Ins	Ins SE	Ins / Leg	Leg NF	TSE
NOT	66,432	0,167	0,109	0,042	0,017	0,087	0,107	0,018	0,137	Ins	Ins LF	Leg	Leg NF	TNE
FOR	72,094	0,134	0,060	0,032	0,042	0,061	0,093	0,007	0,065	Ins	Ins LF / Leg NF	Leg / Ins	Leg NF	TNE
DIV	74,037	0,188	0,094	0,061	0,033	0,042	0,043	0,016	0,064	Ins	Ins SE	Ins	Ins LF / Leg NF	TSE
TES	56,649	0,381	0,121	0,094	0,165	0,081	0,087	0,010	0,211	Esp	Esp I	Esp	Esp I	TE

Tabla 144: Resumen global de valores más representativos para cada variable según tema y tipo de texto según grupo de investigador

Los datos aportados por el investigador continúan con la misma distribución. En este caso, destaca por ser el informante que cuenta con los valores más elevados de reformulación para la informática, particularmente, pero también para la medicina. Este dato está relacionado con el tipo de unidad léxica seleccionada como terminológica por el investigador. En este caso, la unidad léxica especializada estar acompañada de procedimientos de reformulación.

Tema y tipo texto	TTR	IDT				IR				Emisor	Subemisor	Destinatario	Subdestinatario	Clase
		Global	a	b	c	Global	a	b	c					
Derecho														
ART	63,493	0,219	0,115	0,075	0,029	0,097	0,106	0,029	0,154	Esp	Esp I	Esp	Esp I	TE
BLOG	63,935	0,267	0,131	0,105	0,031	0,021	0,026	0,004	0,032	Esp	Esp NI	Ins	Ins LF / Ins SE	TSE
ENT	67,467	0,201	0,099	0,062	0,040	0,024	0,026	0,005	0,048	Esp / Ins	Ins SE	Ins	Ins LF	TNE
NOT	67,471	0,164	0,098	0,054	0,012	0,044	0,062	0,007	0,052	Ins	Ins SE	Leg	Leg NF	TNE
FOR	63,268	0,138	0,085	0,046	0,007	0,011	0,019	0,001	0,007	Leg	Leg NF	Ins	Esp NI / Ins LF / Leg NF	TNE
DIV	68,982	0,210	0,087	0,084	0,038	0,009	0,007	0,002	0,027	Ins	Esp NI / Ins SE	Leg / Ins	Leg NF	TNE / TSE
TES	63,908	0,260	0,144	0,077	0,040	0,085	0,118	0,010	0,135	Esp	Esp I	Esp	Esp I	TE
Informática														
ART	60,358	0,247	0,079	0,072	0,097	0,155	0,174	0,046	0,227	Esp	Esp I	Esp	Esp I	TE
BLOG	65,187	0,225	0,067	0,062	0,096	0,026	0,030	0,010	0,029	Ins / Esp	Ins SE / Esp NI	Ins / Esp	Ins LF	TSE
ENT	67,468	0,231	0,084	0,064	0,082	0,042	0,065	0,006	0,032	Esp / Ins	Ins SE	Ins	Ins LF	TSE
NOT	68,623	0,134	0,043	0,040	0,050	0,152	0,085	0,124	0,113	Ins	Ins SE	Leg	Leg NF	TNE
FOR	64,095	0,150	0,082	0,031	0,037	0,018	0,020	0,007	0,022	Ins	Ins LF	Ins	Ins LF	TNE / TSE
DIV	71,452	0,133	0,069	0,036	0,028	0,032	0,038	0,005	0,064	Ins	Ins SE	Ins	Ins LF	TSE
TES	60,069	0,243	0,074	0,093	0,076	0,102	0,123	0,020	0,175	Esp	Esp I	Esp	Esp I	TE
Medicina														
ART	54,421	0,268	0,121	0,068	0,079	0,075	0,092	0,018	0,105	Esp	Esp I	Esp	Esp I	TE
BLOG	71,555	0,236	0,084	0,068	0,085	0,032	0,046	0,006	0,028	Esp	Esp NI / Ins SE	Leg	Leg NF	TNE / TSE
ENT	65,665	0,245	0,094	0,077	0,074	0,033	0,038	0,009	0,052	Esp / Ins	Ins SE	Leg / Ins	Leg NF	TSE
NOT	66,432	0,151	0,097	0,038	0,016	0,062	0,074	0,014	0,099	Ins	Ins SE	Ins / Leg	Leg NF	TNE
FOR	72,094	0,124	0,052	0,030	0,043	0,056	0,084	0,007	0,059	Leg	Ins LF / Leg NF	Leg	Leg NF	TNE
DIV	74,037	0,176	0,079	0,064	0,033	0,034	0,034	0,014	0,045	Esp / Ins	Esp NI	Ins	Ins LF / Ins SE / Leg NF	TSE
TES	56,649	0,370	0,113	0,093	0,164	0,078	0,084	0,010	0,204	Esp	Esp I	Esp	Esp I	TE

Tabla 145: Resumen global de valores más representativos para cada variable según tema y tipo de texto según grupo de traductores

Como puede comprobarse, las tablas para cada informante suelen mostrar algunos bloques que aglutinan un color u otro, lo que permite identificar los valores máximos o mínimos en un informante. Por ejemplo, el estudiante (Tabla 142) destaca en las formas délficas, especialmente en derecho (es coherente con dos hechos: con que el derecho es el tema que más términos délficos posee, como se observa en la Tabla 125; p. 357; y con que el estudiante sea el grupo de informantes que más términos délficos identifica); pero es el que menos reformulación presenta. En el caso del usuario medio (Tabla 143), comparte con el estudiante los valores mínimos para el derecho. El investigador (Tabla 145), por otra parte, es el que presenta los valores más elevados en reformulación para informática y para medicina.

En cuanto a los valores cualitativos, se observa una tendencia generalizada en la percepción sintética de cada grupo de informantes, puesto que las coincidencias en las clases para cada tipo textual para cada tema suelen coincidir (lo que justifica que defendamos dicha convención colectiva). Los resultados para la percepción analítica para los interlocutores divergen ligeramente entre informantes, pero también queda justificada

por las características que cada uno encuentra en los textos. En definitiva, la caracterización de la reformulación y la densidad terminológica es fruto de la prominencia de rasgos que cada grupo de informantes percibe en cada tema.

3. ANÁLISIS CORRELACIONAL A PARTIR DE APRENDIZAJE AUTOMÁTICO

En esta sección se presentan las diferentes tablas y gráficos que incluyen los resultados al aplicar cuatro algoritmos a la base de datos con los análisis de los informantes.

En resumen, esta sección incluye:

- 1) Las probabilidades de clasificación en las tres clases según las diferentes variables tras aplicar el algoritmo de red bayesiana y realizar los cálculos con el teorema de Bayes.
- 2) Los valores de las variables que son más relevantes en la clasificación de los textos en texto especializado o texto no especializado tras aplicar el algoritmo de regresión logística y el correspondiente regularizador.
- 3) Los árboles de decisión que determinan las diferentes clases de acuerdo con un conjunto de valores condicionantes tras aplicar (y aplicar una poda a) el algoritmo J48.
- 4) Los centroides para las tres clases según el grado de especialización, así como la distribución de patrones en las agrupaciones que realiza el algoritmo no supervisado *simple k-means*.

3.1. Probabilidades de clasificación en clases según variables

El algoritmo de red bayesiana¹⁶⁷ y el teorema de Bayes (*vid. infra* Tabla 62) permiten conocer dos datos:

- a) La probabilidad del valor del atributo conociendo el valor de la clase, es decir, por ejemplo, la probabilidad de que un texto tenga como subemisor a un usuario especialista investigador cuando el texto es especializado;
- b) La probabilidad del valor de la clase conociendo el valor del atributo, es decir, por ejemplo, la probabilidad de que un texto sea clasificado como texto especializado cuando el subemisor es especialista investigador.

$$P(C|At) = \frac{\#At \cap \#C}{\#At}$$

C: valor de la clase

At: valor del atributo

#: número de patrones que cumplen con el valor

Tabla 62: Fórmula del teorema de Bayes aplicado al corpus

¹⁶⁷ *Vid. supra* apartado III. 7.2. (p. 231) para leer la información sobre el método y funcionamiento del teorema de Bayes y su aplicación en esta tesis doctoral.

Dicho algoritmo y teorema ha sido aplicado a la base de datos global con los datos de los cuatro grupos de informantes e individualmente a cada base de datos con los datos de exclusivamente cada uno de los grupos de informantes, por lo que contamos con cinco tablas de probabilidades (Tabla 146 a Tabla 151).

Los resultados muestran cómo existen dos atributos que no contribuyen a la determinación de probabilidades para la clase según el grado de especialización textual, por lo que los resultados de WEKA excluyen dichos atributos en el análisis global. Estos dos atributos no importantes en la probabilidad son la TTR estandarizada, es decir, la variedad léxica, y del IDT a), es decir, de la densidad terminológica de formas délficas. Respectivamente, en los apartados IV. 1.6.7. y IV. 1.4.9. (p. 356 y 323) se determinó que el IDT a) no discrimina las clases según el grado de especialización y que la TTR estandarizada establece una relación inversa con la clase. No obstante, dado que la distribución de número de textos para cada intervalo de IDT a) y de TTR estandarizada es estrechamente similar en cada clase, no resulta pertinente a la hora de calcular la probabilidad de los valores de atributo o valores de clase, como se muestra gráficamente en la Ilustración 25 y la Ilustración 26 (en dichos gráficos el azul oscuro representa el texto especializado, el rojo el texto no especializado y el azul claro el texto semiespecializado, de acuerdo con los colores asignados por WEKA).

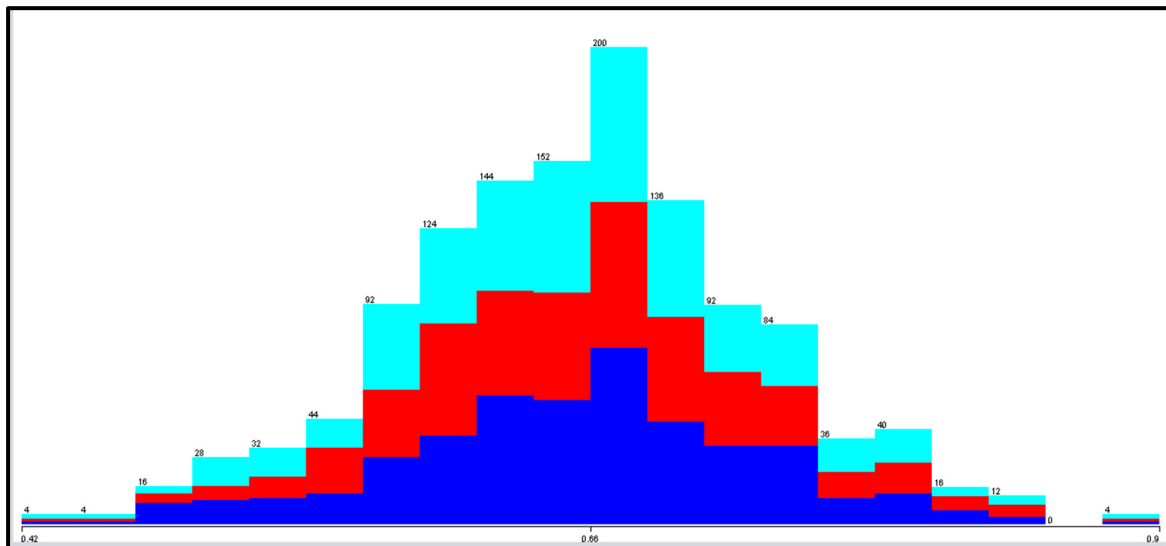


Ilustración 25: Distribución de patrones por TTR según algoritmo de red bayesiana [elaboración propia]

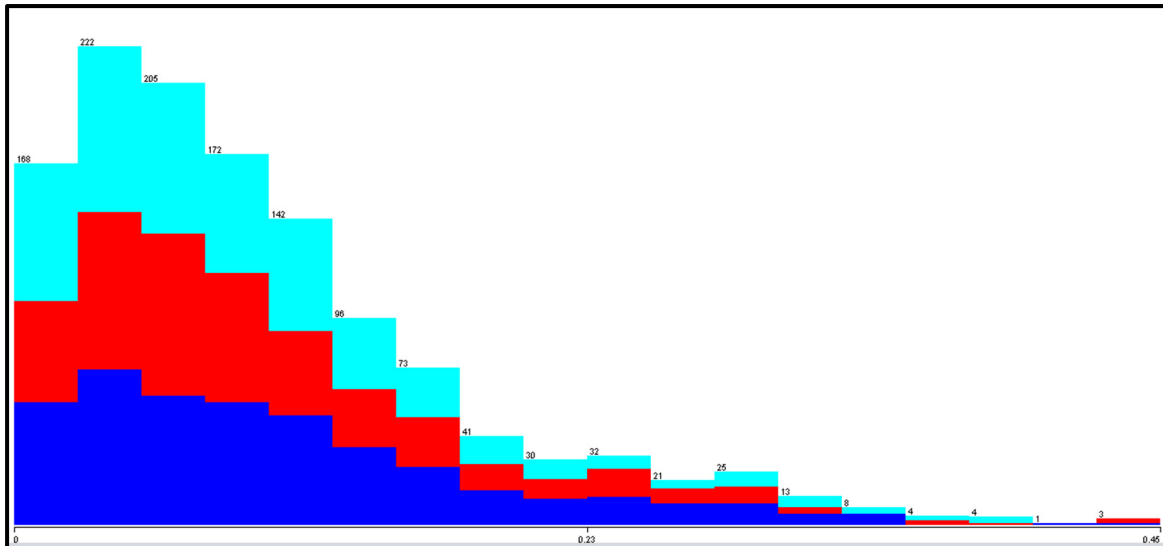


Ilustración 26: Distribución de patrones por IDT a) según algoritmo de red bayesiana [elaboración propia]

Al analizar las cuatro bases de datos individuales por cada informante, WEKA también determina que el IR b) no resulta relevante para determinar la probabilidad de los valores de atributo conociendo los valores de clase, por lo que también quedan excluidos de las tablas de informantes (Tabla 148 a Tabla 151).

La siguiente tabla presenta los datos divididos en dos grandes bloques, que determinan la probabilidad para el atributo o para la clase conociendo el valor del inverso.

En el bloque 1 cada atributo suma 100 % si se suman los valores de las columnas para cada clase (es decir, la suma de los valores que hay en TE para el tema en el bloque 1 suman 100 %), mientras que en el bloque 2 el valor para cada atributo suma 100 % si se suman los valores de las filas para cada clase (es decir, la suma de los valores que hay en TE, TSE y TNE para DER en el bloque 2 suman 100 %). Los datos casan a la perfección con el análisis para cada variable realizado tanto en la sección IV. 1. como en la sección IV. 2. (pp. 244 y 358):

		1) Probabilidad del valor del atributo conociendo el valor de la clase			2) Probabilidad del valor de la clase conociendo el valor del atributo		
Valor del atributo		TE	TSE	TNE	TE	TSE	TNE
Tema	DER	36,26 %	28,17 %	35,66 %	37,38 %	28,57 %	34,05 %
	INF	30,48 %	38,50 %	30,92 %	31,43 %	39,05 %	29,52 %
	MED	33,26 %	33,33 %	33,42 %	34,29 %	33,81 %	31,90 %
Tipo texto	ART	38,11 %	3,29 %	0,25 %	91,67 %	7,78 %	0,56 %
	BLOG	7,62 %	20,89 %	14,46 %	18,33 %	49,44 %	32,22 %
	DIV	7,62 %	24,88 %	10,22 %	18,33 %	58,89 %	22,78 %
	ENT	6,00 %	20,19 %	16,96 %	14,44 %	47,78 %	37,78 %
	FOR	0,69 %	11,27 %	32,17 %	1,67 %	26,67 %	71,67 %
	NOT	2,77 %	15,26 %	25,69 %	6,67 %	36,11 %	57,22 %
	TES	37,18 %	4,23 %	0,25 %	89,44 %	10,00 %	0,56 %
Emisor	Esp	95,15 %	36,62 %	13,72 %	66,13 %	25,04 %	8,83 %
	Ins	4,85 %	61,97 %	71,57 %	3,67 %	46,15 %	50,17 %

	Valor del atributo	1) Probabilidad del valor del atributo conociendo el valor de la clase			2) Probabilidad del valor de la clase conociendo el valor del atributo		
		TE	TSE	TNE	TE	TSE	TNE
		Leg	0,00 %	1,41 %	14,71 %	0,00 %	9,23 %
Subemisor	Esp I	79,21 %	11,03 %	3,49 %	84,90 %	11,63 %	3,47 %
	Esp NI	15,94 %	25,59 %	10,22 %	31,51 %	49,77 %	18,72 %
	Ins SE	4,39 %	46,48 %	40,40 %	5,01 %	52,24 %	42,74 %
	Ins AE	0,46 %	3,99 %	2,00 %	7,41 %	62,96 %	29,63 %
	Ins LF	0,00 %	11,50 %	29,18 %	0,00 %	29,52 %	70,48 %
	Leg NF	0,00 %	1,41 %	14,71 %	0,00 %	9,23 %	90,77 %
Destinatario	Esp	70,21 %	6,34 %	5,24 %	86,36 %	7,67 %	5,97 %
	Ins	29,79 %	80,99 %	37,41 %	20,67 %	55,29 %	24,04 %
	Leg	0,00 %	12,68 %	57,36 %	0,00 %	19,01 %	80,99 %
Subdestinatario	Esp I	54,04 %	1,41 %	1,00 %	95,90 %	2,46 %	1,64 %
	Esp NI	16,17 %	4,93 %	4,24 %	64,81 %	19,44 %	15,74 %
	Ins SE	13,39 %	10,56 %	2,00 %	52,25 %	40,54 %	7,21 %
	Ins AE	13,16 %	22,54 %	2,49 %	34,97 %	58,90 %	6,13 %
	Ins LF	3,23 %	47,89 %	32,92 %	4,00 %	58,29 %	37,71 %
	Leg NF	0,00 %	12,68 %	57,36 %	0,00 %	19,01 %	80,99 %
IDT global	<0,102	3,46 %	8,92 %	33,67 %	7,98 %	20,21 %	71,81 %
	0,1025-0,1945	18,01 %	43,43 %	38,40 %	18,71 %	44,36 %	36,93 %
	0,1945-0,3935	63,05 %	45,07 %	26,68 %	47,73 %	33,57 %	18,71 %
	>0,3935	15,47 %	2,58 %	1,25 %	80,72 %	13,25 %	6,02 %
IDT b)	<0,0075	0,23 %	7,98 %	18,20 %	0,93 %	31,48 %	67,59 %
	0,0075-0,0525	36,03 %	38,50 %	60,10 %	27,81 %	29,23 %	42,96 %
	0,0525-0,1405	41,34 %	50,23 %	18,70 %	38,25 %	45,73 %	16,03 %
	>0,1405	22,40 %	3,29 %	2,99 %	78,86 %	11,38 %	9,76 %
IDT c)	<0,0015	6,93 %	5,63 %	32,42 %	16,30 %	13,04 %	70,65 %
	0,0015-0,0585	36,95 %	60,09 %	59,35 %	24,46 %	39,14 %	36,39 %
	0,0585-0,151	34,41 %	28,17 %	7,98 %	49,50 %	39,87 %	10,63 %
IR global	>0,151	21,71 %	6,10 %	0,25 %	77,69 %	21,49 %	0,83 %
	<0,0608	46,19 %	79,34 %	86,53 %	22,60 %	38,19 %	39,21 %
IR a)	>0,0608	53,81 %	20,42 %	13,47 %	62,13 %	23,20 %	14,40 %
	<0,0753	54,04 %	81,46 %	87,28 %	25,13 %	37,27 %	37,59 %
IR b)	>0,0753	45,96 %	18,54 %	12,47 %	60,49 %	24,01 %	15,20 %
	<0,0016	45,27 %	57,75 %	70,32 %	27,07 %	33,98 %	38,95 %
IR c)	>0,0016	54,73 %	42,02 %	29,68 %	44,22 %	33,40 %	22,20 %
	<0,0091	25,17 %	50,70 %	63,59 %	18,79 %	37,24 %	43,97 %
	0,0091-0,0226	3,23 %	0,00 %	0,00 %	100,00 %	0,00 %	0,00 %
	0,0226-0,2201	48,27 %	45,31 %	33,92 %	38,85 %	35,87 %	25,28 %
	>0,2201	23,33 %	3,99 %	2,24 %	78,91 %	13,28 %	7,03 %

Tabla 146: Probabilidades de clasificación de clase según datos globales

Las probabilidades del tema no resultan del todo relevantes, puesto que la representación de los tres temas en las tres clases es muy equilibrada y, además, ya se defendió que el tema no condiciona la discriminación de clases.

La relación mencionada en el Gráfico 23 (*vid. supra* apartado IV. 1.2.1.) se mantiene en las probabilidades que se observan en el bloque 2): (i) destacan la elevada probabilidad de clasificación en texto especializado del artículo en revista científica y la tesis doctoral (son los tipos textuales que más alejados se encuentran de las otras clases); (ii) el blog, la revista o web de divulgación especializada y la entrevista rondan el 50 % de

probabilidad de ser clasificados como textos semiespecializados, seguido de los no especializados (en efecto, estos tipos textuales son los que se evalúan en la Tabla 87 y la Tabla 88); (iii) por último, la participación en foro y la noticia tienen mayor probabilidad de corresponder con texto no especializado, con un 71,67 % y un 57,22 % de probabilidades respectivamente. Estas probabilidades están muy relacionadas con el destinatario objetivo o prototípico de cada tipo textual, ya que la competencia cognitivo-comunicativa del destinatario parece vincularse con la clase según el grado de especialización (de hecho, como se mencionará en el apartado IV. 3.3. (p. 428), el punto de partida de los árboles de decisión es el destinatario).

Los atributos de emisor y destinatario también muestran diferencias significativas que hacen considerar la relevancia del destinatario como vehículo de la configuración discursiva. De hecho, el destinatario es que el que cuenta con probabilidades más polarizadas: muy elevada para TE para el especialista, media para TSE para el instruido y muy elevada para TNE para el lego. En el caso del emisor instruido, este se vincula más estrechamente con el TNE y el TSE, siendo mucho menor la probabilidad de relacionarse con el TE (*vid. infra* Tabla 149). El caso de subemisor y subdestinatario se ajusta a los datos de las categorías superiores de emisor y destinatario.

	TEXTO ESPECIALIZADO		TEXTO SEMIESPECIALIZADO		TEXTO NO ESPECIALIZADO	
	Emisor	Destinatario	Emisor	Destinatario	Emisor	Destinatario
Especialista	66,13 %	86,36 %	25,04 %	7,67 %	8,83 %	5,97 %
Instruido	3,67 %	20,67 %	46,15 %	55,29 %	50,17 %	24,04 %
Lego	0,00 %	0,00 %	9,23 %	19,01 %	90,77 %	80,99 %

Tabla 147: Probabilidades de clase según emisor y destinatario en el corpus

En relación con el IDT global y los tipos de IDT, así como con el IR global y los tipos de IR, se advierte una clara relación que provoca que aumente la probabilidad de clasificación en un grado de especialización mayor cuando aumentan los valores porcentuales para IDT e IR (como ya se manifestó en las relaciones de los apartados IV. 1.4.9. y IV. 1.5.10., pp. 323 y 349). Gráficamente, en el bloque 2 de la Tabla 146, en las correspondientes filas para tipos de IDT y para tipos de IR, se observa cómo el color verde descende desde la esquina superior derecha (menor valor del atributo y TNE) hacia la esquina inferior izquierda (mayor valor del atributo y TE).

A continuación, se contrastan las tablas para los cuatro grupos de informantes, de forma que puedan identificarse similitudes y diferencias en las probabilidades de clasificación según los informantes.

	Valor del atributo	1) Probabilidad del valor del atributo conociendo el valor de la clase			2) Probabilidad del valor de la clase conociendo el valor del atributo		
		TE	TSE	TNE	TE	TSE	TNE
Tema	DER	35,04 %	33,03 %	31,46 %	39,05 %	34,29 %	26,67 %
	INF	33,33 %	33,94 %	32,58 %	37,14 %	35,24 %	27,62 %
	MED	31,62 %	33,03 %	35,96 %	35,24 %	34,29 %	30,48 %
Tipo texto	ART	30,77 %	7,34 %	1,12 %	80,00 %	17,78 %	2,22 %
	BLOG	9,40 %	19,27 %	14,61 %	24,44 %	46,67 %	28,89 %
	DIV	11,11 %	23,85 %	6,74 %	28,89 %	57,78 %	13,33 %
	ENT	6,84 %	18,35 %	19,10 %	17,78 %	44,44 %	37,78 %
	FOR	2,56 %	11,01 %	33,71 %	6,67 %	26,67 %	66,67 %
	NOT	5,98 %	15,60 %	23,60 %	15,56 %	37,78 %	46,67 %
	TES	33,33 %	4,59 %	1,12 %	86,67 %	11,11 %	2,22 %
Emisor	Esp	94,87 %	41,28 %	21,35 %	63,43 %	25,71 %	10,86 %
	Ins	5,13 %	55,05 %	56,18 %	5,17 %	51,72 %	43,10 %
	Leg	0,00 %	3,67 %	22,47 %	0,00 %	16,67 %	83,33 %
Subemisor	Esp I	66,67 %	22,94 %	7,87 %	70,91 %	22,73 %	6,36 %
	Esp NI	28,21 %	18,35 %	13,48 %	50,77 %	30,77 %	18,46 %
	Ins SE	3,42 %	31,19 %	31,46 %	6,06 %	51,52 %	42,42 %
	Ins AE	1,71 %	9,17 %	3,37 %	13,33 %	66,67 %	20,00 %
	Ins LF	0,00 %	14,68 %	21,35 %	0,00 %	45,71 %	54,29 %
	Leg NF	0,00 %	3,67 %	22,47 %	0,00 %	16,67 %	83,33 %
Destinatario	Esp	58,97 %	14,68 %	6,74 %	75,82 %	17,58 %	6,59 %
	Ins	41,03 %	75,23 %	40,45 %	28,92 %	49,40 %	21,69 %
	Leg	0,00 %	10,09 %	52,81 %	0,00 %	18,97 %	81,03 %
Subdestinatario	Esp I	33,33 %	5,50 %	1,12 %	84,78 %	13,04 %	2,17 %
	Esp NI	25,64 %	9,17 %	5,62 %	66,67 %	22,22 %	11,11 %
	Ins SE	17,95 %	10,09 %	3,37 %	60,00 %	31,43 %	8,57 %
	Ins AE	16,24 %	14,68 %	0,00 %	54,29 %	45,71 %	0,00 %
	Ins LF	6,84 %	50,46 %	37,08 %	8,33 %	57,29 %	34,38 %
	Leg NF	0,00 %	10,09 %	52,81 %	0,00 %	18,97 %	81,03 %
IDT global	<0,3061	52,14 %	76,15 %	88,76 %	27,35 %	37,22 %	35,43 %
	>0,3061	47,86 %	23,85 %	11,24 %	60,87 %	28,26 %	10,87 %
IDT b)	<0,0363	20,51 %	46,79 %	69,66 %	17,52 %	37,23 %	45,26 %
	0,0363-0,1405	58,97 %	52,29 %	28,09 %	45,70 %	37,75 %	16,56 %
	>0,1405	20,51 %	0,92 %	2,25 %	88,89 %	3,70 %	7,41 %
IDT c)	<0,0540	36,75 %	68,81 %	89,89 %	21,72 %	37,88 %	40,40 %
	>0,0540	63,25 %	31,19 %	10,11 %	63,25 %	29,06 %	7,69 %
IR global	<0,065	56,41 %	85,32 %	92,13 %	27,39 %	38,59 %	34,02 %
	>0,065	43,59 %	14,68 %	7,87 %	68,92 %	21,62 %	9,46 %
IR a)	<0,0882	64,10 %	85,32 %	92,13 %	30,00 %	37,20 %	32,80 %
	>0,0882	35,90 %	14,68 %	7,87 %	64,62 %	24,62 %	10,77 %
IR c)	<0,0773	49,57 %	75,23 %	86,52 %	26,73 %	37,79 %	35,48 %
	>0,0773	50,43 %	24,77 %	13,48 %	60,20 %	27,55 %	12,24 %

Tabla 148: Probabilidades de clasificación de clase según datos de estudiantes

		1) Probabilidad del valor del atributo conociendo el valor de la clase			2) Probabilidad del valor de la clase conociendo el valor del atributo		
Valor del atributo		TE	TSE	TNE	TE	TSE	TNE
Tema	DER	36,70 %	26,85 %	36,73 %	38,10 %	27,62 %	34,29 %
	INF	29,36 %	38,89 %	31,63 %	30,48 %	40,00 %	29,52 %
	MED	33,94 %	34,26 %	31,63 %	35,24 %	35,24 %	29,52 %
Tipo texto	ART	38,53 %	2,78 %	0,00 %	93,33 %	6,67 %	0,00 %
	BLOG	8,26 %	20,37 %	14,29 %	20,00 %	48,89 %	31,11 %
	DIV	7,34 %	25,00 %	10,20 %	17,78 %	60,00 %	22,22 %
	ENT	6,42 %	20,37 %	16,33 %	15,56 %	48,89 %	35,56 %
	FOR	0,00 %	12,04 %	32,65 %	0,00 %	28,89 %	71,11 %
	NOT	1,83 %	15,74 %	26,53 %	4,44 %	37,78 %	57,78 %
	TES	37,61 %	3,70 %	0,00 %	91,11 %	8,89 %	0,00 %
Emisor	Esp	97,25 %	40,74 %	15,31 %	64,24 %	26,67 %	9,09 %
	Ins	2,75 %	58,33 %	71,43 %	2,21 %	46,32 %	51,47 %
	Leg	0,00 %	0,93 %	13,27 %	0,00 %	7,14 %	92,86 %
Subemisor	Esp I	83,49 %	10,19 %	3,06 %	86,67 %	10,48 %	2,86 %
	Esp NI	13,76 %	30,56 %	12,24 %	25,00 %	55,00 %	20,00 %
	Ins SE	2,75 %	46,30 %	42,86 %	3,16 %	52,63 %	44,21 %
	Ins AE	0,00 %	3,70 %	2,04 %	0,00 %	66,67 %	33,33 %
	Ins LF	0,00 %	8,33 %	26,53 %	0,00 %	25,71 %	74,29 %
	Leg NF	0,00 %	0,93 %	13,27 %	0,00 %	7,14 %	92,86 %
Destinatario	Esp	69,72 %	2,78 %	5,10 %	90,48 %	3,57 %	5,95 %
	Ins	30,28 %	86,11 %	42,86 %	19,64 %	55,36 %	25,00 %
	Leg	0,00 %	11,11 %	52,04 %	0,00 %	19,05 %	80,95 %
Subdestinatario	Esp I	53,21 %	0,00 %	1,02 %	98,31 %	0,00 %	1,69 %
	Esp NI	16,51 %	2,78 %	4,08 %	72,00 %	12,00 %	16,00 %
	Ins SE	13,76 %	10,19 %	2,04 %	53,57 %	39,29 %	7,14 %
	Ins AE	13,76 %	32,41 %	4,08 %	27,78 %	64,81 %	7,41 %
	Ins LF	2,75 %	43,52 %	36,73 %	3,49 %	54,65 %	41,86 %
	Leg NF	0,00 %	11,11 %	52,04 %	0,00 %	19,05 %	80,95 %
IDT global	<0,0925	4,59 %	12,96 %	47,96 %	7,58 %	21,21 %	71,21 %
	0,0925-0,1803	20,18 %	50,00 %	39,80 %	19,13 %	46,96 %	33,91 %
	>0,1803	75,23 %	37,04 %	12,24 %	61,19 %	29,85 %	8,96 %
IDT b)	<0,053	33,94 %	41,67 %	80,61 %	22,98 %	27,95 %	49,07 %
	>0,053	66,06 %	58,33 %	19,39 %	46,75 %	40,91 %	12,34 %
IDT c)	<0,0016	6,42 %	1,85 %	33,67 %	16,67 %	4,76 %	78,57 %
	0,0016-0,0612	36,70 %	65,74 %	60,20 %	23,53 %	41,76 %	34,71 %
	0,0612-0,1906	40,37 %	32,41 %	6,12 %	51,76 %	41,18 %	7,06 %
	>0,1906	16,51 %	0,00 %	0,00 %	100,00 %	0,00 %	0,00 %
IR global	<0,0643	44,04 %	76,85 %	82,65 %	22,64 %	39,15 %	38,21 %
	>0,0643	55,96 %	22,22 %	17,35 %	59,22 %	23,30 %	16,50 %
IR a)	<0,0752	47,71 %	76,85 %	81,63 %	24,19 %	38,60 %	37,21 %
	>0,0752	52,29 %	23,15 %	17,35 %	57,00 %	25,00 %	17,00 %
IR c)	<0,0102	25,69 %	50,00 %	64,29 %	19,31 %	37,24 %	43,45 %
	>0,0102	74,31 %	50,00 %	34,69 %	47,65 %	31,76 %	20,00 %

Tabla 149: Probabilidades de clasificación de clase según datos de usuarios medios

	Valor del atributo	1) Probabilidad del valor del atributo conociendo el valor de la clase			2) Probabilidad del valor de la clase conociendo el valor del atributo		
		TE	TSE	TNE	TE	TSE	TNE
Tema	DER	36,19 %	27,88 %	35,85 %	36,19 %	27,62 %	36,19 %
	INF	29,52 %	40,38 %	30,19 %	29,52 %	40,00 %	30,48 %
	MED	34,29 %	31,73 %	33,96 %	34,29 %	31,43 %	34,29 %
Tipo texto	ART	41,90 %	0,96 %	0,00 %	97,78 %	2,22 %	0,00 %
	BLOG	5,71 %	23,08 %	14,15 %	13,33 %	53,33 %	33,33 %
	DIV	6,67 %	24,04 %	12,26 %	15,56 %	55,56 %	28,89 %
	ENT	5,71 %	21,15 %	16,04 %	13,33 %	48,89 %	37,78 %
	FOR	0,00 %	11,54 %	31,13 %	0,00 %	26,67 %	73,33 %
	NOT	1,90 %	14,42 %	26,42 %	4,44 %	33,33 %	62,22 %
	TES	38,10 %	4,81 %	0,00 %	88,89 %	11,11 %	0,00 %
Emisor	Esp	94,29 %	36,54 %	11,32 %	66,44 %	25,50 %	8,05 %
	Ins	5,71 %	62,50 %	76,42 %	3,95 %	42,76 %	53,29 %
	Leg	0,00 %	0,96 %	12,26 %	0,00 %	7,14 %	92,86 %
Subemisor	Esp I	80,95 %	5,77 %	1,89 %	91,40 %	6,45 %	2,15 %
	Esp NI	13,33 %	30,77 %	9,43 %	25,00 %	57,14 %	17,86 %
	Ins SE	5,71 %	50,00 %	41,51 %	5,88 %	50,98 %	43,14 %
	Ins AE	0,00 %	1,92 %	1,89 %	0,00 %	50,00 %	50,00 %
	Ins LF	0,00 %	10,58 %	33,02 %	0,00 %	23,91 %	76,09 %
	Leg NF	0,00 %	0,96 %	12,26 %	0,00 %	7,14 %	92,86 %
Destinatario	Esp	73,33 %	3,85 %	4,72 %	89,53 %	4,65 %	5,81 %
	Ins	26,67 %	84,62 %	33,96 %	18,42 %	57,89 %	23,68 %
	Leg	0,00 %	11,54 %	61,32 %	0,00 %	15,58 %	84,42 %
Subdestinatario	Esp I	59,05 %	0,00 %	0,94 %	98,41 %	0,00 %	1,59 %
	Esp NI	14,29 %	3,85 %	3,77 %	65,22 %	17,39 %	17,39 %
	Ins SE	11,43 %	13,46 %	1,89 %	42,86 %	50,00 %	7,14 %
	Ins AE	12,38 %	21,15 %	2,83 %	34,21 %	57,89 %	7,89 %
	Ins LF	2,86 %	50,00 %	29,25 %	3,49 %	60,47 %	36,05 %
	Leg NF	0,00 %	11,54 %	61,32 %	0,00 %	15,58 %	84,42 %
IDT global	<0,100	1,90 %	8,65 %	37,74 %	3,92 %	17,65 %	78,43 %
	0,100-0,1880	17,14 %	43,27 %	39,62 %	17,14 %	42,86 %	40,00 %
	>0,1880	80,95 %	48,08 %	22,64 %	53,46 %	31,45 %	15,09 %
IDT b)	<0,0506	32,38 %	39,42 %	75,47 %	21,94 %	26,45 %	51,61 %
	>0,0506	67,62 %	60,58 %	24,53 %	44,38 %	39,38 %	16,25 %
IDT c)	<0,0014	5,71 %	3,85 %	31,13 %	13,95 %	9,30 %	76,74 %
	0,0014-0,0572	39,05 %	57,69 %	58,49 %	25,15 %	36,81 %	38,04 %
	>0,0572	55,24 %	38,46 %	10,38 %	53,21 %	36,70 %	10,09 %
IR global	<0,0576	39,05 %	76,92 %	86,79 %	19,25 %	37,56 %	43,19 %
	>0,0576	60,95 %	23,08 %	13,21 %	62,75 %	23,53 %	13,73 %
IR a)	<0,0757	50,48 %	80,77 %	87,74 %	23,04 %	36,52 %	40,43 %
	>0,0757	49,52 %	19,23 %	12,26 %	61,18 %	23,53 %	15,29 %
IR c)	<0,1329	58,10 %	87,50 %	92,45 %	24,40 %	36,40 %	39,20 %
	>0,1329	41,90 %	12,50 %	7,55 %	67,69 %	20,00 %	12,31 %

Tabla 150: Probabilidades de clasificación de clase según datos de traductores

	Valor del atributo	1) Probabilidad del valor del atributo conociendo el valor de la clase			2) Probabilidad del valor de la clase conociendo el valor del atributo		
		TE	TSE	TNE	TE	TSE	TNE
Tema	DER	37,25 %	24,76 %	37,96 %	36,19 %	24,76 %	39,05 %
	INF	29,41 %	40,95 %	29,63 %	28,57 %	40,95 %	30,48 %
	MED	33,33 %	34,29 %	32,41 %	32,38 %	34,29 %	33,33 %
Tipo texto	ART	42,16 %	1,90 %	0,00 %	95,56 %	4,44 %	0,00 %
	BLOG	6,86 %	20,95 %	14,81 %	15,56 %	48,89 %	35,56 %
	DIV	4,90 %	26,67 %	11,11 %	11,11 %	62,22 %	26,67 %
	ENT	4,90 %	20,95 %	16,67 %	11,11 %	48,89 %	40,00 %
	FOR	0,00 %	10,48 %	31,48 %	0,00 %	24,44 %	75,56 %
	NOT	0,98 %	15,24 %	25,93 %	2,22 %	35,56 %	62,22 %
	TES	40,20 %	3,81 %	0,00 %	91,11 %	8,89 %	0,00 %
Emisor	Esp	94,12 %	27,62 %	8,33 %	71,64 %	21,64 %	6,72 %
	Ins	5,88 %	72,38 %	79,63 %	3,57 %	45,24 %	51,19 %
	Leg	0,00 %	0,00 %	12,04 %	0,00 %	0,00 %	100,00 %
Subemisor	Esp I	87,25 %	4,76 %	1,85 %	92,71 %	5,21 %	2,08 %
	Esp NI	6,86 %	22,86 %	6,48 %	18,42 %	63,16 %	18,42 %
	Ins SE	5,88 %	59,05 %	44,44 %	5,17 %	53,45 %	41,38 %
	Ins AE	0,00 %	0,95 %	0,93 %	0,00 %	50,00 %	50,00 %
	Ins LF	0,00 %	12,38 %	34,26 %	0,00 %	26,00 %	74,00 %
	Leg NF	0,00 %	0,00 %	12,04 %	0,00 %	0,00 %	100,00 %
Destinatario	Esp	80,39 %	3,81 %	4,63 %	90,11 %	4,40 %	5,49 %
	Ins	19,61 %	78,10 %	33,33 %	14,49 %	59,42 %	26,09 %
	Leg	0,00 %	18,10 %	62,04 %	0,00 %	22,09 %	77,91 %
Subdestinatario	Esp I	73,53 %	0,00 %	0,93 %	98,68 %	0,00 %	1,32 %
	Esp NI	6,86 %	3,81 %	3,70 %	46,67 %	26,67 %	26,67 %
	Ins SE	9,80 %	8,57 %	0,93 %	50,00 %	45,00 %	5,00 %
	Ins AE	9,80 %	21,90 %	2,78 %	27,78 %	63,89 %	8,33 %
	Ins LF	0,00 %	47,62 %	29,63 %	0,00 %	60,98 %	39,02 %
	Leg NF	0,00 %	18,10 %	62,04 %	0,00 %	22,09 %	77,91 %
IDT global	<0,2264	28,43 %	55,24 %	81,48 %	16,57 %	33,14 %	50,29 %
	>0,2264	71,57 %	44,76 %	18,52 %	52,14 %	33,57 %	14,29 %
IDT b)	<0,0559	40,20 %	45,71 %	80,56 %	23,30 %	27,27 %	49,43 %
	0,0559-0,1455	38,24 %	52,38 %	18,52 %	34,21 %	48,25 %	17,54 %
	>0,1455	21,57 %	1,90 %	0,93 %	88,00 %	8,00 %	4,00 %
IDT c)	<0,0013	6,86 %	1,90 %	33,33 %	15,56 %	4,44 %	80,00 %
	0,0013-0,0581	38,24 %	60,00 %	61,11 %	23,21 %	37,50 %	39,29 %
	>0,0581	54,90 %	38,10 %	5,56 %	54,90 %	39,22 %	5,88 %
IR global	<0,0524	39,22 %	80,00 %	87,96 %	18,26 %	38,36 %	43,38 %
	>0,0524	60,78 %	20,00 %	12,04 %	64,58 %	21,88 %	13,54 %
IR a)	<0,0646	48,04 %	81,90 %	88,89 %	21,21 %	37,23 %	41,56 %
	>0,0646	51,96 %	18,10 %	11,11 %	63,10 %	22,62 %	14,29 %
IR c)	<0,0916	51,96 %	80,00 %	89,81 %	22,65 %	35,90 %	41,45 %
	>0,0916	48,04 %	20,00 %	10,19 %	60,49 %	25,93 %	13,58 %

Tabla 151: Probabilidades de clasificación de clase según datos del investigador

Las diferencias entre temas, como en el caso del análisis global, no son relevantes, ya que el tema no determina la clase y viceversa.

En el caso del tipo textual, especialmente en el bloque 1, las diferencias son escasas y también coherentes, puesto que las diferencias se aplican específicamente a la probabilidad para la tesis doctoral y el artículo en revista científica (*vid. infra* Tabla 152), que son las que están estrechamente relacionadas, tanto con la clase de texto especializado como con un elevado IDT. En cuanto al bloque 2, que determina la probabilidad de clasificación en una clase cuando conocemos el tipo textual, dicha probabilidad está compartida entre todos los informantes, especialmente entre el traductor y el investigador, que muestran valores numéricos muy próximos (Tabla 150 y Tabla 151).

	TEXTO ESPECIALIZADO		TEXTO SEMIESPECIALIZADO		TEXTO NO ESPECIALIZADO	
	Tipo texto	Informante	Tipo texto	Informante	Tipo texto	Informante
Menor probabilidad	FOR	todos	TES	Estudiante	No hay diferencias	
			ART	Us. med Traductor Investigador		
Mayor probabilidad	TES	Estudiante	DIV	Todos	No hay diferencias	
	ART	Us. med Traductor Investigador				

Tabla 152: Diferencias para menor y mayor probabilidad de tipo textual según clase según informantes

En relación con el emisor, no existen diferencias para el primer bloque, si bien se identifica una leve divergencia en el estudiante en el segundo bloque. En efecto, el análisis del estudiante diverge en la mayor probabilidad que existe de texto semiespecializado cuando el emisor es instruido (el resto de los informantes determina que la mayor probabilidad corresponde con el texto no especializado). De cualquier forma, tanto la probabilidad de TSE como TNE para el emisor instruido ronda el 50 % por lo que no puede defenderse que el informante estudiante diverja en exceso. En cuanto al subemisor, se repite el patrón del emisor: el informante estudiante diverge del resto de informantes. La diferencia más clara se encuentra en la vinculación del especialista no investigador con el texto especializado (50,77 % para grupo de estudiantes) y no para texto semiespecializado (como en el resto de informantes, que duplican esta probabilidad). Estos datos se relacionan en la Tabla 153.

		Estudiante	Usuario medio	Traductor	Investigador
Subemisor	Especialista investigador	TE 70,91 %	86,67 %	91,40 %	92,71 %
		TSE 22,73 %	10,48 %	6,45 %	5,21 %
		TNE 6,36 %	2,86 %	2,15 %	2,08 %
Especialista no investigador		TE 50,77 %	25,00 %	25,00 %	18,42 %
		TSE 30,77 %	55,00 %	57,14 %	63,16 %
		TNE 18,46 %	20,00 %	17,86 %	18,42 %

Tabla 153: Diferencias para menor y mayor probabilidad de clase según subemisor según informantes

Tanto IDT como IR presentan similitudes relevantes. En el caso de los tipos de IDT, todos los informantes presentan una relación directa entre el aumento del valor de IDT global y el aumento de probabilidad de clasificación en texto especializado y viceversa. En el bloque 1, el estudiante presenta más diferencias para el IDT global, debido a la relevancia de las FLE délficas y el IDT de términos délficos en el cómputo del IDT global. En el caso del IDT de términos semicrípticos, el investigador es el que presenta ligeras divergencias pero los valores porcentuales son tan similares que no se puede defender una divergencia relevante.

En cuanto al IR, existen similitudes entre todos los informantes en los dos bloques. Como en el caso del IDT, el aumento de los valores de IR conlleva mayor probabilidad de TE. Tan solo el usuario medio presenta una diferencia notable en la probabilidad de que el IR por variación sea mayor a 0,0102 cuando se conoce que el texto es TE: mientras que el resto de informantes presenta valores cercanos al 50 %, el usuario medio cuenta con un 74,3 % de probabilidades. Esto puede justificarse teniendo en cuenta que el usuario medio es el informante que presenta más relación entre el texto especializado y la reformulación que emplea formas acronímicas y variantes denominativas.

3.2. Relevancia de los atributos

La regresión logística¹⁶⁸ permite conocer la relevancia o *peso* que tiene cada valor de cada atributo del análisis en la clasificación del texto especializado y del texto no especializado. Esta relevancia no establece una relación directa, es decir, no implica que el valor del atributo conlleve la clasificación en una de las dos clases de forma obligatoria, sino que su presencia o ausencia en el texto está muy relacionada con la clasificación.

La Tabla 154 (*vid. infra*) relaciona los valores numéricos para cada atributo para el TE y para el TNE en el conjunto de los datos globales del corpus. Se marcan en verde los valores más relevantes y en rojo los menos relevantes para cada clase.

Atributo	Valor del atributo	TE	TNE
Tema	Derecho	1,6183	1,1395
	Informática	0,9145	1,0079
	Medicina	0,6757	0,8707
Tipo texto	Artículo en revista científica	12,0622	0,3409
	Blog	0,4100	1,6816
	Entrevista	1,5849	1,1084

¹⁶⁸ *Vid. supra* apartado III. 7.3. (p. 233) para conocer más información sobre el funcionamiento del algoritmo de regresión logística y su relevancia en esta tesis doctoral.

Atributo	Valor del atributo	TE	TNE
	Noticia	0,4983	2,1544
	Participación en foro	0,1838	2,4243
	Revista o web de divulgación especializada	0,3859	0,8876
	Tesis doctoral	3,6088	0,3395
Emisor	Especialista	8,9010	0,7940
	Instruido	0,3048	1,0432
	Legó	0,0058	2,6230
Subemisor	Especialista investigador	5,5855	0,9794
	Especialista no investigador	3,3074	0,6907
	Instruido semiespecialista	33,3019	0,7584
	Instruido aprendiz de especialista	111,8594	0,5273
	Instruido legó formado	0,0001	2,0495
	Legó no formado	0,0058	2,6230
Destinatario	Especialista	19,3088	0,8664
	Instruido	4,1138	0,6342
	Legó	0,0043	2,2647
Subdestinatario	Especialista investigador	18,8402	1,5436
	Especialista no investigador	5,7857	0,2913
	Instruido semiespecialista	7,1861	0,4521
	Instruido aprendiz de especialista	4,2838	0,4721
	Instruido legó formado	1,1686	1,1874
	Legó no formado	0,0043	2,2647
	IDT global	6,9385	0,0105
	IDT a)	0,0565	24,4896
	IDT b)	336521,7447	0,0000
	IDT c)	1342,9754	0,0000
	IR global	3,0645	0,0005
	IR a)	1,3539	6,6982
	IR b)	0,0000	146,9385
	IR c)	20,8491	0,7838
	TTR estandarizada	0,0834	0,9115

Tabla 154: Relevancia de los valores de los atributos en la discriminación del texto especializado y no especializado según datos globales

En la Tabla 155 se contrastan los seis atributos más y menos relevantes para cada clase y se marca en rojo la fuente de los atributos que coinciden en ambas clases. Como se puede comprobar al contrastar estos atributos con fuente en rojo, cuanto más relevante es el atributo para una clase, menos lo es para la clase opuesta.

Atributo	Valor del atributo	TE	TNE	Valor del atributo	Atributo
IR b)		0	146,9385		IR b)
Subemisor	Ins LF	0,0001	24,4896		IDT a)
Destinatario	Leg	0,0043	6,6982		IR a)
Subdestinatario	Leg NF	0,0043	2,623	Leg	Emisor
Emisor	Leg	0,0058	2,623	Leg NF	Emisor
Emisor	Leg NF	0,0058	2,4243	FOR	Tipo texto
Destinatario	Esp	19,3088	0,3395	TES	Tipo texto
IR c)		20,8491	0,2913	Esp NI	Subdestinatario

Atributo	Valor del atributo	TE	TNE	Valor del atributo	Atributo
Subemisor	Ins SE	33,3019	0,0105		IDT global
Subemisor	Ins AE	111,8594	0,0005		IR global
IDT c)		1342,9754	0		IDT b)
IDT b)		336521,745	0		IDT c)

Tabla 155: Contraste de la relevancia de los valores de los atributos en la discriminación del texto especializado y no especializado según datos globales

Así, para el TE, la densidad terminológica de las formas que son menos comunes en el discurso común, es decir, IDT de términos semicrípticos e IDT de términos crípticos, son las más relevantes. Le siguen el subemisor instruido aprendiz de especialista y el subemisor instruido semiespecialista. La relevancia viene determinada de forma negativa; es decir, se trata de dos atributos que son muy relevantes para *no* clasificar un texto como especializado, ya que el porcentaje de TE sobre el total con dichos subemisores no es representativo (*vid. infra* Tabla 156).

		TE	Total de textos	% de TE sobre el total
Subemisor	Instruido aprendiz de especialista	2	27	7,4 %
	Instruido semiespecialista	19	379	5 %

Tabla 156: Porcentaje de textos especializados sobre el total de textos emitidos por instruido aprendiz de especialista o instruido semiespecialista

Dentro de los atributos más representativos también se encuentra la reformulación con procedimientos de variación (IR c)) y el destinatario especialista (el 86 % de los textos con destinatario especialista está clasificado como texto especializado) (*vid. supra* Tabla 146).

Por el contrario, los atributos más relevantes en el texto no especializado corresponden con el IR b) e IDT a), que son los atributos que WEKA determinó como no válidos en el cálculo de la probabilidad (*vid. supra* IV. 3.1., p. 410.). En el caso del IDT a), la relación se establece con los tipos textuales que cuentan con un elevado IDT a) y un IDT b) o IDT c) reducido, como son los textos clasificados típicamente como no especializados (principalmente noticia y participación en foro, entre otros). De hecho, la participación en foro es otro de los atributos más relevantes para el texto no especializado, con un 71,7 % de representación de textos no especializados sobre el total de participaciones en foro. En cuanto al IR a), la relación también se establece de forma negativa, puesto que el IR a) cuenta con una relación directa con el grado de especialización textual. Por último, los dos atributos restantes son el emisor lego (90,9 % de textos no

especializados sobre el total de textos emitidos por legos) y el lego no formado (89,4 % de textos no especializados sobre el total de textos emitidos por legos no formados).

A continuación, se presentan las cuatro tablas (Tabla 157 a Tabla 160) con todos los valores de la ratio de cocientes para cada atributo para cada clase según los cuatro grupos de informantes.

Atributo	Valor del atributo	TE	TNE
Tema	Derecho	0,9836	0,7267
	Informática	1,3075	1,1068
	Medicina	0,7776	1,2432
Tipo texto	Artículo en revista científica	5,8255	1,1278
	Blog	0,2571	0,6663
	Entrevista	0,9126	0,8364
	Noticia	0,6796	1,1338
	Participación en foro	0,8651	2,2402
	Revista o web de divulgación especializada	0,2624	0,3673
	Tesis doctoral	4,7415	1,7054
Emisor	Especialista	19,2616	0,8375
	Instruido	0,6102	1,0350
	Lego	0,0002	1,6623
Subemisor	Especialista investigador	5,4254	0,7570
	Especialista no investigador	8,2747	1,1265
	Instruido semiespecialista	20,5603	1,1610
	Instruido aprendiz de especialista	129,6365	0,2968
	Instruido lego formado	0,0002	1,4749
	Lego no formado	0,0002	1,6623
Destinatario	Especialista	7,9553	1,0127
	Instruido	3,5991	0,3634
	Lego	0,0070	5,2728
Subdestinatario	Especialista investigador	4,5991	1,1071
	Especialista no investigador	6,8571	0,9208
	Instruido semiespecialista	5,6155	7,1135
	Instruido aprendiz de especialista	3,8850	0,0001
	Instruido lego formado	1,0722	10,6012
	Lego no formado	0,0070	5,2728
IDT global	IDT global	1,3962	0,0171
	IDT a)	0,0053	26,6278
	IDT b)	81145402,5112	0,0002
	IDT c)	367,5400	0,0000
IR global	IR global	0,3226	0,1144
	IR a)	0,0811	12054,0000
	IR b)	0,0000	0,0502
	IR c)	0,6984	0,2137
TTR estandarizada		1,3656	0,9776

Tabla 157: Relevancia de los valores de los atributos en la discriminación del texto especializado y no especializado según informantes estudiantes

Atributo	Valor del atributo	TE	TNE
Tema	Derecho	2,1976	0,8354
	Informática	0,5340	0,9836
	Medicina	0,8521	1,2170
Tipo texto	Artículo en revista científica	1603520,3283	1834,7793
	Blog	0,0280	0,0267
	Entrevista	1,3328	0,1059
	Noticia	0,0046	0,0259
	Participación en foro	0,0000	0,0443
	Revista o web de divulgación especializada	0,1340	0,0946
	Tesis doctoral	183470,7380	1774,2616
Emisor	Especialista	5,8046	1,3835
	Instruido	0,3508	0,8817
	Lego	0,0139	0,3074
Subemisor	Especialista investigador	2,2034	0,9498
	Especialista no investigador	5,5087	1,8211
	Instruido semiespecialista	11,0241	1,3345
	Instruido aprendiz de especialista	0,0001	0,8055
	Instruido lego formado	0,0027	0,4118
	Lego no formado	0,0139	0,3074
Destinatario	Especialista	119,3082	0,0471
	Instruido	1,5130	7,6203
	Lego	0,0015	1,7757
Subdestinatario	Especialista investigador	381,9821	0,0011
	Especialista no investigador	1,5095	410,6924
	Instruido semiespecialista	56,7752	20,4149
	Instruido aprendiz de especialista	2,3474	4,4313
	Instruido lego formado	0,1756	1,2846
	Lego no formado	0,0015	1,7757
IDT global	IDT global	5161190,9514	1787,8181
	IDT a)	0,0000	0,0000
	IDT b)	12553840440,9713	290390,8578
	IDT c)	3689545682187270,0000	3437873780,4473
IR global	IR global	0,0013	1,1601
	IR a)	852,9021	1,6228
	IR b)	0,0000	1,7897
	IR c)	508,4364	1,0414
TTR estandarizada		0,0000	0,0145

Tabla 158: Relevancia de los valores de los atributos en la discriminación del texto especializado y no especializado según informantes usuarios medios

Atributo	Valor del atributo	TE	TNE
Tema	Derecho	1,7537	0,7140
	Informática	1,1510	1,1387
	Medicina	0,4954	1,2299
Tipo texto	Artículo en revista científica	549065,3553	118,1915
	Blog	0,1120	0,0766
	Entrevista	0,5076	0,0898
	Noticia	0,1048	0,0545
	Participación en foro	0,0000	0,0953
	Revista o web de divulgación especializada	0,0562	0,0904
	Tesis doctoral	13062,5532	2621,8129
Emisor	Especialista	7,5692	1,2955
	Instruido	0,2410	0,9014
	Legó	0,0298	0,4027
Subemisor	Especialista investigador	2,9594	0,8280
	Especialista no investigador	6,7349	2,0345
	Instruido semiespecialista	47,0536	1,9662
	Instruido aprendiz de especialista	0,0008	1,6728
	Instruido legó formado	0,0001	0,2354
	Legó no formado	0,0298	0,4027
Destinatario	Especialista	245,0641	0,0649
	Instruido	2,1318	7,7172
	Legó	0,0010	1,1921
Subdestinatario	Especialista investigador	1317,0502	0,0025
	Especialista no investigador	0,4284	467,2746
	Instruido semiespecialista	3,7173	5,7070
	Instruido aprendiz de especialista	5,6610	4,2058
	Instruido legó formado	0,6005	2,9812
	Legó no formado	0,0010	1,1921
IDT global	IDT global	44029,4615	195,6691
	IDT a)	0,1868	0,2891
	IDT b)	9699908807,0687	24596,9900
	IDT c)	510246252,1574	115450,2951
IR global	IR global	0,0001	25,7277
	IR a)	16335,8224	9,4192
	IR b)	0,0000	0,0066
	IR c)	57912749,3006	60,8198
TTR estandarizada		16,2767	200,7133

Tabla 159: Relevancia de los valores de los atributos en la discriminación del texto especializado y no especializado según informantes traductores

Atributo	Valor del atributo	TE	TNE
Tema	Derecho	1,8480	0,7209
	Informática	0,6802	1,0000
	Medicina	0,7955	1,3872
Tipo texto	Artículo en revista científica	625,4337	77303,3041
	Blog	1,8336	0,0116
	Entrevista	3,9083	0,0233
	Noticia	0,4732	0,0086
	Participación en foro	0,0000	0,0070
	Revista o web de divulgación especializada	0,8454	0,0230
	Tesis doctoral	144,7646	34882,3479
Emisor	Especialista	18,5518	0,9814
	Instruido	0,3411	2,7941
	Legó	0,0000	0,0018
Subemisor	Especialista investigador	156,0808	0,2408
	Especialista no investigador	0,0348	16,4513
	Instruido semiespecialista	1,8230	3,0105
	Instruido aprendiz de especialista	0,0000	1,2966
	Instruido legó formado	0,1348	0,9829
	Legó no formado	0,0000	0,0018
Destinatario	Especialista	1519,1309	0,0170
	Instruido	1,3487	12,4168
	Legó	0,0004	2,9866
Subdestinatario	Especialista investigador	1483,9970	0,0010
	Especialista no investigador	40,7949	12932,9992
	Instruido semiespecialista	3615,5974	9,4291
	Instruido aprendiz de especialista	1356,6803	5,6537
	Instruido legó formado	0,0026	5,0375
	Legó no formado	0,0004	2,9866
IDT global	IDT global	1279,2208	5269,5060
	IDT a)	0,0000	0,0028
	IDT b)	111862373056,7	6244,4025
	IDT c)	176991472626,9	26185453655,5191
IR global	IR global	412,2163	412,6989
	IR a)	233,0507	22,1144
	IR b)	0,0159	0,0854
	IR c)	34343,5108	6,8388
TTR estandarizada		0,0208	3,7723

Tabla 160: Relevancia de los valores de los atributos en la discriminación del texto especializado y no especializado según investigador

A continuación, se relacionan las tablas que contrastan los valores más relevantes para clasificar el TE y el TNE según los cuatro grupos de informantes (Tabla 161 a Tabla 164) con el objetivo de contrastar si los cuatro grupos coinciden en la relevancia de los atributos para cada clase.

Atributo	Valor del atributo	TE	TNE	Valor del atributo	Atributo
IR b)		0,0000	12054,000		IR a)
Emisor	Leg	0,0002	26,6278		IDT a)
Subemisor	Ins LF	0,0002	10,6012	Ins LF	Subdestinatario
Subemisor	Leg NF	0,0002	7,1135	Ins SE	Subdestinatario
IDT a)		0,0053	5,2728	Leg	Destinatario
Destinatario	Leg	0,0070	5,2728	Leg NF	Subdestinatario
Subdestinatario	Leg NF	0,0070	0,1144		IR global
Subemisor	Esp NI	8,2747	0,0502		IR b)
Emisor	Esp	19,2616	0,0171		IDT global
Subemisor	Ins SE	20,5603	0,0002		IDT b)
Subemisor	Ins AE	129,6365	0,0001	Ins AE	Subdestinatario
IDT c)		367,5400	0,0000		IDT c)
IDT b)		81145402,511			

Tabla 161: Contraste de la relevancia de los valores de los atributos en la discriminación del texto especializado y no especializado según informantes estudiantes

Atributo	Valor del atributo	TE	TNE	Valor del atributo	Atributo
Tipo texto	FOR	0,0000	3437873780		IDT c)
IDT a)		0,0000	290390,8578		IDT b)
IR b)		0,0000	1834,7793	ART	Tipo texto
TTR estandarizada		0,0000	1787,8181		IDT global
Subemisor	Ins AE	0,0001	1774,2616	TES	Tipo texto
IR global		0,0013	410,6924	Esp NI	Subdestinatario
IR a)		852,9021	0,0443	FOR	Tipo texto
Tipo texto	TES	183470,7380	0,0267	BLOG	Tipo texto
Tipo texto	ART	1603520,3283	0,0259	NOT	Tipo texto
IDT global		5161190,9514	0,0145		TTR estandarizada
IDT b)		12553840440,9713	0,0011	Esp I	Subdestinatario
IDT c)		3689545682187270	0,0000		IDT a)

Tabla 162: Contraste de la relevancia de los valores de los atributos en la discriminación del texto especializado y no especializado según informantes usuarios medios

Atributo	Valor del atributo	TE	TNE	Valor del atributo	Atributo
Tipo texto	FOR	0,0000	115450		IDT c)
IR b)		0,0000	24596,990		IDT b)
Subemisor	Ins LF	0,0001	2621,8129	TES	Tipo texto
IR global		0,0001	467,2746	Esp NI	Subdestinatario
Subemisor	Ins AE	0,0008	200,7133		TTR estandarizada
Destinatario	Leg	0,0010	195,6691		IDT global
Subdestinatario	Leg NF	0,0010	0,0898	ENT	Tipo texto
IR a)		16335,8224	0,0766	BLOG	Tipo texto
IDT global		44029,4615	0,0649	Esp	Destinatario
Tipo texto	ART	549065,3553	0,0545	NOT	Tipo texto
IR c)		57912749,300	0,0066		IR b)
IDT c)		510246252	0,0025	Esp I	Subdestinatario
IDT b)		9699908807			

Tabla 163: Contraste de la relevancia de los valores de los atributos en la discriminación del texto especializado y no especializado según informantes traductores

Atributo	Valor del atributo	TE	TNE	Valor del atributo	Atributo
Tipo texto	FOR	0,0000	26185453655		IDT c)
Emisor	Leg	0,0000	77303,3041	ART	Tipo texto
Subemisor	Ins AE	0,0000	34882,3479	TES	Tipo texto
Subemisor	Leg NF	0,0000	12932,9992	Esp NI	Subdestinatario
IDT a)		0,0000	6244,4025		IDT b)
Destinatario	Leg	0,0004	5269,5060		IDT global
Subdestinatario	Lego NF	0,0004	0,0086	NOT	Tipo texto
Subdestinatario	Esp I	1483,9970	0,0070	FOR	Tipo texto
Destinatario	Esp	1519,1309	0,0028		IDT a)
Subdestinatario	Ins SE	3615,5974	0,0018	Leg	Emisor
IR c)		34343,5108	0,0018	Leg NF	Subemisor
IDT b)		111862373056	0,0010	Esp I	Subdestinatario
IDT c)		176991472626			

Tabla 164: Contraste de la relevancia de los valores de los atributos en la discriminación del texto especializado y no especializado según investigador

Los cuatro grupos de informantes muestran como criterio común en el texto especializado la relevancia de la densidad terminológica de las formas más opacas o poco frecuentes en el discurso común, como son el IDT b) y el IDT c). Otros criterios no relevantes comunes para el texto especializado son la participación en foro (excepto para el grupo de informantes) y el destinatario lego (excepto para el grupo de usuarios medios). En contraste, para el texto no especializado, los atributos más relevantes funcionan de forma negativa: se trata del IDT c) y de la tesis doctoral (excepto para el grupo de estudiantes). En este caso, la ausencia de estos valores contribuye a la clasificación del texto no especializado. Llama la atención que en el caso del grupo de estudiantes el IDT b) e IDT c) sean precisamente los criterios menos relevantes para el texto no especializado. Esto se justifica atendiendo a que se trata del grupo de informantes en el que la clase viene determinada de forma positiva por la presencia de otros elementos, como son un IDT a) elevado o un destinatario o subdestinatario lego y lego no formado respectivamente.

3.3. Árboles de decisión

A continuación, se presenta un árbol de decisión¹⁶⁹ dividido en siete partes (Ilustración 27 a Ilustración 34). Cada árbol presenta una distribución que atiende a los atributos (colores más oscuros) y a los valores para dichos atributos (colores más claros) hasta llegar a las *hojas*. Las hojas son el final del árbol de decisión y representan la

¹⁶⁹ Vid. *supra* apartado III. 7.4. (p. 235) para conocer con detalle el funcionamiento del algoritmo J48 y su aplicación en esta tesis doctoral.

correspondiente predicción de clase para el cumplimiento de todos los criterios o, en su defecto, cuando el árbol es excesivamente complejo o largo, representa otro atributo que inicia otro árbol de decisión (como, por ejemplo, Subemisor [1], IDT b) [2], etc.), que están marcados numéricamente entre corchetes (puede comprobarse en la rama intermedia de la Ilustración 27).

Dichos atributos marcados entre corchetes corresponden con cada uno de los árboles de decisión que se presentan entre la Ilustración 28 y la Ilustración 34. En la hoja de clase (la que finaliza una rama del árbol) se incluye tanto el valor (TE, TSE o TNE) como la relación de aciertos y errores del árbol de decisión, ya que a este se le aplicó la opción de poda. Así, la primera clase, TE, fruto de la relación destinatario especialista-emisor especialista (parte izquierda de la Ilustración 27), cuenta con 315 TE y 14 textos que cumplen con dichos criterios, pero no con la clasificación de TE.

Cuando la hoja de clase contiene únicamente un único valor numérico entre paréntesis quiere decir que el 100 % de los textos con dicha clase cumple con dichas condiciones. Por ejemplo, en la parte izquierda de la Ilustración 27, los textos con destinatario especialista, emisor instruido, con un IR b) menor que 0,0061 y un IDT b) menor que 0,0577 son clasificados como TNE. En nuestro corpus existen 6 textos con estas características y, además, todos los textos que cumplen esta conjunción de reglas son TNE.

Si la clase predicha contiene el valor (0.0) quiere decir que ninguno de los textos del corpus cumple con dicha regla, es decir, que no existen textos con esas condiciones en el corpus. Para facilitar la identificación de estos textos se han marcado en amarillo las hojas de clase que no cumplen con las reglas del árbol de decisión.

Cuantas más subdivisiones parten de cada atributo, mayor es la complejidad del árbol y la predicción de la clase. Las dos predicciones más simples (Ilustración 27) corresponden con *destinatario instruido + emisor lego = texto no especializado* y *destinatario especialista + emisor especialista = texto especializado*.

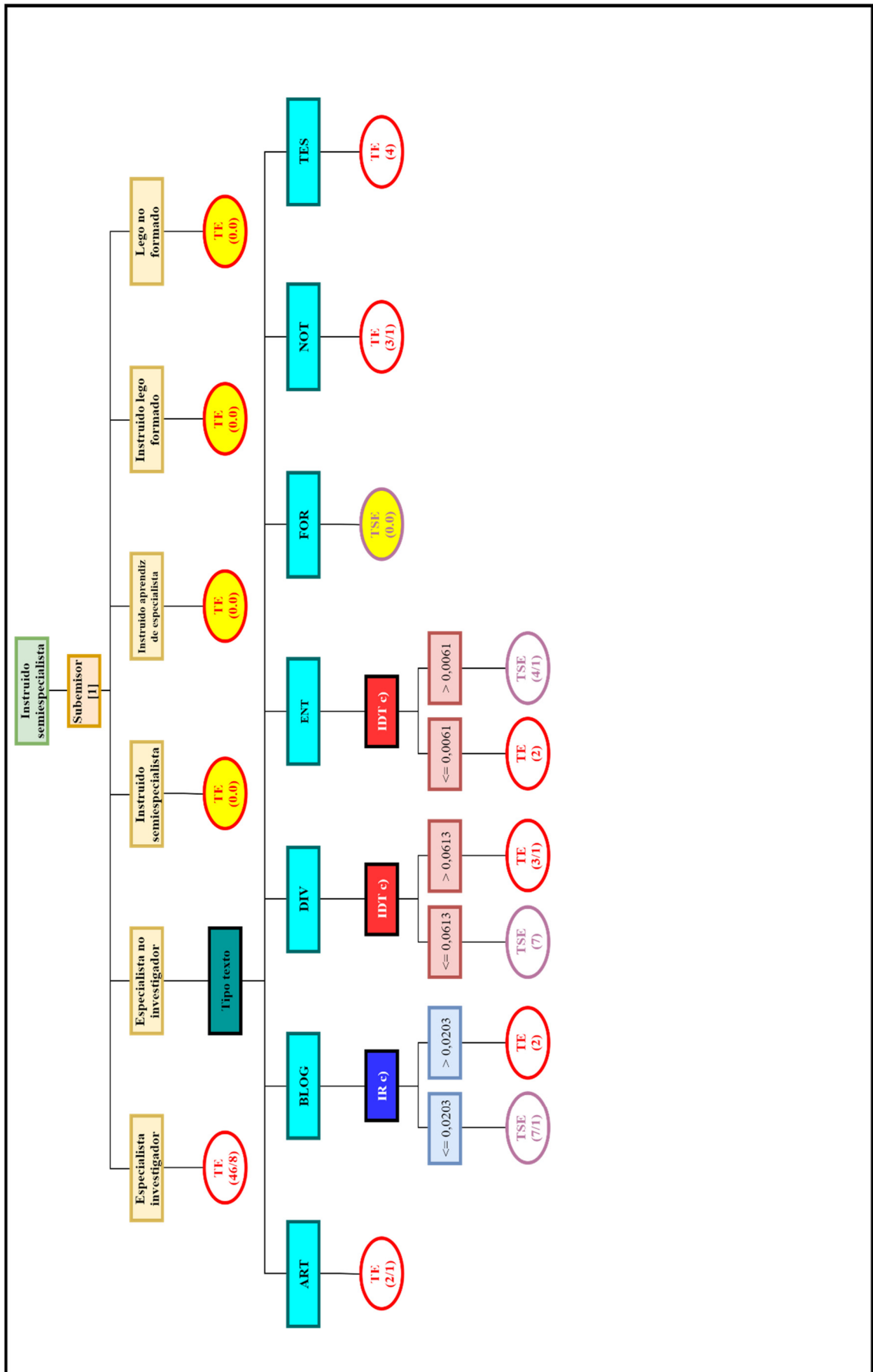


Ilustración 28: Árbol de decisión de clases según atributos [1]

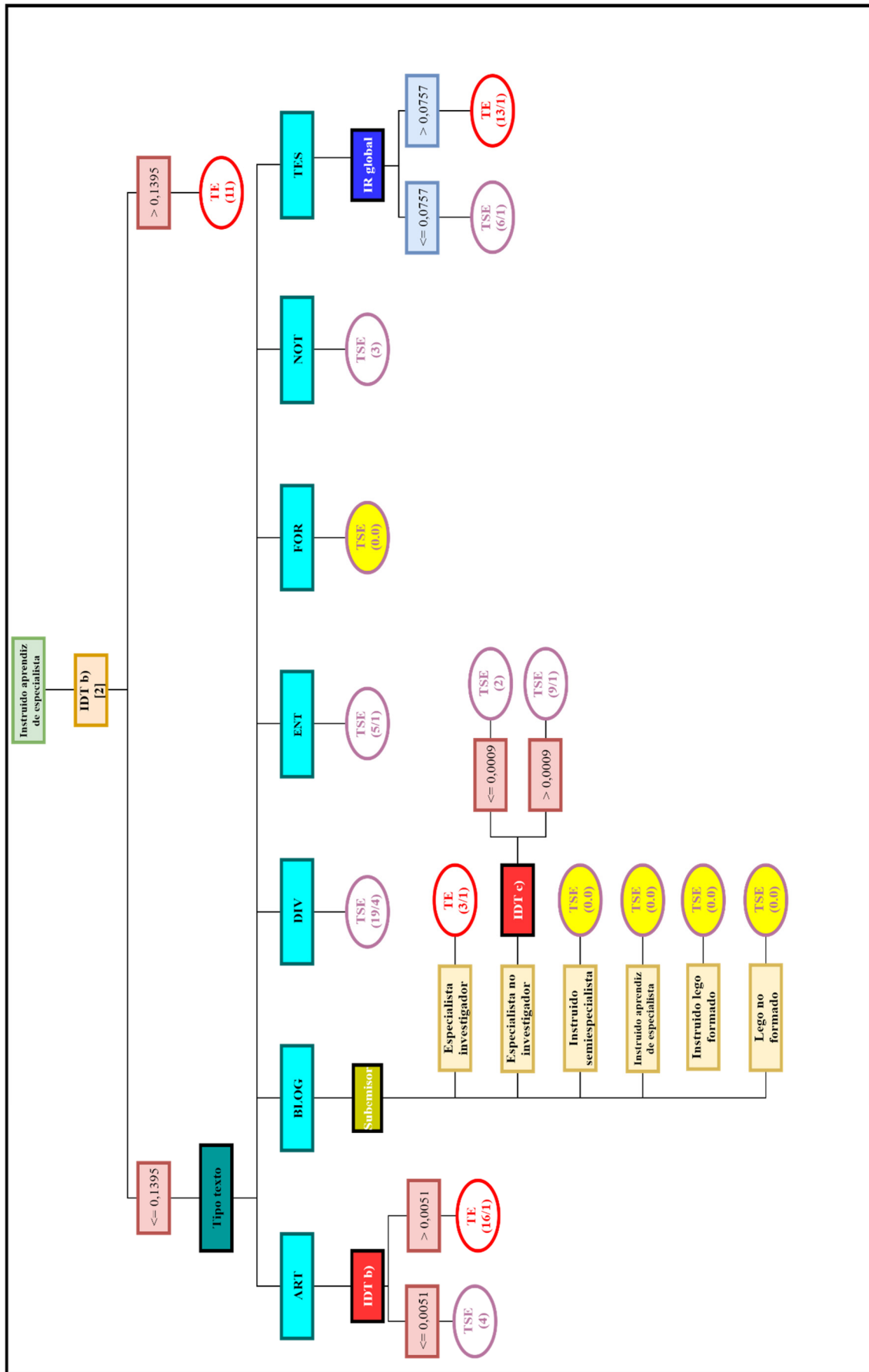


Ilustración 29: Árbol de decisión de clases según atributos [2]

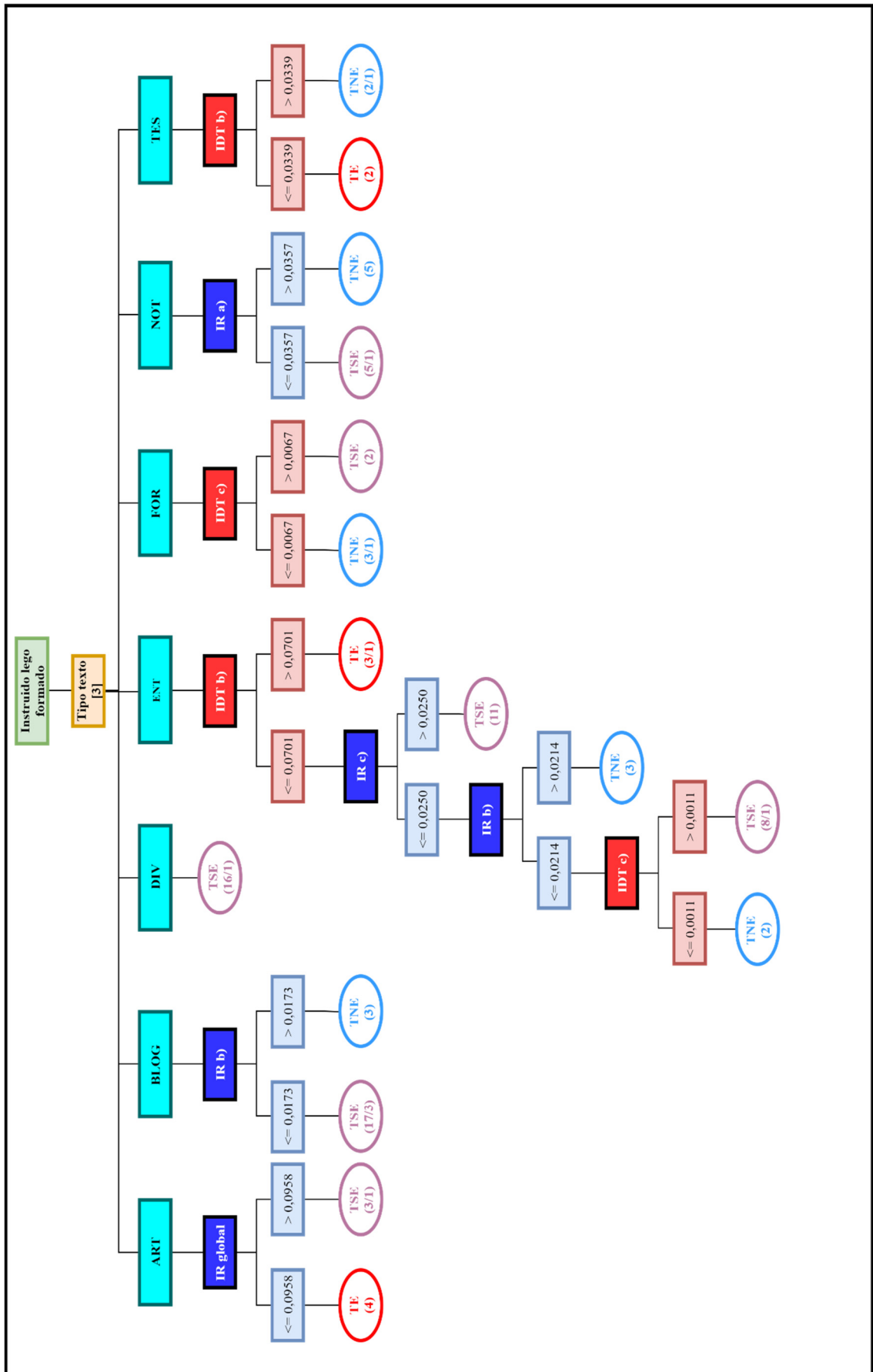


Ilustración 30: Árbol de decisión de clases según atributos [3]

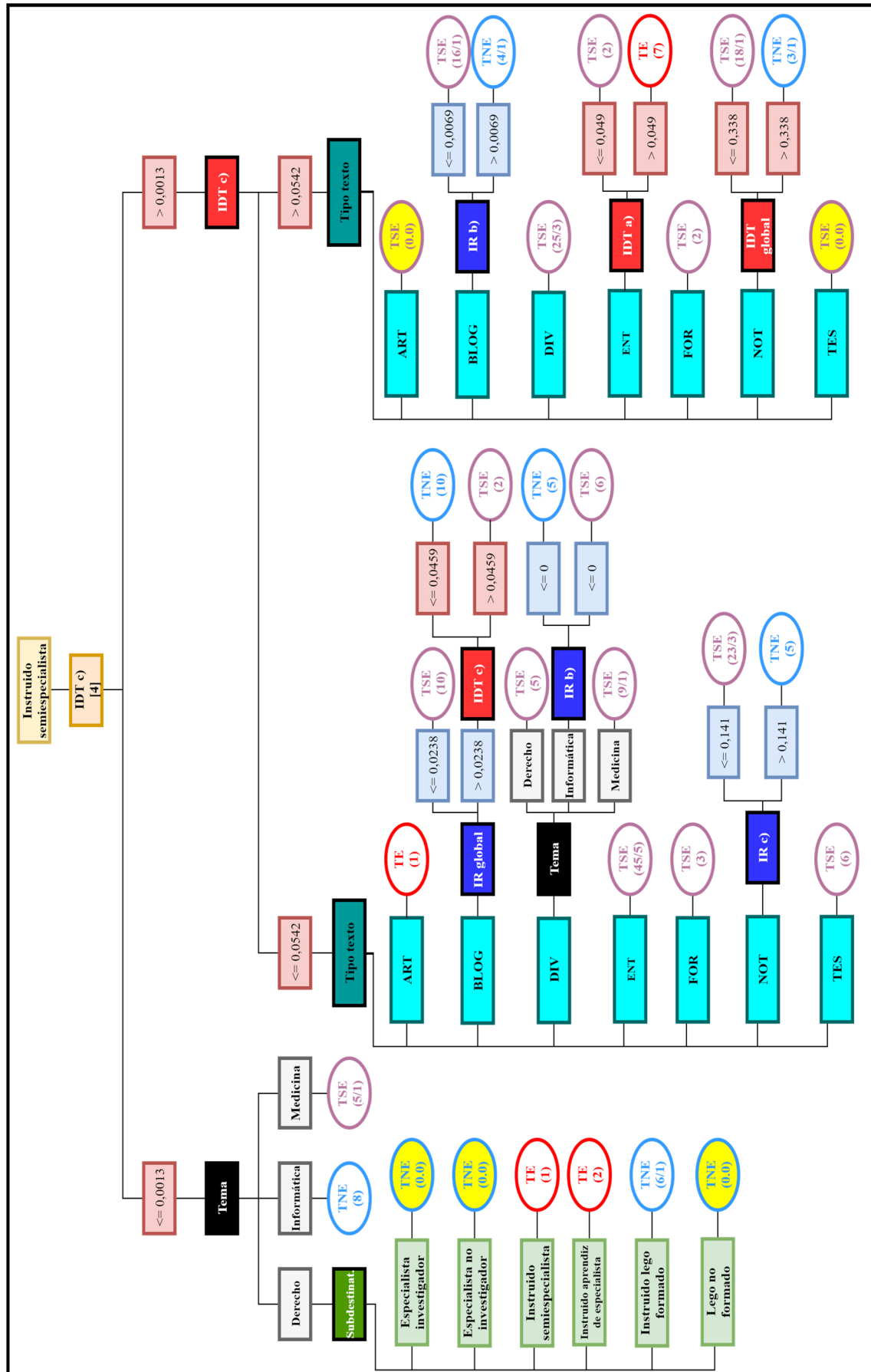


Ilustración 31: Árbol de decisión de clases según atributos [4]

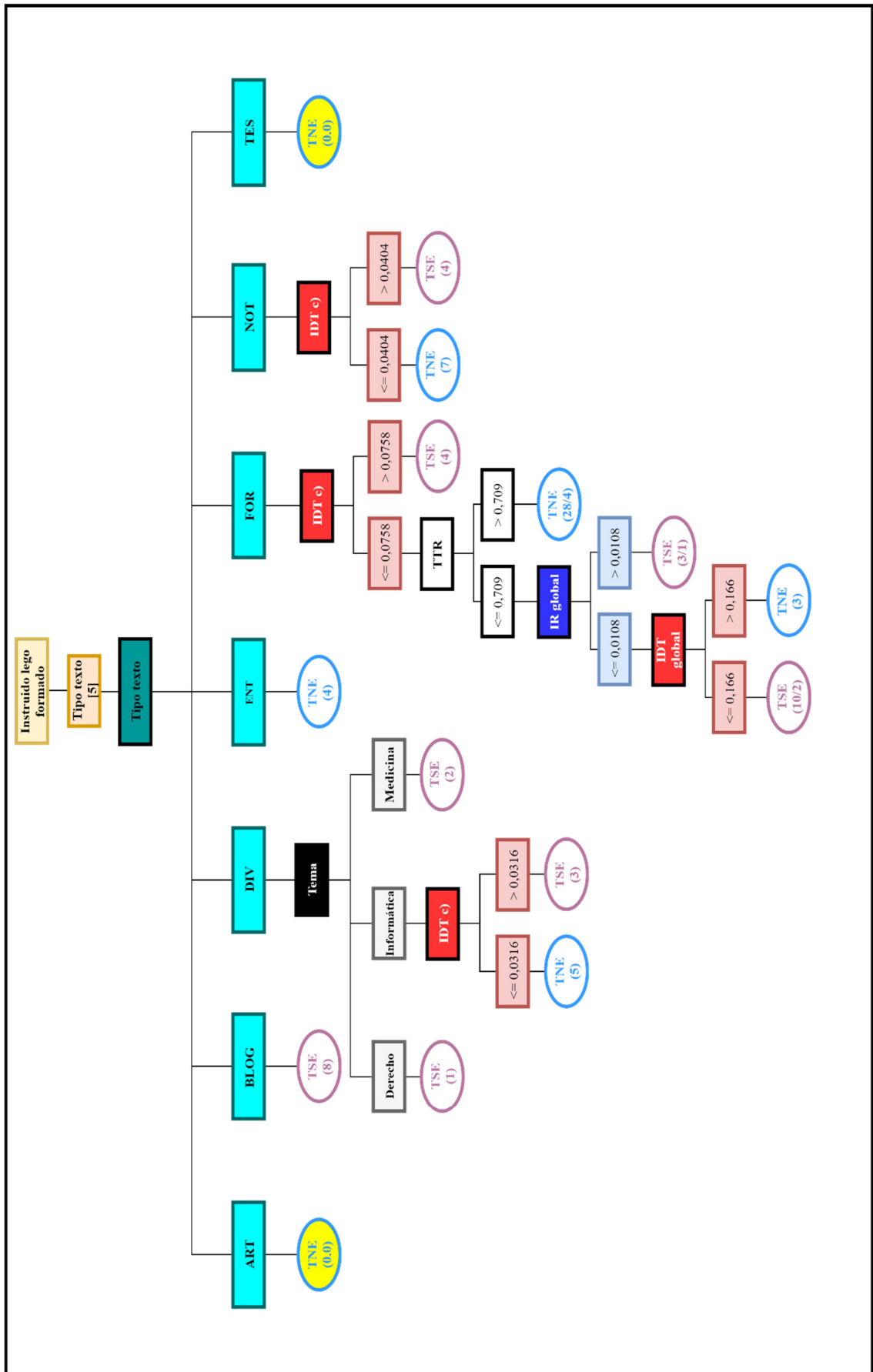


Ilustración 32: Árbol de decisión de clases según atributos [5]

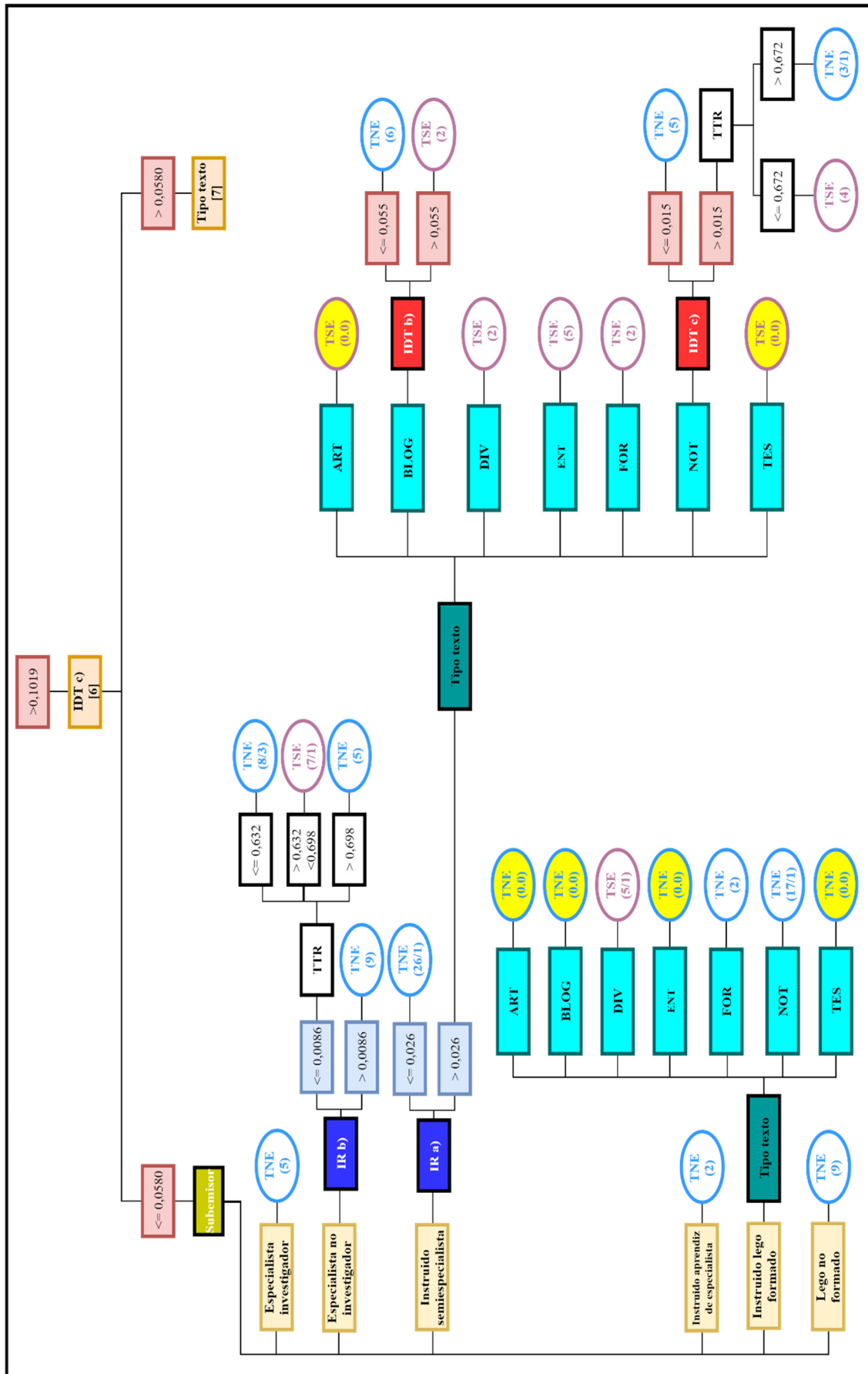


Ilustración 33: Árbol de decisión de clases según atributos [6]

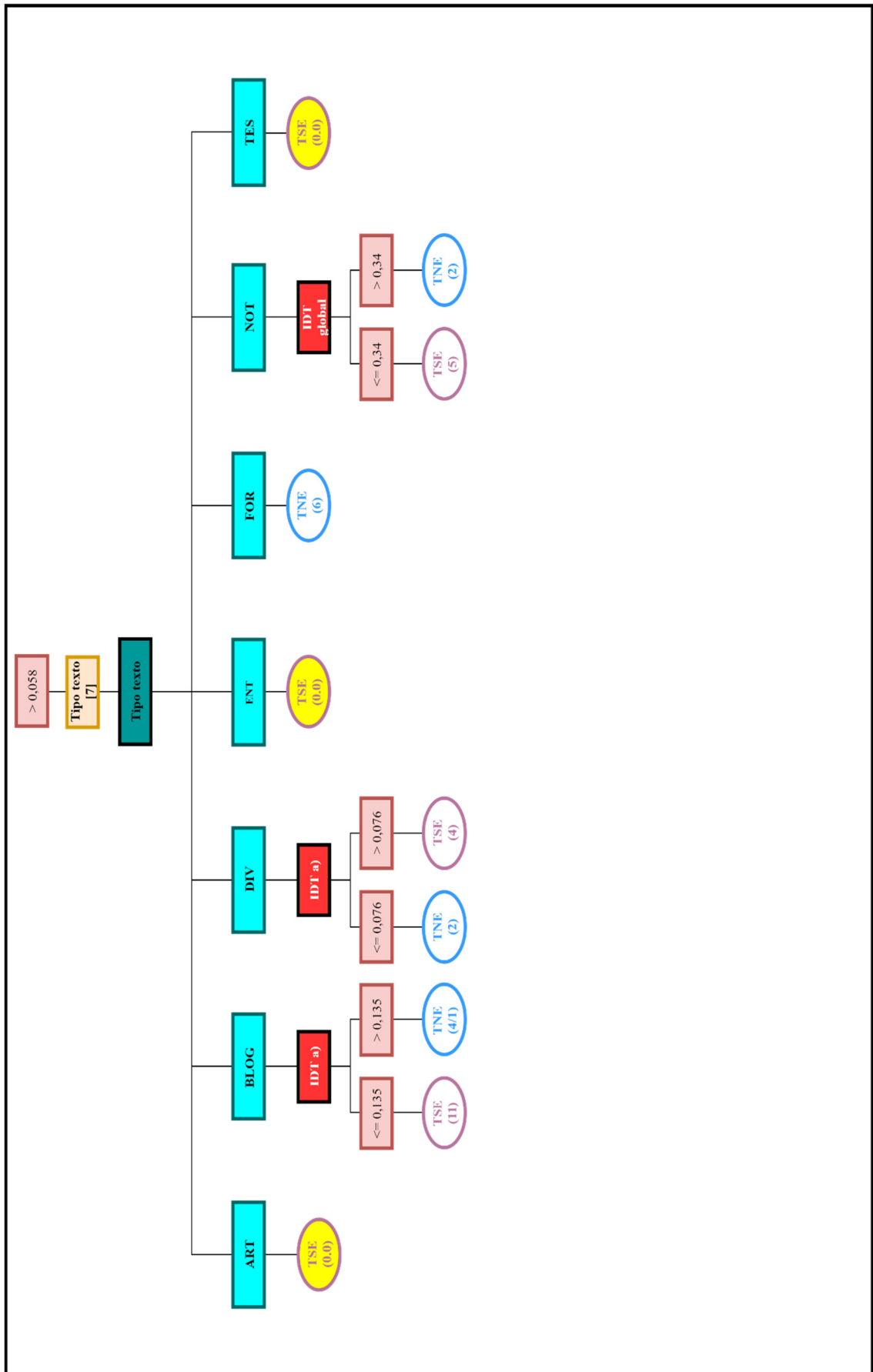


Ilustración 34: Árbol de decisión de clases según atributos [7]

El atributo más importante según el algoritmo J48 es el destinatario, a partir del cual se estructura todo el árbol. Tras este, el caso del destinatario especialista y el destinatario instruido, el atributo de emisor desempeña el segundo papel más relevante. Así, puede defenderse que la relación discursiva entre destinatario (como condicionante de la configuración discursiva) y emisor (como actualizador de la competencia cognitivo-comunicativa en el discurso) es la que establece las bases más importantes para distinguir los tres grados de especialización. En el caso del destinatario lego, el IDT c) y el IDT global son los atributos que discriminan las clases¹⁷⁰.

Como se comprueba en dicho árbol, los atributos numéricos suelen servir para discriminar entre dos clases atribuibles a una misma regla (como ocurre, por ejemplo, en la Ilustración 30) y suelen encontrarse próximos a la hoja, es decir, se integran en el árbol como parte final de la decisión. En cuanto a las propias clases, sirviéndonos de los colores para TSE y TNE, la polarización observada en el Gráfico 23 (*vid. supra* apartado IV. 1.2.1., p. 273) se aprecia al comprobar que el TE queda restringido principalmente a la relación discursiva especialista-especialista (*vid. supra* Ilustración 27, p. 430) y algunos otros casos más reducidos o dispersos (Ilustración 28, entre otras) y que el TSE y el TNE suelen compartir espacios, como se puede observar en la Ilustración 30, la Ilustración 31 o la Ilustración 32.

3.4. Grupos identificados según el aprendizaje automático

Los datos arrojados por el algoritmo simple k-means¹⁷¹ muestran tres grupos que coinciden con las tres clases según el grado de especialización. La Tabla 165 muestra los centroides o patrones representativos o prototípicos para dichas agrupaciones (o *clusters*), así como para el conjunto del corpus.

Atributo	Corpus	Cluster 0	Cluster 1	Cluster 2
Número de patrones	1260	452 (36 %)	471 (38 %)	335 (27 %)
Tema	Derecho	Derecho	Informática	Derecho
Tipo texto	Artículo en revista científica	Artículo en revista científica	Entrevista	Noticia
TTR	0,6555	0,6584	0,6508	0,6583

¹⁷⁰ En este caso no existe contradicción con los datos aportados por la regresión logística (*vid. supra* Tabla 155), en los que se determina que IDT global e IDT c) no son relevantes para determinar el TNE. En primer lugar, porque los datos de la Tabla 155 se aplican al conjunto de todo el corpus, en los que se determina que son otros los atributos que contribuyen a identificar el TNE. En segundo lugar, en la Ilustración 27 la combinación del IDT c) (en concreto de un IDT c) muy reducido) y el IDT global es la que conforma la regla que predice la clase de TNE.

¹⁷¹ *Vid. supra* apartado III. 7.5. (p. 235) para profundizar en el funcionamiento de simple k-means y los datos que permite obtener en esta tesis doctoral.

Atributo	Corpus	Cluster 0	Cluster 1	Cluster 2
IDT global	0,2175	0,2841	0,1907	0,1654
IDT a)	0,0998	0,1114	0,0863	0,1032
IDT b)	0,0624	0,0865	0,0544	0,0412
IDT c)	0,0553	0,0862	0,05	0,021
IR global	0,0555	0,0786	0,0516	0,0297
IR a)	0,0643	0,0949	0,0534	0,0384
IR b)	0,0156	0,0158	0,0216	0,0067
IR c)	0,0813	0,1349	0,0605	0,0385
Emisor	Especialista	Especialista	Instruido	Instruido
Subemisor	Especialista investigador	Especialista investigador	Instruido semiespecialista	Instruido semiespecialista
Destinatario	Instruido	Especialista	Instruido	Lego
Subdestinatario	Instruido lego formado	Especialista investigador	Instruido lego formado	Lego no formado
Clase	Texto especializado	Texto especializado	Texto semiespecializado	Texto no especializado

Tabla 165: Centroides para los tres grupos de clasificación según el algoritmo k-means

Los centroides que aquí se muestran tienen estrecha relación con todo el análisis llevado a cabo hasta el momento, puesto que suelen cumplirse el resto de reglas o condiciones estudiadas hasta ahora, tanto mediante aprendizaje automático como mediante análisis manual. En primer lugar, el número de patrones identificado es similar, si bien existen algunas diferencias, como la descompensación a favor del TSE y en contra del TNE (*vid. infra* Tabla 166), lo que contribuye a defender la estrecha relación entre estas clases.

		TE	TSE	TNE
Número de patrones	Según k-means	36 %	38 %	27 %
	Según corpus	34,4 %	33,8 %	31,8 %

Tabla 166: Contraste de porcentajes de representación de TE, TSE y TNE según el algoritmo k-means y los datos del corpus

En segundo lugar, los tipos textuales más representativos no corresponden con los más frecuentes en dichas clases (*vid. supra* Gráfico 23; apartado IV. 1.2.1., p. 273; aunque en comparación con las otras clases, sí que son más frecuentes en la clase del centroide) sino que se trata de aquellos tipos textuales que mejor representan (en conjunción con otros elementos) la clase en cuestión. Por ejemplo, la revista o web de divulgación especializada o el blog son los tipos textuales más frecuentes en el texto semiespecializado, pero también tienen presencia en el texto no especializado y en el texto especializado y, de hecho, comparten muchas características con dichas clases, por lo que no se constituyen como representativos del texto semiespecializado. Lo mismo sucede con el caso de la participación en foro en el texto no especializado, que, a pesar de ser la más frecuente, no es la más representativa, ya que cuenta con características que comparte con el texto semiespecializado. En estos casos, se demuestra cómo la frecuencia y la probabilidad no

necesariamente condicionan la caracterización de la clase (pero sí condicionan la clasificación dentro de una clase, lo cual es diferente). En otras palabras, la frecuencia de aparición de un valor de un atributo sobre el total del corpus sí que implica cierta probabilidad (*vid. supra* Tabla 146, p. 413) y esto conlleva cierta tendencia hacia una clase, mas dicha probabilidad no condiciona la caracterización de la clase (es decir, la conjunción de elementos representativos que se atribuyen o perfilan la clase).

El IDT y sus tipos y el IR y sus tipos también decrecen cuando la clase de las agrupaciones de k-means decrece¹⁷² (*vid. supra* Tabla 165), tal y como se defiende en los apartados IV. 1.4.2. y IV. 1.5.2. (pp. 310 y 326).

En cuanto a la relación de emisor y destinatario, también coinciden con los análisis llevados a cabo a lo largo de esta sección.

A continuación, se presentan diversos gráficos que contienen la información de la distribución de patrones por cada *cluster* (*vid. infra* Ilustración 35 a Ilustración 50). En estas ilustraciones se puede observar gráficamente la prominencia de un valor de un atributo en dicha agrupación o *cluster* (o, por el contrario, la ausencia de una vinculación estrecha con un *cluster*). En dichas ilustraciones, el color rojo representa el TE, el color morado representa el TSE y el color azul representa el TNE.

El primer caso más destacable es el que cruza las tres clases según el grado de especialización con las tres agrupaciones identificadas por k-means (*vid. infra* Ilustración 35). No solo los centroides (*vid. supra* Tabla 165) quedan vinculados con un valor de la clase, sino que gráficamente se puede observar la concentración de patrones en tres puntos: el cluster0 en el TE, el cluster1 en el TSE y el cluster2 en el TNE, que son prototípicos o representativos según la Tabla 165.

También se observa la relevancia del TSE como objeto de transición, puesto que los patrones excepcionales para el cluster0 y el cluster2 se posicionan en el TSE. Además, la relación de TSE con TNE también se identifica en la Ilustración 35, puesto que los patrones excepcionales tienen prominencia en el cluster1.

¹⁷² Esto no sucede con el IR b), que, recuérdese, no es un atributo relevante en el corpus (*vid. supra* apartados IV. 1.5.2., p. 329, y IV. 3.2., p. 425).

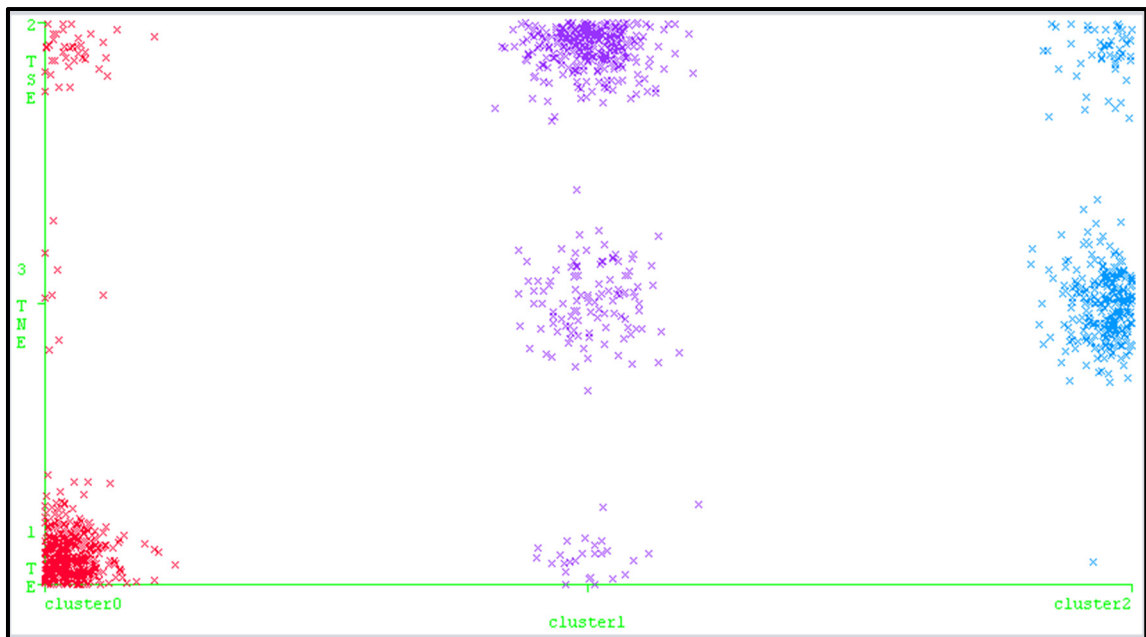


Ilustración 35: Agrupamiento de patrones por clase según algoritmo k-means

La imposibilidad de discriminación de clases del tema se identifica en la Ilustración 36, en el que todos los *clusters* tienen una representación más o menos homogénea en los tres temas del corpus, como se indica en el apartado IV. 1.1.1. (p. 244).

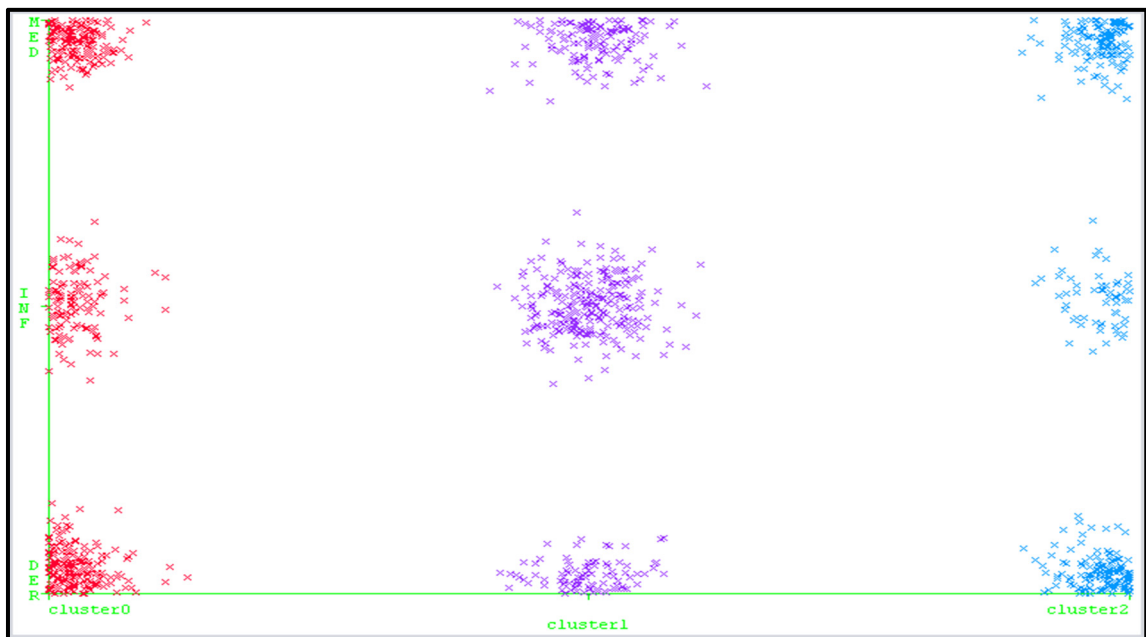


Ilustración 36: Agrupamiento de patrones por tema según algoritmo k-means

Esta ausencia de relación discriminadora también se repite en la distribución del agrupamiento por patrones de la TTR estandarizada (Ilustración 37), como se indicaba en el apartado IV. 1.6.7. (p. 356).

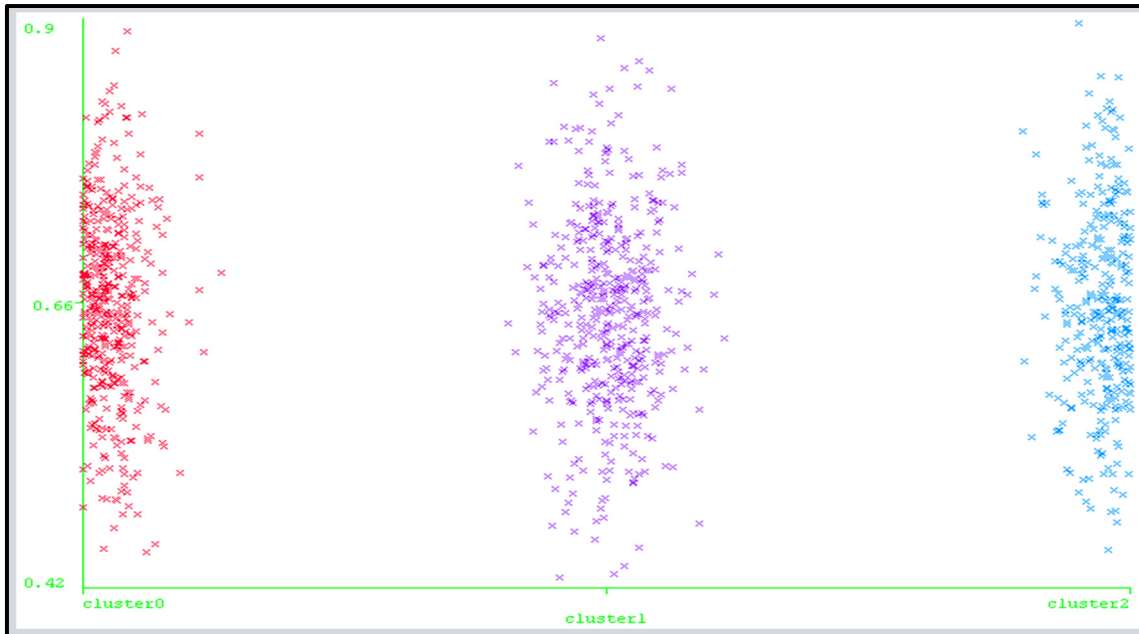


Ilustración 37: Agrupamiento de patrones por TTR estandarizada según algoritmo k-means

La distribución observada en el Gráfico 23 (apartado IV. 1.2.1., p. 273) se repite en la Ilustración 38: concentración de patrones del cluster0 en la tesis doctoral y el artículo en revista científica y distribución homogénea y similar en el cluster1 o cluster2.

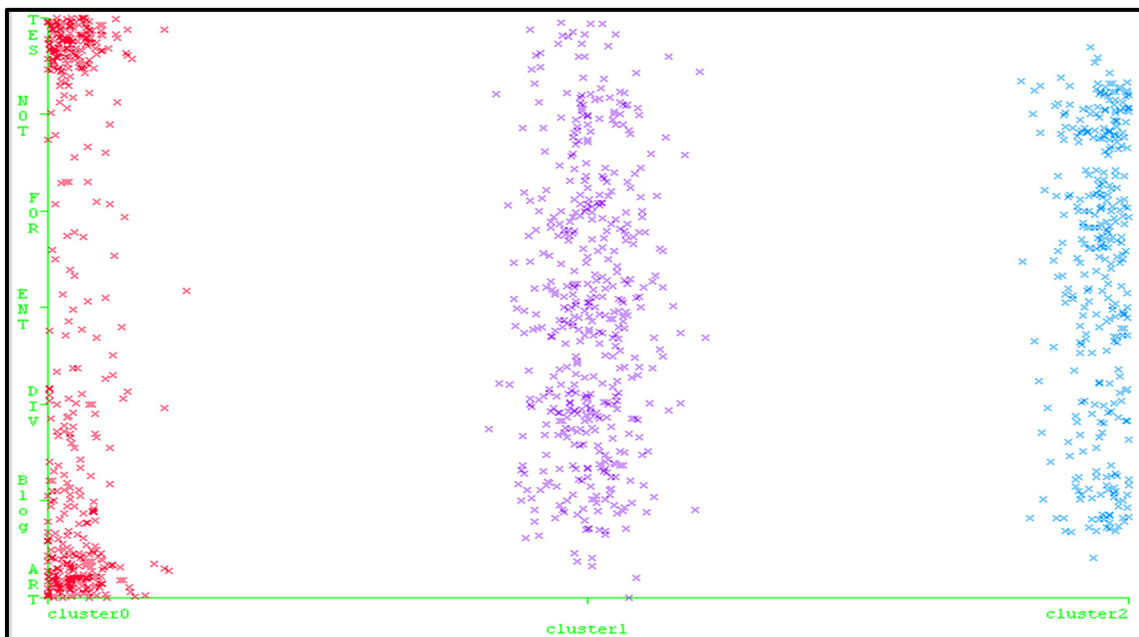


Ilustración 38: Agrupamiento de patrones por tipo de texto según algoritmo k-means

En los casos de IDT y sus tipos (Ilustración 39 a Ilustración 42), la concentración de patrones se identifica en las bases, es decir, en el eje horizontal, más próximo a 0. No obstante, en los tres casos se observa que la distribución de patrones para los valores más elevados del eje vertical desciende desde el cluster0 (es decir, el *cluster* que corresponde

con el TE) al cluster2 (el que corresponde con el TNE), verificando la relación directa defendida en el apartado IV. 1.4.9. (p. 323).

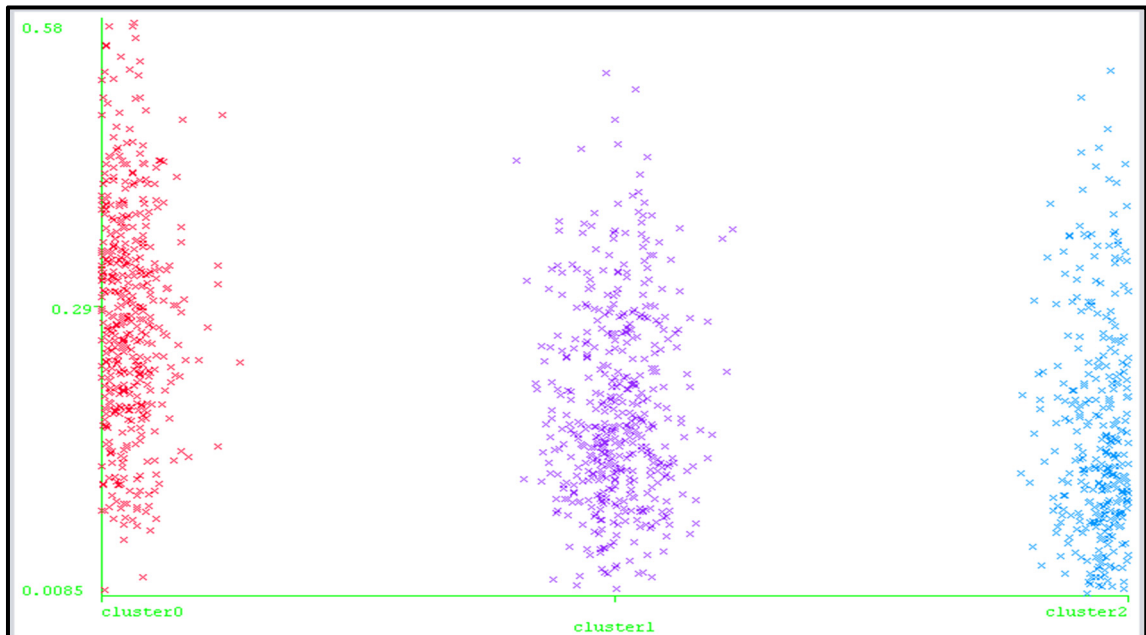


Ilustración 39: Agrupamiento de patrones por IDT global según algoritmo k-means

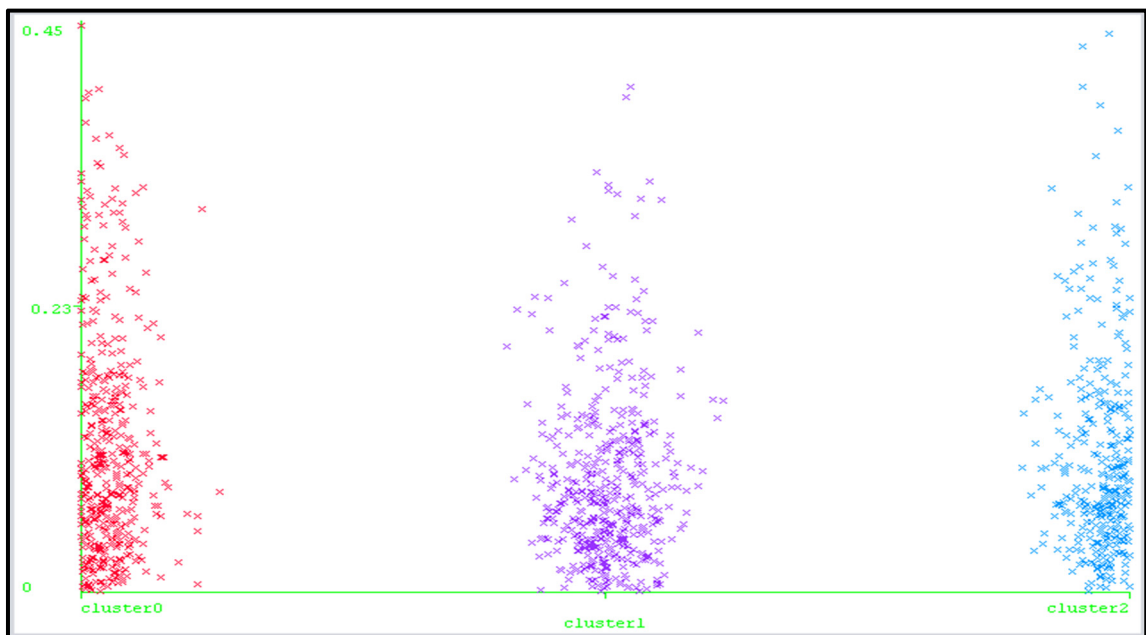


Ilustración 40: Agrupamiento de patrones por IDT a) según algoritmo k-means

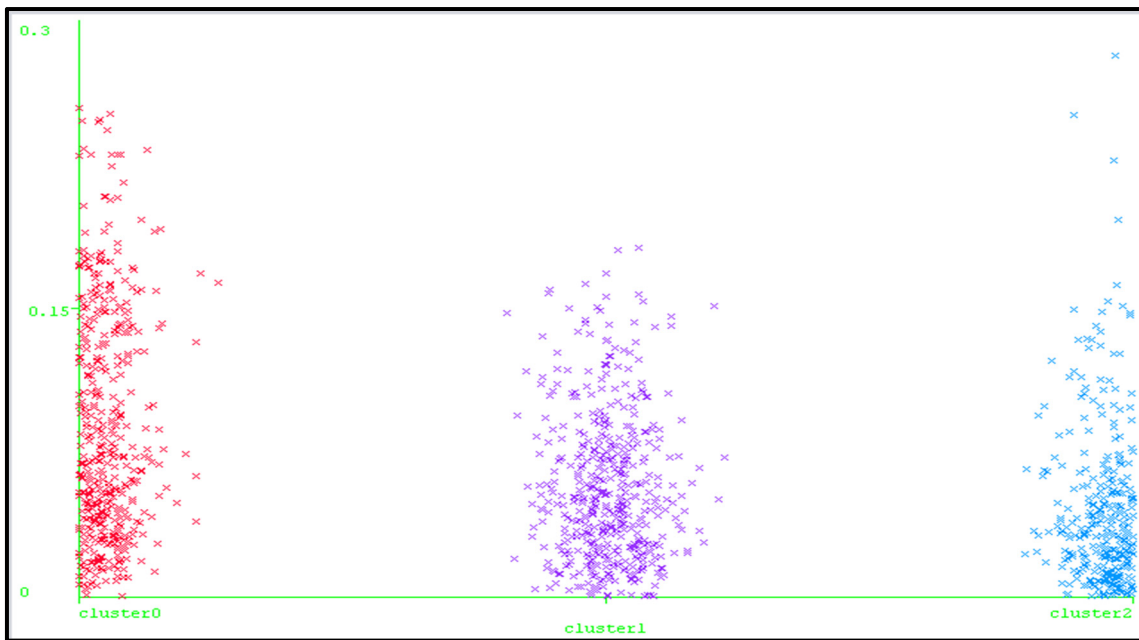


Ilustración 41: Agrupamiento de patrones por IDT b) según algoritmo k-means

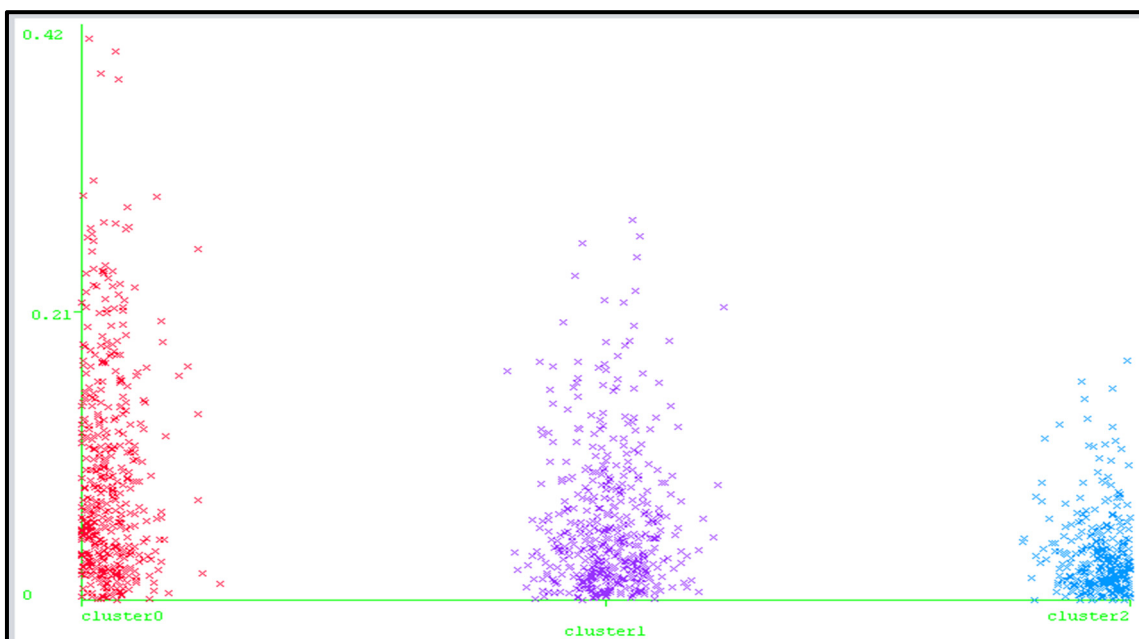


Ilustración 42: Agrupamiento de patrones por IDT c) según algoritmo k-means

Al igual que con el IDT, el IR y sus tipos muestran la misma distribución de relación directa (mucho más notable en la Ilustración 44 y la Ilustración 46), si bien la concentración en los valores inferiores a 0,5 es muy superior, lo que viene justificado por el procedimiento de normalización al que se someten los valores de IR.

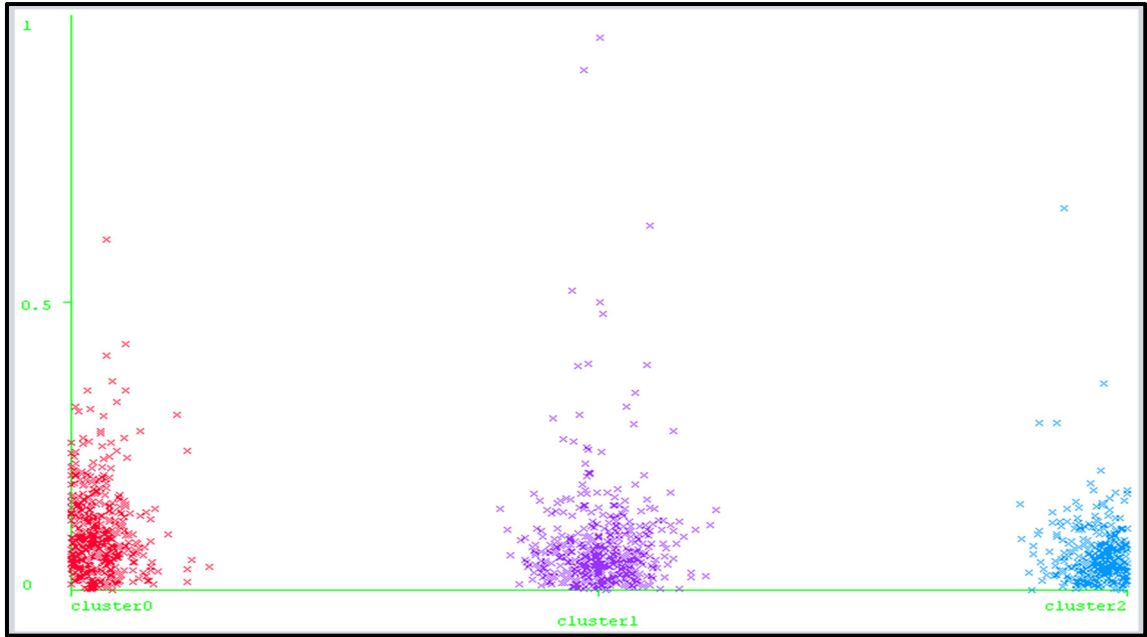


Ilustración 43: Agrupamiento de patrones por IR global según algoritmo k-means

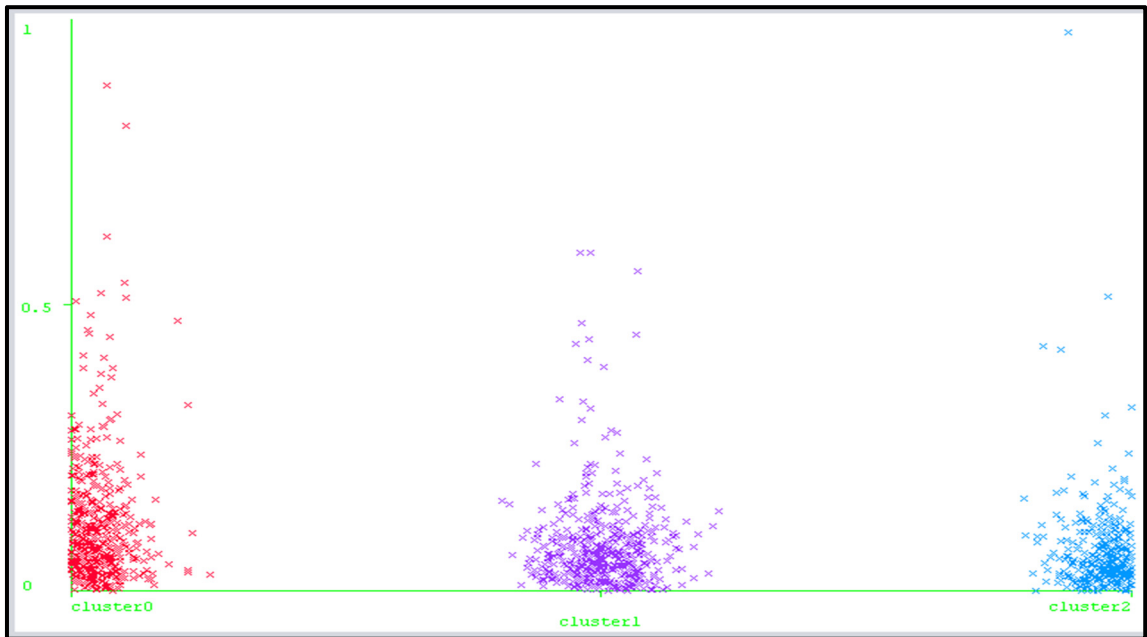


Ilustración 44: Agrupamiento de patrones por IR a) según algoritmo k-means

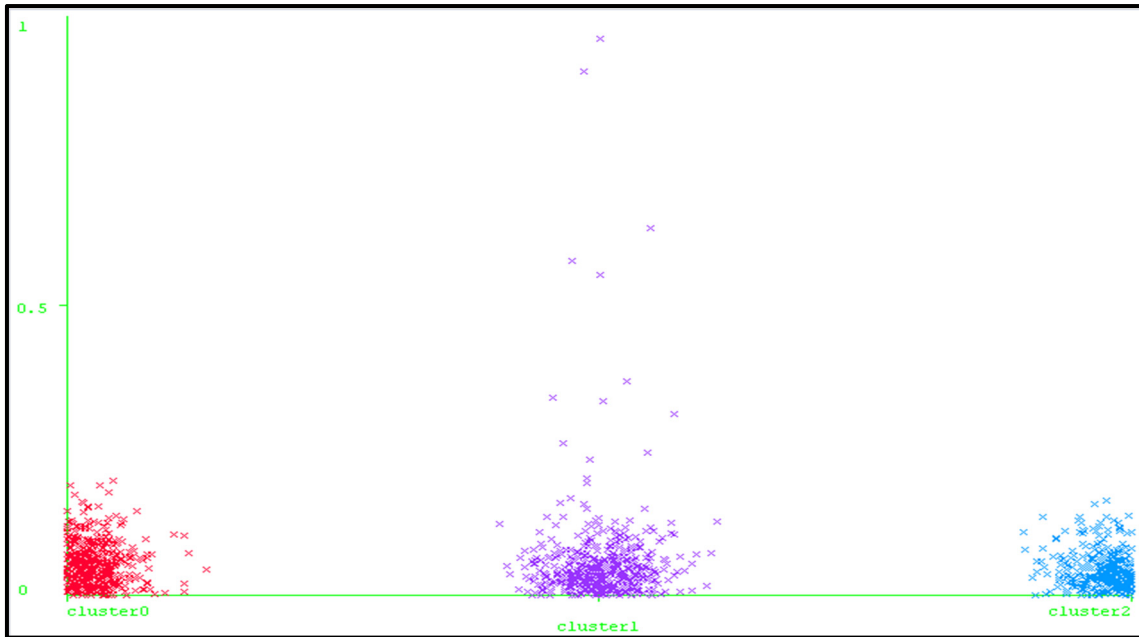


Ilustración 45: Agrupamiento de patrones por IR b) según algoritmo k-means

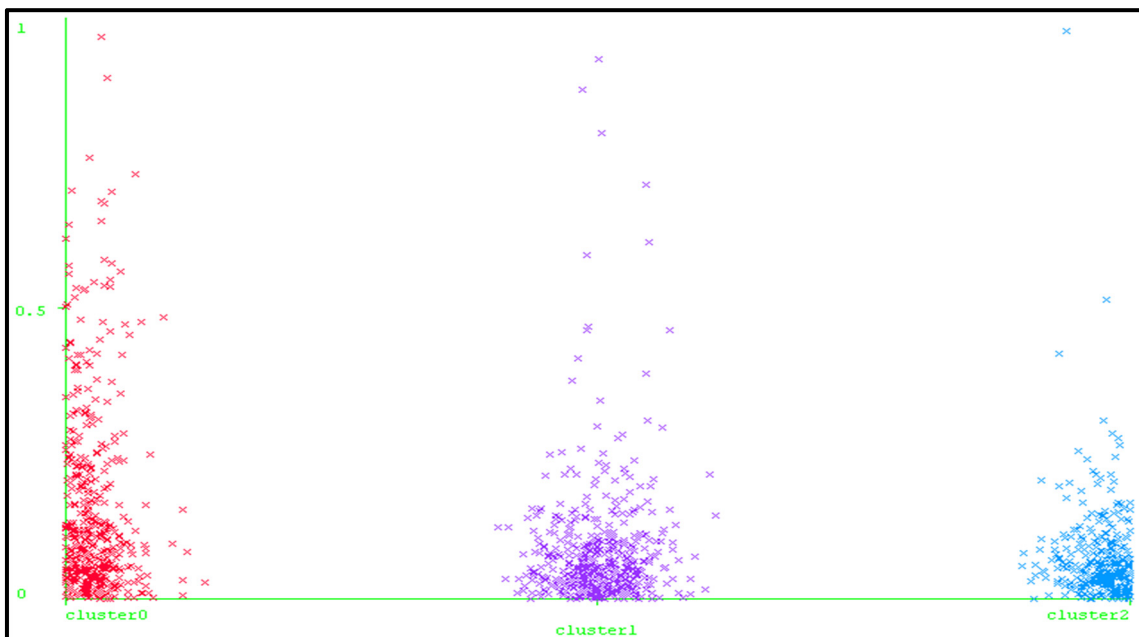


Ilustración 46: Agrupamiento de patrones por IR c) según algoritmo k-means

Los casos de emisor y destinatario (Ilustración 47 e Ilustración 48, respectivamente) presentan una clara agrupación según cada *cluster*.

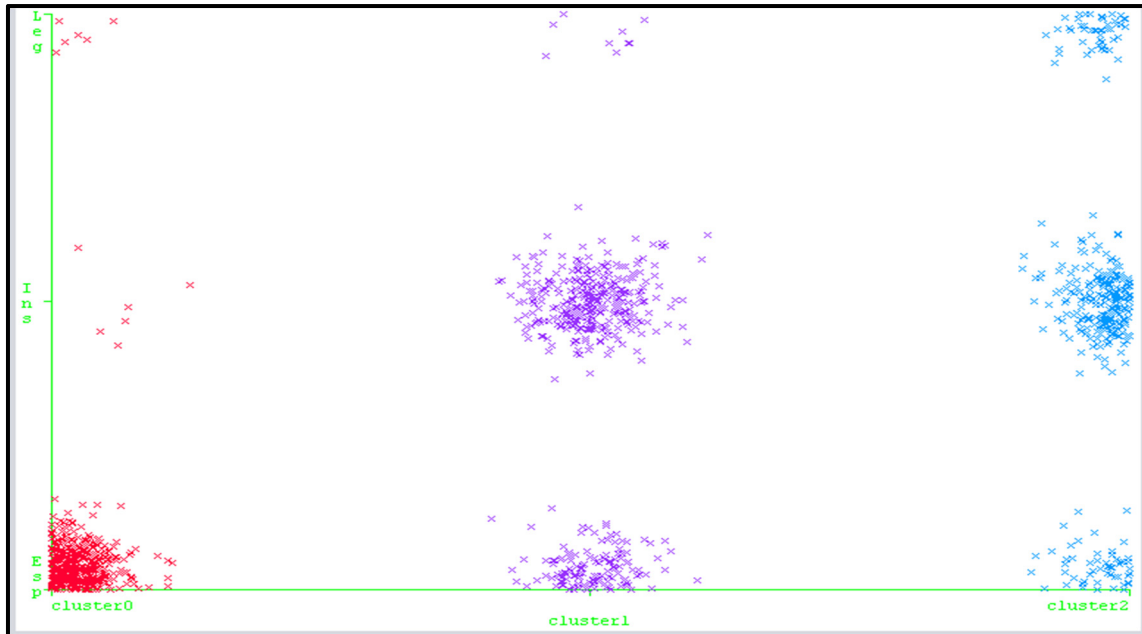


Ilustración 47: Agrupamiento de patrones por emisor según algoritmo k-means

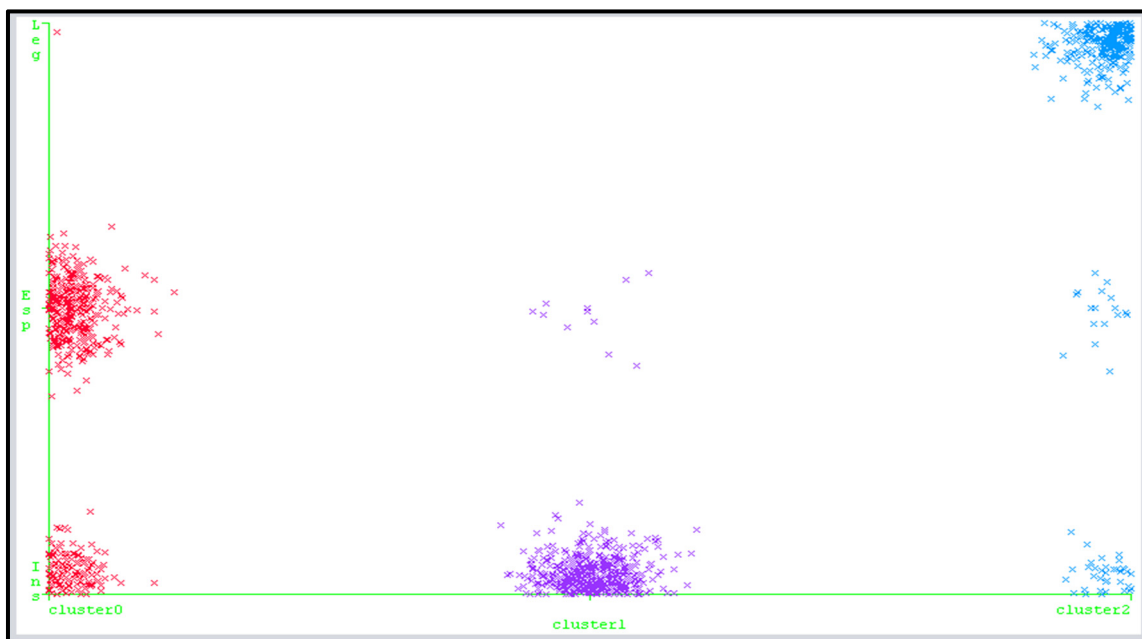


Ilustración 48: Agrupamiento de patrones por destinatario según algoritmo k-means

Estos datos coinciden con la probabilidad del valor de emisor o destinatario cuando se conoce la clase, como se sintetiza a continuación:

		1) Probabilidad del valor del atributo conociendo el valor de la clase		
		TE	TSE	TNE
Emisor	Valor del atributo			
	Esp	95,15 %	36,62 %	13,72 %
	Ins	4,85 %	61,97 %	71,57 %
	Leg	0,00 %	1,41 %	14,71 %
Destinatario	Valor del atributo			
	Esp	70,21 %	6,34 %	5,24 %
	Ins	29,79 %	80,99 %	37,41 %
	Leg	0,00 %	12,68 %	57,36 %

Tabla 167: Probabilidades de clasificación de emisor y destinatario según clase según datos globales

Al igual que en el caso anterior, los casos de subemisor y subdestinatario (Ilustración 49 e Ilustración 50) se concentran en tres puntos, si bien la dispersión es mucho mayor, como también lo es la probabilidad para cada valor (*vid. infra* Tabla 168).

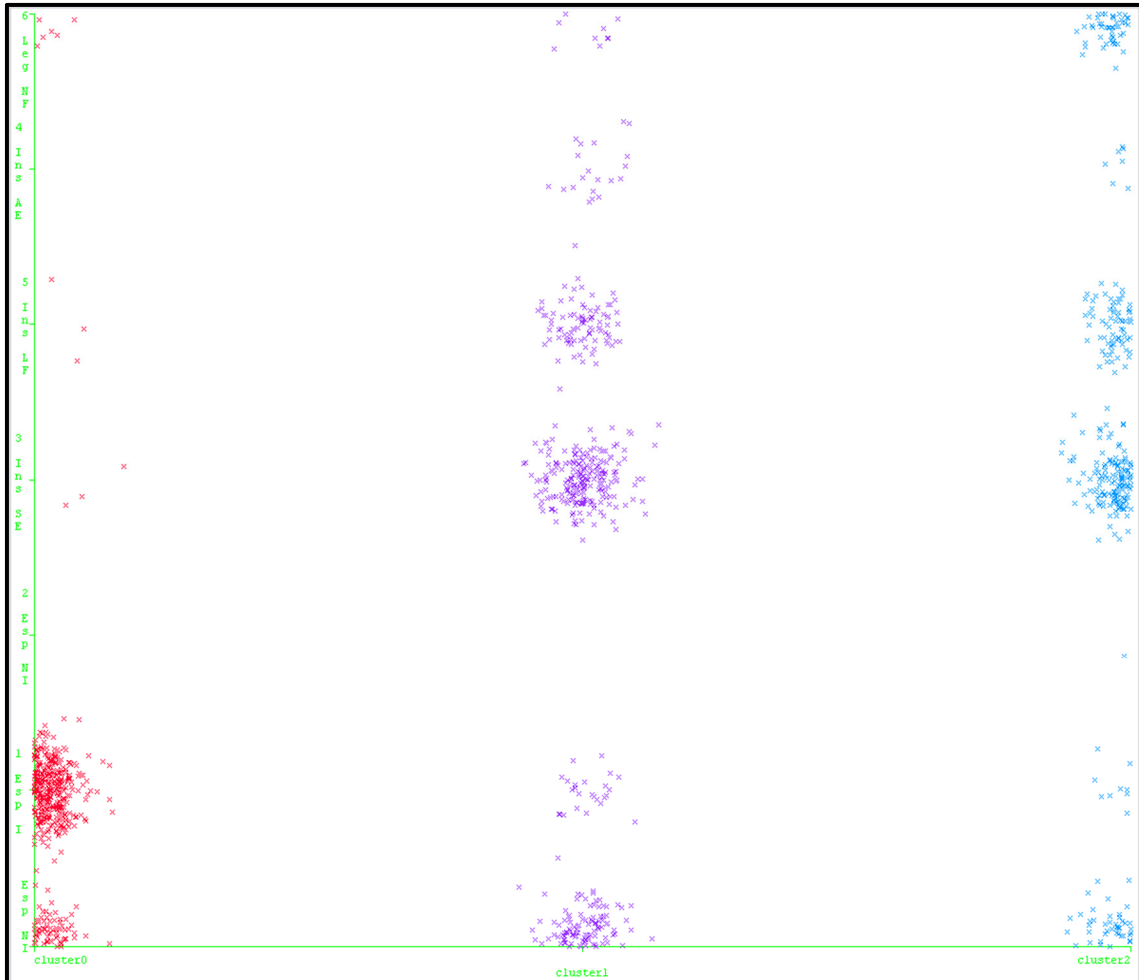


Ilustración 49: Agrupamiento de patrones por subemisor según algoritmo k-means

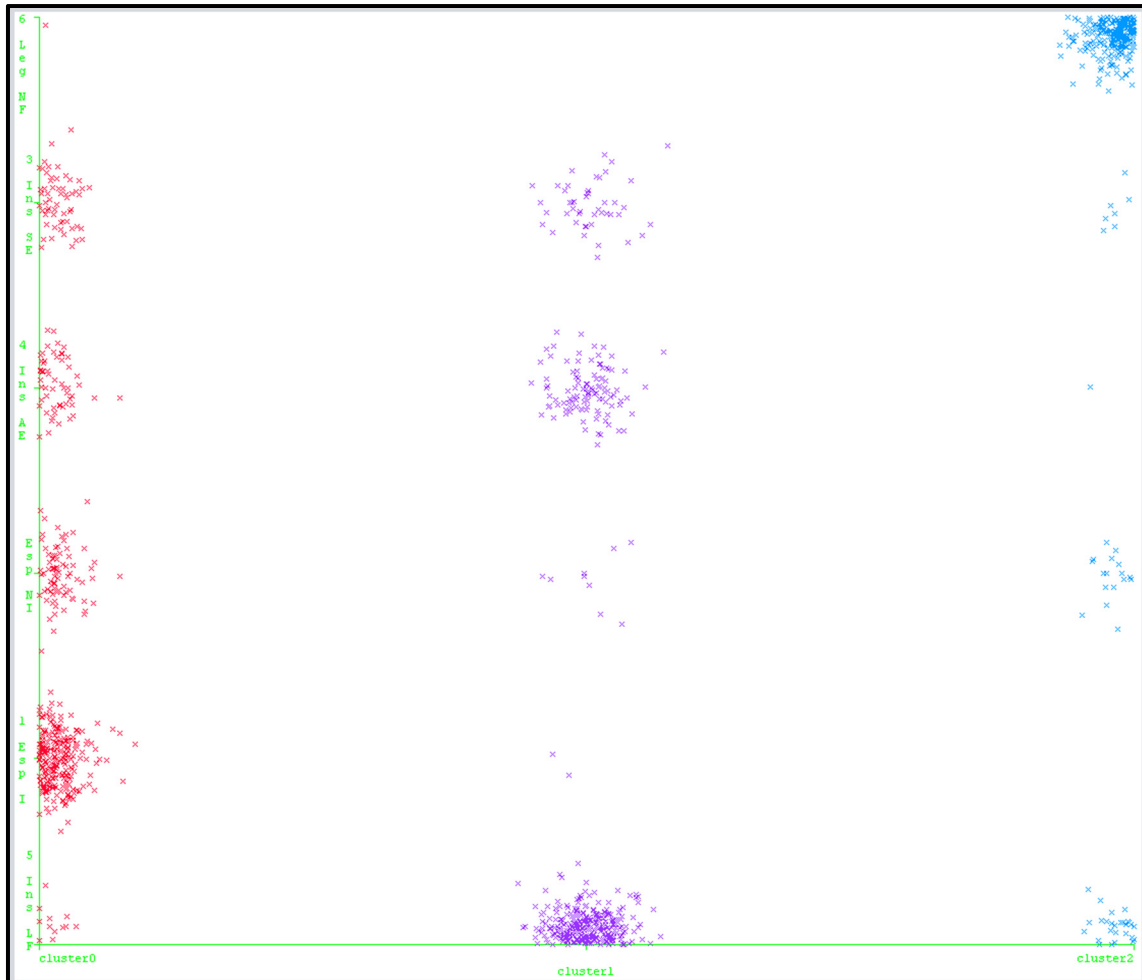


Ilustración 50: Agrupamiento de patrones por subdestinatario según algoritmo k-means

		1) Probabilidad del valor del atributo conociendo el valor de la clase		
		TE	TSE	TNE
Subemisor	Valor del atributo			
	Esp I	79,21 %	11,03 %	3,49 %
	Esp NI	15,94 %	25,59 %	10,22 %
	Ins SE	4,39 %	46,48 %	40,40 %
	Ins AE	0,46 %	3,99 %	2,00 %
	Ins LF	0,00 %	11,50 %	29,18 %
Leg NF	0,00 %	1,41 %	14,71 %	
Subdestinatario	Esp I	54,04 %	1,41 %	1,00 %
	Esp NI	16,17 %	4,93 %	4,24 %
	Ins SE	13,39 %	10,56 %	2,00 %
	Ins AE	13,16 %	22,54 %	2,49 %
	Ins LF	3,23 %	47,89 %	32,92 %
	Leg NF	0,00 %	12,68 %	57,36 %

Tabla 168: Probabilidades de clasificación de subemisor y subdestinatario según clase según datos globales

V. SUMMARIZATION AND CONCLUSIONS

1. SUMMARIZATION

This doctoral thesis is based on the problem of text classification according to the degree of specialization. The main problem lies in the identification of the intermediate degree and its differentiation regarding the opposite poles, the specialized text and the non-specialized text.

The notion of continuum constitutes the backbone of the thesis, since it considers that the texts can be actualized in different degrees of specialization. The aforementioned intermediate degree is the least studied degree in the scientific literature.

In our opinion, the characterization of the three stages of text specialization degree contributes to the predictive capacity of the discursive actualization of the different types of texts according to the degree.

The hypotheses outlined are related to: (i) the possibility of finding linguistic evidences that conform to each of the specialization degrees; (ii) the relationship of similarity between the semi-specialized text and the non-specialized text; (iii) the relevance of the topic or the type of text in the discrimination of the specialization degree; (iv) the perceptive similarity of the users when it comes to distinguishing the degree; and (v) the relevance of the speaker-recipient relationship and the terminology as distinguishing factors of the specialization degree.

As a general framework, this thesis uses applied linguistics as a dimension of linguistics responsible for contributing to solving the communicative problems of the human being (II. 1., p. 25). From our point of view, the discrimination of text specialization degree not only constitutes a theoretical problem, but its resolution also contributes to methodological decisions in corpus linguistics, terminography or translation, to mention only some fields.

The discursive actualization of the text specialization degree is understood within the framework of a communicative process, characterized mainly by the informative intentions of the speaker and the informative expectations of the recipient. The differences in the various cognitive-communicative motivations of the speakers make us propose the existence of three profiles of speaker and recipient (specialist, learned and lay users), which are specified in six sub-profiles (the non-researcher specialist, the researcher specialist, the learned semi-specialist user, the learned specialist-apprentice user, the learned lay user and the unlearned lay user) (II. 2.4., p. 46). These profiles are understood in a continuum framework and have prototypical and peripheral elements dependent on the perception of

the informants. In relation to the predictive capacity that knowledge of the characteristics of an object can offer, we have shown interest in the description of notions related to the text such as the text genre, the text typology or the type of texts (II. 3.3., p. 58), and especially on the theoretical problems and classification proposals that entails the construction of text typologies (II. 3.4., p. 59). This debate is subsequently applied to the compilation of the research corpus (III. 3.3., p. 184).

Going into detail about the linguistic field dedicated to the study of specialized discourse, terminology, we have carried out a brief review of the main theoretical approaches to terminology (II. 4., p. 63), as well as justified the adoption of the Communicative Theory of Terminology as a fundamental theoretical basis to understand, interpret and justify our discursivist position on specialized discourse and its degree (II. 5., p. 71).

Attending one of the characteristic elements of specialized discourse, both from a lexicalist and discursivist approach, we have studied the term (or specialized lexical unit) using traditional models and recent research that attempts to oppose term and non-term (that is, non-specialized lexical unit) (II. 5.3., p. 77).

The scarcity of clear conclusions about the definition of the term has led us to delve deeper into the notion of specialized value and situated sense (II. 6., p. 90) for which we have resorted to the characterization of specialized knowledge as opposed to common knowledge (II. 8., p. 126).

In this framework, from a theoretical point of view, and based on the principle of variation, we have analyzed some of the relationships that are established between the denomination of a term, the concept to which it refers and the discourse in which it is inserted. These relationships include terminologization, vulgarization or denominative or conceptual variation (II. 7., p. 104). This consideration, in a global way, has led us to propose a set of criteria to distinguish whether a lexical unit activates or not a specialized value (II. 8.1., p. 126) and what happens when the communicative situation is not prototypical in this activation (III. 8.2., p. 128). This block continues with a review around the features of the specialized text and the description of the models on the continuum of the specialized text (II. 9., p. 136).

The block of theoretical foundations concludes with the section dedicated to the reformulation and the proposal of a model for the identification of markers and reformulation means, which is motivated by the role played by the reformulation in

specialized discourse, in which reformulation serves as a mean of access to specialized knowledge of terminology not understandable by the recipient (II. 10., p. 151).

The difficulties of the theoretical description of our object of study are reiterated in the justification of the methodological foundations. The main methodological framework of the thesis corresponds to the conjunction of empirical work with corpus and the sociocognitive theory of perception.

From the point of view of the corpus compilation, we adopt the perspective that considers corpus linguistics as an instrument of inductive work (III. 1., p. 165). From the perspective of the theory of perception, we adopt the relevance that collectively-accepted subjective traits are not the result of chance. We use the foundations around analytical perception and synthetic perception (III. 2., p. 168).

The materials used in this thesis are composed of a corpus that has 315 texts and half a million words. Its qualitative representativeness to study the text specialization degree lies in that it is based on three different subjects (socially recognized, such as law, computer science and medicine). Subsequently, we have selected seven possible types of texts in the three subjects (article in scientific journal, blog post, scientific-divulgence magazine or web, news, interview, participation in a forum and doctoral thesis). From this combination of type of text and subject, 15 texts have been collected (III. 3., p. 181). From the classificatory point of view, our corpus is a written, specialized, generic, contemporary, synchronic, restricted, closed, equivalent, single-issue, monolingual, reference and simple corpus (III. 3.6., p. 190).

Our study adopts a mixed quantitative approach with eminently correlational but also descriptive scope. A transversal non-experimental design of a descriptive and correlational type is followed (III. 4., p. 197).

In this corpus, several variables are studied, both quantitative and qualitative. Among the qualitative variables we find the nominal variables (the type of text and the subject of the text), which are control variables that regulate the results of the study. Among the qualitative ordinal variables, we find the speaker, the sub-speaker, the recipient, the sub-recipient and the class of text (this latter variable is the dependent variable that is investigated).

Among the quantitative variables we find the number of lexical forms, the standardized ratio between types of lexical forms and the total of lexical forms, the terminological density index (and its three types), the number of reformulation means and the reformulation index (and its three types) (III. 5.1. to III. 5.5., pp. 201 to 216).

Some of these variables are provided by the researcher (III. 6.8., p. 226), while another part is provided by the groups of informants (III. 6.7., p. 222.). These latter variables are the types of terms (understood as specialized lexical forms), divided into delphic, semi-cryptic and cryptic terms (following the criterion of perceptive occurrence-prominence), the profile of the speaker and the recipient and the class of text as a result of a synthetic analysis.

The analysis of the four groups of informants (student, average user, translator and researcher) yields a database with 1260 patterns (or texts) that is analyzed through the manual correlation of two or more variables and through machine learning. The analysis using machine learning techniques employs Bayesian network algorithm (which calculates the classification probabilities), logistic regression (which determines the most relevant attributes for classification), decision tree J48 (which allows to design a rules tree) to classify texts) and the unsupervised k-means clustering algorithm (which allows to recognize the most representative characteristics of our three classes according to the specialization degree) (III. 7., p. 228).

2. CONCLUSIONS

The results analyzed, both globally and contrastively (manually and automatically) shed light on the hypotheses proposed in section I. 3. (p. 18). Thus, in this section we proceed to provide the most relevant data or interpretations that contribute to verify or refute each of the hypotheses. It should be noted that these data must necessarily be linked to the type of corpus and the set of types of texts that form it, although we understand that its property of representativeness would allow generalizations of interpretations to other possible texts.

a) **Hypothesis 1:** It is possible to use the variables related to the perception of the lexicon and the cognitive-communicative profile of the speakers and recipients of the texts to perceive the text object from the consideration of an object marked by the specialization degree in the next text categories: non-specialized text, semi-specialized text and specialized text, that is, marked by the continuum.

The variables used have allowed to link, generally, certain prototypical values to each of the classes according to the specialization degree (not only from the point of view of machine learning through the characterization of the centroids, but also by means of the global manual analysis). For example, the intermediate values of terminological density, reformulative density and standardized TTR, as well as the speaker or recipient, are closely linked to the semi-specialized text. The same applies to the specialized text and the non-specialized text, which are linked to higher and lower values, respectively.

In relation to the notion of continuum, we proposed the idea that the semi-specialized text is constituted as that textual object that shares characteristics with both the specialized text and the non-specialized text, although it can be identified as an individual textual object, as we propose in section II. 9.4. (see below, Ilustración 13, p. 147):

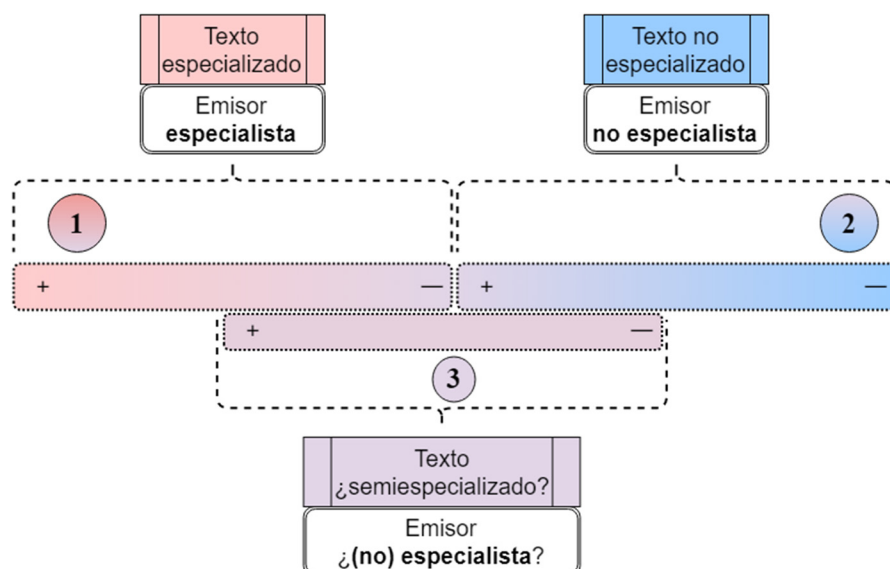


Ilustración 13: Gradación en y entre polos opuestos según nuestro modelo

The block number 3 is confirmed in the sense that the semi-specialized text is constituted as an object of transition between the opposite poles, mainly if the terminological density index and the reformulation index are taken into account, even the standardized TTR value, since in the three cases the semi-specialized text has the intermediate values. In the case of the type of speaker and of the recipient, it is closely related to the learned user (which is, in turn, the intermediate value). This illustration should have a modification, since we have verified that the recipient is the starting attribute in the classification of the specialization degree, which is why it regulates the discursive act and its actualization (and, therefore, the subsequent classification in the three classes):

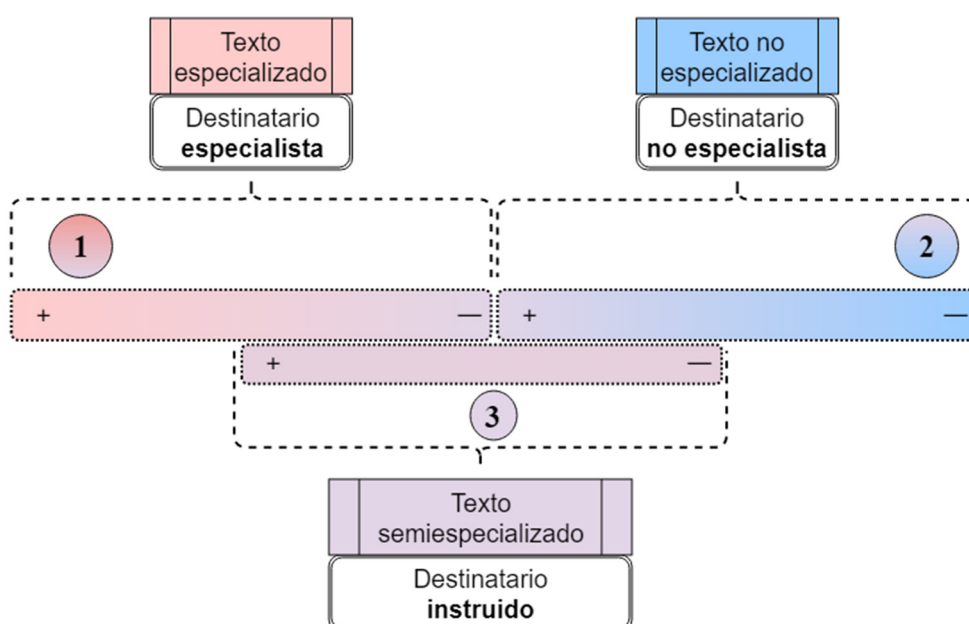


Ilustración 51: Gradación en y entre polos opuestos según nuestro modelo [elaboración propia]

This illustration shows only a prototypical image of the most representative elements of each class, since, as it has been proved in the corpus, depending on each type of text, a specialist is capable of designing specialized, semi-specialized and non-specialized texts, independently of the conceptual domain that is exercised and that reflects the text (as it is defended in section II. 9.4., p. 144).

b) **Hypothesis 2:** There is sufficient evidence in the configuration of a text to classify it as semi-specialized text in the same way that they exist (and have been investigated) for the specialized text.

c) **Hypothesis 3:** The semi-specialized text is closer to the formal features of the non-specialized text. It only shares semantic criteria with the specialized text.

d) **Hypothesis 4:** The types of texts that the semi-specialized text adopts are identical to those that the non-specialized and the specialized can adopt.

The problem of finding the signs of classification lies in the limits of classification between specialized text and semi-specialized text, and between semi-specialized text and non-specialized text. Firstly, the criteria for discriminating the degree of these three classes must be determined: in our case, the relation of the recipient with the speaker and the terminological density, which are between the specialized and the non-specialized text.

Indeed, the semi-specialized text shares with the non-specialized text the most frequent types of texts in which it is usually actualized (in order: scientific-divulgate magazine or web, blog, interview, news and participation in a forum). The specialized text establishes a clear relationship with the article in a scientific journal and with the doctoral thesis. Such is this relationship that they are practically restricted to this type of text (only 1% of the articles and another 1% of the doctoral theses are classified as semi-specialized). This is justified by the words of Loureda (2002: 159), when he defends that the textual types respond to the expressive needs of the interlocutors and that «se presenta como horizonte de expectativas para el receptor o destinatario» with the aim of guaranteeing the functionality of speaking (idem).

Secondly, the most representative types of texts do not correspond to the most frequent types in those classes (although in comparison with the other classes, they are

more frequent in the class of the centroid) but rather those types of texts that best represent (in conjunction with other elements) the class in question. For example, the scientific-divulagation magazine or web or the blog are the most frequent type of text in the semi-specialized text, but they also have presence in the non-specialized text and in the specialized text and, in fact, they share many characteristics with these classes, so they are not constituted as representative of the semi-specialized text. The same happens with the case of participation on a forum in the non-specialized text, which, despite being the most frequent, it is not the most representative, since it has characteristics that shares with the semi-specialized text. In these cases, it is shown how frequency and probability do not necessarily condition the characterization of the class (but they do determine the classification within a class, which is different). In other words, the frequency of the appearance of a value of an attribute over the total of the corpus does imply a certain probability and this implies a tendency towards a class, but this probability does not condition the characterization of the class (that is, the conjunction of representative elements that are attributed to the class).

Also from the formal point of view, the semi-specialized text shares with the non-specialized text the reduced number of specialized lexical forms (19% and 10% of the corpus, respectively), especially the delphic forms, which have a homogeneous presence in these two classes. The occurrence of terminology in the semi-specialized text and the non-specialized text is similar both quantitatively and in the different subjects although there are some minor differences. The formal difference (which actually becomes a conceptual criterion) is found in the frequency of semi-cryptic and cryptic forms, which are much greater than in the non-specialized text. This increase in frequency (not comparable, of course, to the specialized text) is what justifies the conceptual approximation of the semi-specialized text to the specialized text. And, although the semi-specialized text shares the types of text with the non-specialized text, some texts have recipients with different informational expectations, which justifies the use of semi-cryptic and cryptic terms. This fact also entails an increase in reformulation in the semi-specialized text, which, although it approximates the average value of the corpus, is in fact similar to that presented by the non-specialized text. These are the elements that we proposed to identify as intervening «en la identificación del texto semiespecializado y que resultan de posible aplicación tanto en la configuración del texto especializado como en la del divulgativo» of the objectives section (I. 4., p. 19).

The similarities between non-specialized text and semi-specialized text are also observable graphically, especially in the decision trees and the colors for the class leaves: it is possible to confirm the polarization observed in Gráfico 23 (see above section IV. 1.2.1.), as specialized text is restricted mainly to the specialist-specialist discursive relationship (see above Ilustración 28, p. 430) and some other more reduced or scattered cases (Ilustración 28, among others) and that the semi-specialized text and the non-specialized text usually share spaces, as it can be seen in Ilustración 30, Ilustración 31 or Ilustración 32.

There are also some elements that deserve attention, especially if they are contrasted with other studies:

- a) As defended in Rodríguez-Tapia (2016b), the semi-specialized text presents an intermediate IDT between specialized and non-specialized texts, which favors the continuum hypothesis.
- b) Ciapuscio (2003: 90), on the other hand, defends that at the medium specialization degree «escasea la terminología y se perciben formas parafrásticas, definiciones, ejemplos, etc., y en el nivel más divulgativo se evita la terminología o se define y explica», which cannot be demonstrated with the data that is handled in this investigation, or, at least, not in a clear way. At least initially, in the semi-specialized text there is no scarcity of terminology, nor in the non-specialized text is it avoided; rather, there are qualitative differences, above all, when compared with the specialized text. The paraphrased reformulation means are similar in the semi-specialized text and the non-specialized text, although expansion means almost doubled in the latter one. This does not mean that the non-specialized text is reformulated more, since it has the lowest reformulation indexes of the corpus.
- c) Neither it is confirmed the fact that «en los textos más especializados la variación denominativa es un fenómeno menos frecuente que en los textos no especializados, tanto porque existe un número menor de nociones representadas por más de una denominación, como porque los grupos polidenominativos presentan un número inferior de denominaciones diferentes» (Freixa, 2002b: 265). As it has been proved, the means of reformulation by variation are much more numerous in the specialized texts, including the index of reformulation by variation, which is justified by the use of alternative denominative forms to denominate concepts, including the acronyms. In this sense, this conclusion

contrasts with the traditional idea that the reformulation is linked to the non-specialized text and the absence of alternative forms and explanations and descriptions with the specialized text. According to the data obtained, the reformulation is associated to the terminology and, the more terminology exists for each lexical form, the greater tendency to reformulate exists in the text. This phenomenon normally takes place in the specialized text.

- d) On the contrary, the words of Ciapuscio and Kuguel (2002: 43) and Cabré and Estopà (2005: 36) are confirmed, which defend, respectively (i) that specialized texts select «clases textuales específicas» (*types of text* according to the metalanguage that we use) of specialized discourse (as the research article in a journal) or (ii) that «desde el punto de vista formal, los textos especializados no se distinguen de los generales más que por el hecho de seleccionar géneros textuales específicos» (*types of texts*).

In relation to the characterization of the semi-specialized text, Rodríguez-Tapia (2016b: 246) established the following definition:

Hasta el momento, los corpus que hemos estudiado nos permiten sostener que el texto semiespecializado se define como aquel texto redactado por un especialista y dirigido indistintamente a un público lego o especialista, en el que predominan, de forma independiente o en conjunto, las funciones representativa y comunicativa, cuyo índice de densidad terminológica no es alto (0,12 de media) y que se asocia a la superestructura de los glosarios, artículos divulgativos y artículos científicos.

Considering the lexical and discursive characteristics stated in the thesis, we consider that a more optimal definition would be the following: the semi-specialized text is defined as that discursive act prototypically issued by a speakers with cognitive-communicative profile of a learned user and addressed to another learned user; it has a terminological density lower than the specialized text but higher than the non-specialized text, it is associated with types of texts such as blogs, interviews or scientific-divulcation magazines or websites, in which terms that are frequent in common discourse stand out, but also has cryptic and semi-cryptic forms that are subject to reformulation, mainly by expansion or variation.

The main difficulty in defining the semi-specialized text is that it must be defined as opposed to a maximum or minimum value. It could be affirmed, keeping in mind the exceptions that may exist, that the semi-specialized text is a specialized text that takes the structure and formal features of a non-specialized text.

e) **Hypothesis 5:** The terminological-conceptual configuration of the different disciplines directly affects the classification of the texts according to their specialization degree based on the recipient's perception.

From our point of view, it is not possible to verify this hypothesis or, at least, not completely. In the first place, the three subjects have a homogenous representation in the three classes according to the specialization, so it cannot be defended that the conceptual organization according to the discipline guides the texts towards one class or another. Some differences in the type of speaker do have been identified (for example, the researcher specialist in medicine, the non-researcher specialist in law and the learned lay user in computer science). In fact, when analyzing the classes according to the specialization degree, there are more similarities than differences, such as those observed when combining the speaker and the classes or the recipient and the classes. Computer science is the subject that presents the highest divergence, since it tends in all the cases towards the learned speaker or towards the learned recipient.

Alonso and DeCesaris (2006: 270) defend the fact that the differentiation of the degree of specialization does not have to be the same in all the subject fields. Given the homogeneity in the classification of the subjects in the three classes by the four groups of informants, it can be argued that the differences are not found in the distinction of the classes, but in other elements. In fact, these elements are those that we mentioned in section III. 3.2.: «sí que pueden existir diferencias entre disciplinas a la hora de graduar la especialización: en el número de términos que emplean en sus textos, en los procedimientos de reformulación que emplean, los tipos textuales que seleccionan, etc.» (p. 183).

In the case of the vocabulary, differences are observed in the number of terms in each subject (in order, from highest to lowest: medicine, law and computer science). And not only there are quantitative differences, but they also differ in the distribution of the three types, because law stands out in the delphic terms as opposed to the cryptic terms, while computer science and medicine, despite highlighting the delphic terms, have less presence of semi-cryptic terms. These differences are also observed in the reformulation: law is the subject that is least reformulated and computer science and medicine have similar values. The absence of reformulation for law can be linked to the fact that law stands out, above all, in the delphic terms, that are those lexical forms recognizable as very frequent in common discourse, so that understanding is more accessible to the recipient.

However, the linguistic differences of the three subjects do not imply a different classificatory distribution in the three classes: it simply implies that each subject, by representing a different field of knowledge, resorts to different resources (in this case, the delphic terms, semi-cryptic terms, the reformulation, etc.).

f) **Hypothesis 6:** There are no perceptual differences between different social groups when classifying the texts according to the specialization degree, although they may present divergences in the criteria they apply to distinguish types of terms.

The data shows that there are more similarities than divergences in the criterion applied by the informants to assign values to each of the variables. That is, the percept of informants has more common characteristics, so it is possible to notice a certain degree of convention or collective acceptance when analyzing the text specialization degree.

This has to do with the cognitive dimension of each individual (Caravedo, 2014: 36, 37, 46), in which the measurement of probabilities of the phenomena (that is, how each of the variables are actualized in the discourse) is part of the speaker's own communicative (and cognitive) competence, by employing the criterion of occurrence-prominence over the phenomena he or she analyzes. In this case, the informants coincide, to a greater or lesser extent, in analyzing the elements mentioned by Cabré (2002a: 89):

[...] el emisor y el receptor (tipo de emisor y receptor, intenciones y nivel de conocimiento que ambos poseen sobre el tema), la situación (el medio en que se produce la comunicación y el sistema de transferencia utilizado), el propósito y las expectativas de emisor y receptor con relación a su interacción. Cada uno de estos factores posee un determinado valor en cada acto comunicativo y su conjunto explica la configuración de un tipo de texto que pretende ser adecuado a las circunstancias en las que se produce.

The group of students is the one that presents more evident divergences, sometimes together with the average user. This pair usually contrasts with the similar observable criteria between the researcher and the group of translators.

Among the common criteria, it is observed that the informants do not use the terminological density (nor particularly the terminological density of cryptic terms) to classify the texts according to the degree of specialization, but that the determined class is the result of a synthetic analysis that adds other elements to discriminate the classes.

Among the differences, classification tendencies are noticed in the case of sub-recipients (the group of students tends towards the researcher specialist and the rest of the informants opt for the semi-specialist learned user instead of the non-researcher specialist).

These tendencies may be related to a logical decision on the part of the four groups of informants: it is easier to have three assignable values, since in them a maximum, intermediate and minimum value can be identified. Thus, in our opinion, the frequency of three types of sub-recipients (researcher specialist, semi-specialist learned user and learned lay user) is linked to the supercategories of specialist, learned user and lay user. Even in the case of the learned lay user, which in our classification is considered a learned user, is relevant in the perceptive decisions of the informants, since it is coherent that the recipient of a text has a minimum of knowledge on a subject to be able to understand it. This presupposition of knowledge on the part of the speaker or the difficulty of finding a clear limit between subcategories could justify the absence of the specialist-apprentice subcategory.

In relation to specialized lexical forms, especially in relation to delphic forms (and particularly when compared with the group of students), it is necessary to consider the reference factor that each group takes to regulate their perception. From our point of view, the average user assimilates the specialized value of a lexical unit to its semantic transparency through morphology: the opaquer a lexical unit is, the greater the probability that the average user classifies that unit as specialized. This fact contrasts with the criterion applied by the student, which more frequently identifies lexical forms of common discourse that activate a specialized value, in our opinion, due to his recent instruction in terminology and approach to the notion of specialized value as situated sense. Thus, it can be argued that the prototypical discourses that informants take as a reference to measure the frequency of a lexical unit in the common discourse are different in the student and in the average user. Thus, it can be determined that the factor that each group of informants takes to regulate their perception may be different when carrying out the analytical analysis for the lexical forms, although it does not necessarily have to be when determining the class through synthetic analysis, since, as it has been observed, it is usually similar among informants.

In short, it is possible to defend that users share a perceptual criterion based on linguistic tests that allows them to discriminate the text specialization degree.

g) **Hypothesis 7:** The discourse relation between speaker and recipient is the factor that regulates the rest of the variables, and, therefore, the discursive configuration of the text. Thus, this relation can condition the specialization degree by combining it with other variables.

Adelstein (2007a: 276 and 277) points out the communicative situation, especially the interlocutors, as the main factor that determines the generation of specialized meaning, but she does not position herself in favor of the speaker or the recipient. Domènech (2006: 50), on the contrary, argues that the profile of the speaker is not the only pragmatic characteristic of specialized texts, but it is the most decisive one. However, the data handled here, especially those obtained using the decision tree algorithm J48, indicate that it is the recipient, in the first instance, the most determining factor.

From our point of view, the recipient of the text constitutes the starting point of the discourse configuration. Thus, it can be affirmed that the specialization degree of a text depends on the features of its recipient (Domènech, 2006: 72), since, in short, it is the main axis of acceptability (Beaugrande and Dressler, 1997: 35-45). In specialized communication, the recipient expects relevant information related to specialized knowledge, although this information is adjusted to the recipient's knowledge. This last expectation is what is observed in the classificatory decision of the informants, when they opt for the learned recipient instead of the unlearned lay recipient: the recipient has a minimum knowledge of a topic to be able to understand it.

In fact, this is defended by Domènech (2006: 72) when she resorts to idealized models: «l'emissor adapta el contingut del text que produeix al domini cognitiu idealitzat especialitzat que atribueix al seu receptor» both from a quantitative (density of knowledge) and a qualitative point of view (structuring and relevance of knowledge). The groups of informants in this research consider that the recipients of the texts are subject to encyclopedic or shared knowledge according to social groups, which serves as a condition for the speaker to reconfigure the discourse (selection of the types of terms and their density, the types of reformulation means, their frequency and their relation to the terms, etc.).

The most important attribute according to the J48 algorithm is the recipient, from which the whole tree is structured. After this, the speaker attribute plays the second most relevant role. Thus, it can be argued that the discourse relation between the recipient (as a conditioner of the discourse configuration) and the speaker (as the actualizer of the cognitive-communicative competence in discourse) is the one that establishes the most

important bases to distinguish the three specialization degrees. In particular, the competence of the learned recipient plays an important role as a regulating factor in the reformulation, especially in cases where a term is followed by a description or an explanation.

In the case of the lay recipient, the IDT of cryptic terms and the global IDT are the attributes that discriminate the classes. In general, numerical attributes serve to discriminate between two classes attributable to the same rule and are usually found next to the leaf, that is, they are integrated into the tree as the final part of the decision.

h) Hypothesis 8: The most relevant factor to distinguish the text specialization degree is the terminological density of the lexical units that are perceived as opaque semantically or as rare in common discourse.

As mentioned in the sections corresponding to the theoretical foundations, it is widely recognized by the scientific community that the lexical level of specialized discourse constitutes the essential characteristic of the specialized text (Monterde Rey, 2002b: 115, Cabré *et al.*, 2007). The specialized value of a lexical unit plays an important role in specialized discourse, since when a sense integrated by a selection of precise semantic features and corresponding to the specialized meaning of a unit in a specific specialized field becomes active, other senses are inhibited. These senses are conformed by a selection of these same traits together with other features not activated.

The IDT of the cryptic forms and the IDT of the semi-cryptic forms constitute the most relevant attributes to discriminate the specialized text and the non-specialized text, according to the logistic regression algorithm. Not only this, but, as for the class, there is a direct relationship between the global IDT of a text and its specialization degree, which allows to defend that the terminological density contributes to discriminate the text specialization degree. Only the IDT of delphic terms in a text does not contribute to class discrimination. This direct relationship is maintained between the IDT and the type of speaker or sub-speaker according to the cognitive-communicative competence of the user, since, as a general rule, the global IDT decreases as the aforesaid competence decreases. And, even though the direct relationship is also established with the recipient, the differences are so subtle that, in this case, it cannot be defended that the IDT contributes to discriminate sub-recipients.

This direct relation between class and IDT contrasts with the problem solved in other empirical studies that do not demonstrate this fact. For example, Domènech (2006: 207, 212-213) considers that the terminological density seems to be an insufficient measure to obtain a valid quantitative parameter to discriminate a text according to its specialization degree. In fact, Domènech's doctoral thesis (2006: 208, 212) not only shows that the differences in terminological density between degrees are not significant, but rather precisely, inversely proportional to the specialization degree. In our opinion, this is due to the method used, which uses identification software for term candidates, that is, by frequencies of appearance in discourse, and does not use a sociocognitive study of perception like the one we carry out. The differences are, then, methodological.

Domènech (*idem*) does highlight the terminological diversity (in particular, it deals with the number of terminological occurrences of a text, the number of different terms, the frequency of occurrence of each term and the overall proportion within the text of the distribution of this frequency) as a statistically valid quantitative parameter to discriminate a text according to its specialization degree (*ibid.*: 216). In our case, it has only been inversely linked that the specialization degree of a text increases when the lexical diversity (which is not terminological) decreases.

Regarding the relevance of the type of terminology that is relevant to discriminate classes, our hypothesis is based on the fact of specialized lexical forms that are semantically opaque or rare in common discourse. In fact, the delphic terms, that is, the frequent forms in common discourse with a specialized value, are not very relevant in our study, since they usually appear in all classes and do not determine the classification in any way. As we advance in section II. 7.8., (p. 123), the delphic terms are a problem for the terminological work, either for the systematic terminologist or for the interlinguistic mediator, since sometimes, due to the ignorance of the thematic field (as Picht already advanced, 1984: 215), some terms are not identified as such.

If we look at the elements in which the IDT is not relevant, it can be argued that it does not allow to discriminate the subject of a text, although a reduced number of cryptic forms helps to identify texts on law. In general, the IDT of delphic forms does not contribute to discriminate any class, type of text or discourse relation.

i) **Hypothesis 9:** The terminological reformulation of a text is inevitably linked to the terminological density of that text and its values increase when the text specialization degree increases.

The reformulation means are similar to the terminology of the texts in the specialization degree, since the specialized text has 66% of the reformulation means (70% of terms), the semi-specialized text accounts for 23% of the means (19% of terms) and non-specialized text, 11% (10% of terms).

In fact, such is its link with the specialization degree, that the most specialized texts (article in scientific journal and doctoral thesis, in computer science and medicine respectively) are undoubtedly the most reformulated. This contradicts the idea that highly specialized texts do not suffer reformulation. However, an exception to these cases should be noted, for which we resort to the words of Cabré (1998: 63):

El grado máximo de variación de la terminología lo cumplirán los términos de las áreas más banalizadas del saber y los que se utilizarían en el discurso de registro comunicativo de divulgación de la ciencia y de la técnica; el grado mínimo de la variación sería propio de la terminología normalizada por comisiones de expertos; el grado intermedio, la terminología usada en la comunicación natural entre especialistas.

In our study, the maximum degree of variation is found in the specialized text, and not in the non-specialized text (which would correspond to the texts that represent the most vulgarized areas of knowledge), although it should be stressed that the specialized text, institutionally standardized, constitutes an exception for the reformulation if compared with our study, since, according to Cabré (*idem*), these types of text, in a certain way, deviate from the natural communication between specialists.

Specifically, with respect to the frequency of means, the most used mean is expansion, followed by variation and, finally, reduction. In the specialized text, expansion and variation means stand out above the other two types.

Having said that, specifically, only the reformulative density by expansion (and more timidly by variation) contributes to discriminate text specialization degree (in the same way that happens with the IDT). The IR by reduction does not contribute to distinguish the classes and, in addition, in general, the reformulation does not allow to distinguish the subjects either.

Finally, depending on the type of interlocutor, both in the case of the sub-speaker and in the case of the sub-recipient, the reformulative density associated with the terminology

decreases in the unlearned lay user, so it is not possible to defend that the greater the lack of knowledge of a speaker on a topic, the greater the reformulation to which the discourse on that topic is subject. In fact, it is noted that the role played by the reformulation by expansion contributes to discriminate sub-recipients, since the first increases when the competence of the second does so. In addition, it is also noted that word variation or acronymy procedures, which are the main procedures identified by variation, are not frequent in the texts received by a user with a lower cognitive-communicative competence.

3. CONTRIBUTION OF THIS THESIS TO THEORETICAL AND APPLIED LINGUISTICS

This doctoral thesis contributes in two specific aspects to the linguistic discipline, especially if we take into account that it is the result of the theoretical reflection applied to a corpus, that is, that it is an empirical study: (i) in the theoretical dimension, it stands out the role that description has in the predictive capacity of textual analysis; (ii) in the applied dimension, this predictive capacity stands out in the reconstruction of patterns in communication studies.

From the point of view of theoretical linguistics, the most interesting aspect of the results of this research corresponds to the predictive capacity provided by the theoretical description of the specialization degree, in other words, the information that previously facilitates the characterization of the three stages of the specialization degree, from a prototypical or representative point of view of the category and also attending to their exceptions or differential elements. This notion of the prototypical point of view is important, since, as Diaz Rojo (2000: 78) mentions, «la pertenencia de un elemento a una categoría se establece a partir del grado de similitud con el prototipo, sin necesidad de que todos los miembros de una clase posean todas las características de aquel». This representativity is what is achieved through the characterization of the centroid using the k-means algorithm, in which one text serves as a model for other real textual objects, due to the similarity of one text with the attributes of another (Loureda, 2002: 158).

However, it should be noted that the ability to predict can work with trends or generalizations (Cabré, 2002b; 21) and that the complex, dynamic and open nature of the communicative act makes it impossible to «predecir o describir de forma totalmente precisa» the systems (Bernárdez, 1995: 137).

In particular, this thesis provides information on the relations (direct or inverse) that exist between the three classes and the different variables, especially the terminological density and the recipient, as determining variables. In addition, it has been suggested that the perception of the informants coincides in a generalized way in the selection of values for the analysis variables, which allows having another very relevant criterion when carrying out future work.

From a methodological point of view, this doctoral thesis improves the method used in Rodríguez-Tapia (2016a), which was complex from a quantitative point of view, mainly due to the strict limits to determine the specialization degree. Thus, the conjunction of

corpus linguistics and the theory of perception appear to be a method that allows for analyzable results from a more descriptive and orientative perspective, which abandons the restriction of previous quantitative methods. It also improves the data on which the analysis is carried out, which increases in patterns, types of texts and subjects, so that representativeness is guaranteed. And, finally, it has information provided by different real users, which allows to contrast sociocognitive aspects.

On the other hand, **from the point of view of applied linguistics**, the results of this thesis can use the theoretical characterization of the specialization degree for the design of criteria that contribute to the resolution of communicative problems.

The predictive ability plays an important role for applied linguistics, since it allows the reconstruction of discursive patterns in the texts, as Loureda (2002: 161) defends: the classification of the genres allows knowing the applicability of the rules for the production and reception of discourses to (or some) speech acts and, in our opinion, the most productive aspect in our study (in the sense of automation) would be «hacer predicciones sobre la organización y el contenido de los textos a partir de sus estructuras».

Thus, this knowledge would allow, for example, to consider the type of text of a text to imitate the characteristics of the prototype. This also applies to specialized discourse and its actualization degrees. In our opinion, the characteristics related to the terminological density of semi-cryptic and cryptic forms are especially relevant and the reformulation, especially if linked to the recipient, as a regulating factor of the specialization degree. This application of theoretical knowledge is adjusted, from our point of view, to all activities of discursive creation, whether it be scientific writing, scientific divulgation or translation. In particular, translation can use the representativeness of types of lexical items in the types of text according to the specialization degree to transfer an informative content (which does not necessarily have to coincide formally between languages) from one language to another.

From the point of view of corpus compilation, the theoretical knowledge allows to consider that the selection of texts belonging to a specific specialization degree has some characteristics that represent it and that restrict the discursive possibilities to that object, so that the theoretical and predictive knowledge is also useful when filtering texts as candidates to the sample, rejecting candidates and proposing new ones. For example, knowing that specialized texts about law contain a multitude of delphic term, if we want to analyze the nature of these units in specialized discourse, the selection of patterns from

doctoral theses and scientific articles is optimal, since they are the most representative for the highest specialization degree and for a terminological density of high delphic forms. On the other hand, if we are interested in knowing the behavior of the cryptic units in the specialization degrees of divulgation, we will select types of text such as the blog or scientific-divulgation magazine or web, which are the types of texts not classified as specialized with more representativeness of cryptic and semi-cryptic lexical forms. The same happens if we intend to analyze the competence of speakers or recipients. In short, the knowledge of the characteristics that describe a degree according to specialization, a type of text, etc. allows for criteria based on scientific evidence to be used as an example or as textual support material, both in professional and research work.

4. FUTURE LINES OF RESEARCH

After indicating the main conclusions regarding to the hypotheses proposed and showing some of the contributions of the thesis to theoretical and applied linguistics, it is worth mentioning the most relevant future lines of research that should be projected in the following works in order to complement or contrast these conclusions. In our opinion, there are three aspects that should be addressed: (i) the theoretical foundation, (ii) the method and (iii) the work by analyzing perception.

On the theoretical foundations

The direct relationship observed between IDT and class and the inverse relationship between IDT and IR in some texts may be due to a change in the discursive conventions of the last decades. In our opinion, it would be convenient to carry out a comparative study with two different corpus types: one that includes texts from the 50s to 2000 and another with texts from 2000 to check if there has been a change in the paradigm of discursive traditions. This is only an approach, since the periods of compilation of the corpus must be submitted to scientific evaluation, for example, using as criteria those that use corpora such as CREA or CORPES XXI.

On the principle of activation of specialized value in common discourse, it should be noted that this principle explains the phenomenon of terminological vulgarization but does not explain the terminology and its variations. Questions such as when specialized lexical units lose their specialized value when vulgarized chronically should be answered, as can occur, for example, with the term *fiebre*, which according to the user could be considered term or non-term. If we use the terminology employed in the corresponding section, the hypothesis could be explained in the following way: if the absolute frequency of use of the unit considered T_1 in Y_n increases, it is possible that T is no longer used with the Z_aX_a characteristics, so that T would not activate its specialized value (T_0) (because the speaker no longer considers the unit as a term). In our opinion, it is necessary to go deeply into the sociocognitive aspect in a communicative framework, since it would allow to know in depth the relationship between the conceptual organization of a system in a discourse, the lexical use of this system by the speakers, its effective use by the society and its perception by the individual.

About the method

Firstly, it would be pertinent to repeat the study conducted in other languages, mainly English, for two fundamental reasons: a) its character as an instrument of universal communication in various scientific fields (especially in computer science and medicine); and b) the use of more numerous delphic lexical forms in specialized discourse. In this case it would be possible to compare the results according to the number of delphic or cryptic terms and if the informants of that language identify the specialized values (especially in the delphic terms) in the same way as Spanish speakers, both in the classes, in the subjects and in the types of text.

Also, considering the lexical perspective (but from the theory of perception) it is necessary to study terminological diversity (and not lexical diversity), as Domènech (2006: 216) argues: the number of terminological occurrences of a text must be considered, the number of different terms, the frequency of occurrence of each term and the overall proportion within the text of the distribution of this frequency to check if it is a statistically valid quantitative parameter to discriminate a text according to its specialization degree (ibid. : 216).

From another point of view, it should be useful to identify to what extent each type of term is reformulated; that is, to check which type of reformulation is most frequently associated with each type of term. Thus, it could be known if the cryptic terms receive more attention than the delphic ones, for example, by the reformulation means. In this way, it could be recognized if indeed the semantic opacity or the absence of frequency in the common discourse is a criterion used by the speakers to work with reformulation means.

One of the aspects not analyzed by lack of attention in the research work of specialized discourse analysis is syntax. In our opinion, relating the syntactic component with the rest of the variables can provide important information about the discursive configuration of the texts according to the specialization degree. It would be pertinent to perform a quantitative analysis of the frequency of appearance of the types of sentences; as qualitative, on the representativity of simple or complex sentences, and of coordinated or subordinate clauses (including the type of subordination that each text presents, linked to the type of text, the subject, the class, etc.).

Finally, we can resort to the application of ordinal classification methods (Sáez *et al.*, 2016) of machine learning, since actually the nominal variables we use are ordinal, or the conditional random field models (Wallach, 2004), which could expand expectations in the analysis of results.

On perception

Considering the subclassification of interlocutors, it has been observed that the analysis carried out by the informants tends towards three concrete sub-profiles, which, in our opinion, would be constituted as prototypes for the supercategories of specialist, learned user and lay user. Subclassifications deserve attention in future work to try to check the limits of the continuum they present and their prototypical and peripheral nature regarding supercategories.

Also related to the perception of the speakers, it is pertinent to analyze the objective frequency of the delphic, semi-cryptic and cryptic terms, that is, to check with real examples extracted from the corpus of analysis and compare them with the frequency of reference corpus of Spanish, as CORPES XXI. In this way, in our opinion, it would be possible to design a model that allows combining the subjective factor with the objective evidence.

On the other hand, an additional group of informants could be formed by experts in the subjects that constitute the corpus, so that it could be measured and evaluated if the experts determine the specialization degree, the specialized value of the lexical forms and other relations of the same way in which the rest of groups of informants do it or if, in some way, they are conditioned by the frequency of appearance of said terms in the texts with which they have contact in their daily professional work.

Finally, it would be interesting to analyze the centroid or prototypical texts that characterize the text specialization degree. This research would consist in measuring such prototypical texts that speakers use as reference according to their profile, since all speakers use perception to measure if an object corresponds with a value or another with respect to a cognitive model that has been previously created (which has the prototypical characteristics that lean toward one value or another).

VI. REFERENCIAS BIBLIOGRÁFICAS

1. REFERENCIAS BIBLIOGRÁFICAS

- ADAM, Jean-Michel (1990): *Éléments de linguistique textuelle. Théorie et pratique de l'analyse textuelle*. Liège: Mardaga.
- ADAM, Jean-Michel (2001 [1992]): *Les textes. Types et prototypes*. Paris: Nathan-Her.
- ADELSTEIN, Andreína (1998): «Banalización de términos con formantes de origen grecolatino». En: RED IBEROAMERICANA DE TERMINOLOGÍA (ed.), *Actas del V Simposio Iberoamericano de Terminología-Riterm*. México: Riterm, pp. 12-17.
- ADELSTEIN, Andreína (2004): *Unidad léxica y valor especializado: estado de la cuestión y observaciones sobre su representación*. Trabajo de investigación de doctorado. Barcelona: Institut Universitari de Lingüística Aplicada. Universitat Pompeu Fabra. Disponible en: http://elies.rediris.es/miscelanea/misce_7/adelstein.pdf.
- ADELSTEIN, Andreína (2007a): *Unidad léxica y significado especializado: modelo de representación a partir del nombre relacional madre*. Tesis doctoral. Barcelona: Institut Universitari de Lingüística Aplicada. Universitat Pompeu Fabra. Disponible en: <http://www.tdx.cat/bitstream/handle/10803/7505/taa.pdf?sequence=1>.
- ADELSTEIN, Andreína (2007b): «Valor especializado de nombres relacionales en compuestos terminológicos». En: LORENTE, Mercè; ESTOPÀ, Rosa; FREIXA, Judit; MARTÍ, Jaume y TEBÉ, Carles (eds.): *Estudis de lingüística i de lingüística aplicada en honor de M.^a Teresa Cabré Castellví*. Vol. 2. Barcelona: Institut Universitari de Lingüística Aplicada. Universitat Pompeu Fabra, pp. 165-178.
- ALARCÓN, Rodrigo; BACH, Carme y SIERRA, Gerardo (2007): «Extracción de contextos definitorios en corpus especializados: hacia la elaboración de una herramienta de ayuda terminográfica», *Revista Española de Lingüística* 37, pp. 247-277. Disponible en: <http://www.sel.edu.es/pdf/ene-dic-07/RSEL-37-Alarcon-Bach-Sierra.pdf>.
- ALCARAZ VARÓ, Enrique (1990): *Tres paradigmas de la investigación lingüística*. Alcoy: Marfil.
- ALCARAZ VARÓ, Enrique y MARTÍNEZ LINARES, María Antonia (2004): *Diccionario de lingüística moderna*. Barcelona: Ariel.
- ALONSO, Araceli y DECESARIS, Janet (2006): «El valor especializado de las unidades léxicas en el área del medio ambiente». En: CABRÉ, María Teresa; ESTOPÀ, Rosa y TEBÉ, Carles (eds.): *La terminología en el siglo XXI. Contribución a la cultura de la paz, la diversidad y la sostenibilidad*:

- actas del IX Simposio Iberoamericano de Terminología RITERM04*. Barcelona: Institut Universitari de Lingüística Aplicada. Universitat Pompeu Fabra, pp. 269-288.
- ALVAR, Manuel (1976): *Lengua y sociedad*. Barcelona: Planeta.
- ÁLVAREZ, Gerardo (2004 [1996]). *Textos y discursos. Introducción a la lingüística del texto*. Concepción: Universidad de Concepción.
- ÁLVAREZ, Piedad (2008): *La sintaxis del lenguaje de los textos científicos. Los tipos oracionales y los giros de participio y gerundio: estudio de un corpus ruso-español de textos médicos*. Trabajo de investigación. Barcelona: Universitat de Barcelona. Disponible en: https://ddd.uab.cat/pub/trerecpro/2008/hdl_2072_170118/La_sintaxis_del_lenguaje_de_los_textos_cientificos.pdf.
- ARNTZ, Reiner y PICT, Heribert (1995): *Introducción a la terminología*. (A. Irazábal, trad.). Madrid: Fundación Germán Sánchez Ruipérez.
- AUSTIN, John (1962): *How to Do Things with Words*. London: Clarendon Press.
- BACH, Carme (2001): «La equivalencia parafrástica en los textos especializados en vista de la detección de información paralela». En: CABRÉ, María Teresa y FELIU, Judit (eds.): *La terminología científico-técnica: reconocimiento, análisis y extracción de información formal y semántica*. Barcelona: Institut Universitari de Lingüística Aplicada. Universitat Pompeu Fabra, pp. 217-226.
- BACH, Carme; FREIXA, Judit y SUÁREZ DE LA TORRE, Mercedes (2003): «Equivalencia conceptual y reformulación parafrástica en terminología». En: CORREIRA, Margarita (ed.): *Terminologia e Industria da Lingua*. Lisboa: ILTEC, pp. 173-184.
- BAJTÍN, Mijaíl M. (1999 [1979]): «El problema de los géneros discursivos». En: BAJTÍN, Mijaíl M. (ed.): *Estética de la creación verbal*. México D. F.: Siglo XXI, pp. 248-293.
- BALBONI, Paolo (1986): «LGP versus LSP». *Unesco ALSED-LSP Newsletter* 9(1), pp. 2-8.
- BALDINGER, Kurt (1952): «Die Gestaltung des wissenschaftlichen Wörterbuchs», *Romanistisches Jahrbuch* 5, pp. 65-94.
- BALLMER, Thomas (1985): «The Psychology of Context Change». En: BALLMER, Thomas (ed.): *Linguistic Dynamics. Discourses, Procedures and Evolution*. Berlin/New York: Walter de Gruyter, pp. 322-355.
- BALLMER, Thomas y WILDGEN, Wolfgang (eds.). (1987): *Process Linguistics: Exploring the processual aspects of language and language use, and the methods of*

- their description*. Tübingen: Niemeyer.
- BAYES, Thomas (1763): «An Essay towards solving a Problem in the Doctrine of Chances», *Philosophical Transactions of the Royal Society of London* 53, pp. 370-418. Disponible en: <http://rstl.royalsocietypublishing.org/content/53/370.full.pdf+html>.
- BEAUGRANDE, Robert-Alain (1988): «Teoría lingüística y metateoría para una ciencia del texto». En: BERNÁRDEZ, Enrique (ed.): *Lingüística del texto*. Madrid: Arco Libros, pp. 35-94.
- BEAUGRANDE, Robert-Alain y DRESSLER, Wolfgang Ulrich (1997): *Introducción a la lingüística del texto*. (S. Bonilla, trad.). Barcelona: Ariel.
- BECERRA-BONACHE, Leonor y JIMÉNEZ-LÓPEZ, María Dolores (2015): «Un modelo de aprendizaje automático para medir la complejidad de las lenguas», *E-AESLA* 1, pp. 1-10. Disponible en: <http://cvc.cervantes.es/lengua/eaesla/pdf/01/53.pdf>.
- BEN-GAL, Irad (2007): «Bayesian Networks». En: RUGGERI, Fabrizio; KENNETT, Ron S. y FALTIN, Frederick W. (eds.): *Encyclopedia of Statistics in Quality and Reliability*. John Wiley & Sons. DOI: 10.1002/9780470061572.eqr089.
- BENSON, Morton *et al.* (1986): *The BBI combinatorial dictionary of English: A guide to word combinations*. Amsterdam: John Benjamins.
- BERNÁRDEZ, Enrique (1982): *Introducción a la lingüística del texto*. Madrid: Espasa-Calpe.
- BERNÁRDEZ, Enrique (1987): «Introducción». En: BERNÁRDEZ, Enrique (ed.): *Lingüística del texto*. Madrid: Arco/Libros, pp. 7-18.
- BERNÁRDEZ, Enrique (1995): *Teoría y epistemología del texto*. Madrid: Cátedra.
- BERNS, Margie y MATSUDA, Paul Kei (2006): «Applied Linguistics: Overview and History». En: BROWN, Keith (ed. in chief): *Encyclopedia of Language and Linguistics*. Amsterdam: Elsevier, pp. 394-405. DOI: <http://dx.doi.org/10.1016/B0-08-044854-2/00599-X>.
- BIBER, Douglas (1988): *Variation across Speech and Writing*. Cambridge: Cambridge University Press.
- BIBER, Douglas (1990): «Methodological issues regarding corpus-based analyses of linguistic variation», *Literary and Linguistic Computing* 5(4), pp. 257-269.
- BIBER, Douglas (1993): «Representativeness in Corpus Design», *Literary and Linguistic Computing* 8(4), pp. 243-257.
- BIBER, Douglas; CONRAD, Susan y REPPEN, Randi (1998): *Corpus*

- Linguistics. Investigating Language Structure and Use.* Cambridge: Cambridge University Press.
- BISHOP, Cristopher M. (2010): *Pattern recognition and machine learning.* New York: Springer.
- BLECUA, José M. (1996): «Reflexiones al margen de los corpus escritos». En: PAYRATÓ, Lluís; BOIX, Emili; LLORET, M. Rosa y LORENTE, Mercè (eds.): *Corpus, corpora.* Barcelona: Promociones y publicaciones universitarias.
- BLOOMFIELD, Leonard (1984 [1933]): *Language.* Chicago/London: University of Chicago.
- BRIGHT, William (1992): *International Encyclopedia of Linguistics.* New York: Oxford University.
- BRINKER, Klaus (1988): *Linguistische Textanalyse.* Berlin: E. Schmidt.
- BROWN, H. Douglas (1987): *Principles of Language Learning and Teaching.* Englewood Cliffs, NJ: Prentice Hall.
- BRUMFIT, Cristopher (1997): «How Applied Linguistics Is the Same as Any Other Science», *International Journal of Applied Linguistics* 7, pp. 86-94.
- BRUMFIT, Cristopher (2004): «Applied linguistics in 2004: Unity in diversity?», *AILA Review* 17, pp. 133-136.
- BÜHLER, Karl (1979 [1934]): *Teoría del lenguaje.* (J. Marias, trad.). Madrid: Alianza.
- BYGATE, Martin (2005): «Applied linguistics: a pragmatic discipline, a generic discipline?», *Applied Linguistics* 26, pp. 568-581.
- CABRÉ, María Teresa (1993): *La terminología: teoría, metodología, aplicaciones.* Barcelona: Antártida/Empúries.
- CABRÉ, María Teresa (1998): «Elementos para una teoría de la terminología: hacia un paradigma alternativo», *El Lenguaraz. Revista Académica Del Colegio de Traductores Públicos de La Ciudad de Buenos Aires* I(1), pp. 59-78.
- CABRÉ, María Teresa (1999): *La terminología: representación y comunicación. Elementos para una teoría de base comunicativa y otros artículos.* Barcelona: Institut Universitari de Lingüística Aplicada. Universitat Pompeu Fabra.
- CABRÉ, María Teresa (2000a): «Terminología i diccionaris III. A propòsit del Diccionari de la llengua catalana». En: GINEBRA, Jordi; MARTÍNEZ GILI, Raül-David y PRADILLA, Miquel Àngel (eds.): *Symposia Philologica III. La lingüística de Pompeu Fabra.* Alicante: Universitat Rovira i Virgili, pp. 231-270.
- CABRE, María Teresa (2000b): «Terminologie et linguistique: la

- théorie des portes», *Terminologies nouvelles* 21, pp. 10-15.
- CABRÉ, María Teresa (2002a): «Análisis textual y terminología, factores de activación de la competencia cognitiva en la traducción». En: ALCINA, Amparo y GAMERO, Silvia (eds.): *La traducción científico-técnica y la terminología en la sociedad de la información*. Castelló de la Plana: Publicacions de la Universitat Jaume I, pp. 87-105.
- CABRÉ, María Teresa (2002b): «Textos especializados y unidades de conocimiento: metodología y tipologización». En: GARCÍA, Joaquín y FUENTES, María Teresa (eds.): *Texto, terminología y traducción*. Salamanca: Almar, pp. 15-36.
- CABRÉ, María Teresa (2007a): «Constituir un corpus de textos de especialización: condiciones y posibilidades». En: BALLARD, Michel y PINEIRA-TRESMONTANT, Carmen (ed.): *Les corpus en linguistique et en traductologie*. Arras: Artois Presses Université, pp. 89-106.
- CABRÉ, María Teresa (2007b): «Términos y palabras en los diccionarios». En: EMSEL, Martina y CUARTERO OTAL, Juan (eds.): *Vernetzungen: Bedeutung in Wort, Satz und Text. Festschrift für Gerd Wotjak zum 65. Geburtstag*. Frankfurt am Main: Peter Lang, pp. 71-84.
- CABRÉ, María Teresa (2008): «El principio de poliedricidad: la articulación de lo discursivo, lo cognitivo y lo lingüístico en Terminología», *Ibérica: Revista de la Asociación Europea de Lenguas para Fines Específicos (AELFE)* 16, pp. 9-36.
- CABRÉ, María Teresa (2009): «La Teoría Comunicativa de la Terminología, una aproximación lingüística a los términos», *Revue française de linguistique appliquée* 2(XIV), pp. 9-15. Disponible en: www.cairn.info/revue-francaise-de-linguistique-appliquee-2009-2-page-9.htm.
- CABRÉ, María Teresa y BACH, Carme (2004): «El corpus tècnic del IULA: corpus textual especializado plurilingüe», *Panace@* V(16), pp. 173-176. Disponible en: <https://repositori.upf.edu/bitstream/handle/10230/6165/pana6165.pdf?sequence=1&isAllowed=y>.
- CABRÉ, María Teresa y ESTOPÀ, Rosa (2002): «El conocimiento especializado y sus unidades de representación: diversidad cognitiva», *Sendebarr* 13, pp. 141-153.
- CABRÉ, María Teresa y ESTOPÀ, Rosa (2005): «Unidades de conocimiento especializado: caracterización y tipología». En: CABRÉ, María Teresa y BACH, Carme (eds.): *Coneixement, llenguatge i discurs especialitzat*. Barcelona: Institut Universitari de Lingüística

- Aplicada. Universitat Pompeu Fabra, pp. 69-93.
- CABRÉ, María Teresa; BACH, Carme; CASTELLÀ, Josep M. y MARTÍ, Jaume (2007). «La caracterización lingüística del discurso especializado». En: *Aprendizaje de lenguas, uso del lenguaje y modelación cognitiva: actas del XXIV Congreso Nacional de AESLA*. Madrid: UNED, pp. 851-857.
- CABRÉ, María Teresa; BACH, Carme; DA CUNHA, Iria; MORALES, Albert y VIVALDI, Jorge (2010): «Comparación de algunas características lingüísticas del discurso especializado frente al discurso general: el caso del discurso económico». En: *Proceedings of the XXVII AESLA international conference: Ways and modes of human communication (AESLA 2009)*. Ciudad Real: Universidad de Castilla-La Mancha, pp. 453-460.
- CABRÉ, María Teresa; DA CUNHA, Iria; SAN JUAN, Eric; TORRES-MORENO, Juan-Manuel y VIVALDI, Jorge (2011): «Automatic specialized vs. non-specialized texts differentiation: a first approach». En: TALAVÁN, Noa; MARTÍN MONJE, Elena y PALAZÓN, Francisco (eds.): *Technological innovation in the teaching and processing of LSPs: Proceedings of TISLID'10*. Madrid: Universidad Nacional de Educación a Distancia (UNED), pp. 301-310.
- CABRÉ, María Teresa; DA CUNHA, Iria; SAN JUAN, Eric; TORRES-MORENO, Juan-Manuel; VIVALDI, Jorge (2014): «Automatic specialized vs. non-specialized text differentiation: the usability of grammatical features in a Latin multilingual context». En: BÁRCENA, Elena; READ, Timothy y ARÚS, Jorge (eds.): *Languages for Specific Purposes in the Digital Era*. Berlín: Springer, pp. 223-241.
- CABRÉ, María Teresa; DOMÈNECH, Meritxell; MOREL, Jordi y RODRÍGUEZ, Carlos (2001): «Las características del conocimiento especializado y la relación con el conocimiento general». En: CABRÉ, María Teresa y FELIU, Judit (eds.): *La terminología científico-técnica: reconocimiento, análisis y extracción de información formal y semántica*. Barcelona: Institut Universitari de Lingüística Aplicada. Universitat Pompeu Fabra, pp. 173-86.
- CABRÉ, María Teresa; FELIU, Judit y TEBÉ, Carles (2001): «Bases cognitivas de la terminología: hacia una visión comunicativa del concepto», *Sendebarr* 12, pp. 301-310.
- CALSAMIGLIA, Helena y TUSÓN, Amparo (2007): *Las cosas del decir: manual de análisis del discurso*. Barcelona: Ariel.
- CAMPOS ANDRÉS, Olga (2013): «Procedimientos de desteminologización: traducción y redacción de guías para pacientes»,

- Panace@* XIV(37), pp. 48-52. Disponible en: <http://www.tremedica.org/panacea/IndiceGeneral/n37-tradyterm-OCamposAndres.pdf>.
- CAMPOS PLAZA, Nicolás y CAMPOS MARTÍN, Natalia (2009): «La terminología jurídica (francés-español) en clave traductológica: términos délficos, términos crípticos. Tipología textual y cultura jurídica». En: SANTA, Àngels y SOLÉ, Cristina (eds.): *Texto y sociedad en las letras francesas y francófonas*. Lleida: Universidad de Lleida, pp. 686-97.
- CARAVEDO, Rocío (1990): «Los espacios de variabilidad en fonología», *Voz y Letra* II(1), pp. 17-38.
- CARAVEDO, Rocío (1990): *Sociolingüística del español de Lima*. Lima: Pontificia Universidad Católica del Perú.
- CARAVEDO, Rocío (1999): *Gramática española: enseñanza e investigación. Apuntes metodológicos. Lingüística del corpus. Cuestiones teórico-metodológicas aplicadas al español*. Salamanca: Ediciones Universidad de Salamanca.
- CARAVEDO, Rocío (2003): «Problemas conceptuales y metodológicos de la lingüística de la variación». En: MORENO FERNÁNDEZ, Francisco *et al.* (eds.): *Lengua, variación y contexto. Estudios dedicados a Humberto López Morales*. Madrid: Arco Libros, pp. 541-558.
- CARAVEDO, Rocío (2005): «Sobre el corpus en la dialectología actual». En: LOZANO, Mariano (dir.): *Homenaje a José Joaquín Montes Giraldo*. Bogotá: Instituto Caro y Cuervo, pp. 94-107.
- CARAVEDO, Rocío (2014): *Percepción y variación lingüística: enfoque sociocognitivo*. Madrid/Frankfurt am Main: Iberoamericana/Vervuert.
- CARMONA, Elena (2015): *El lector (inter)activo en la prensa escrita española*. Sevilla: Universidad de Sevilla. Tesis doctoral. Disponible en: <https://idus.us.es/xmlui/handle/11441/33271>.
- CARTER, Ronald (1987): *Vocabulary. Applied linguistic perspectives*. London: Allen & Unwin Publishers Ltd.
- CASAS GÓMEZ, Miguel (2003): «Hacia una tipología de la variación». En: MORENO FERNÁNDEZ, Francisco; GIMENO, Francisco; SAMPER, José Antonio; GUTIÉRREZ, M.^a Luz; VAQUERO, María y HERNÁNDEZ, César (coords.): *Lengua, variación y contexto. Estudios dedicados a Humberto López Morales*, vol. 2. Madrid: Arco/Libros, pp. 559-574.
- CHAREAUDEAU, Patrick (1992): *Grammaire du sens et de l'expression*. Paris: Hachette.
- CHOMSKY, Noam (1965): *Aspects of a Theory of Syntax*. Cambridge, Mass: MIT Press.

- CHOMSKY, Noam (1985 [1957]): *Syntactic Structures*. The Hague: Mouton.
- CHOMSKY, Noam (2006 [1968]): *Language and Mind*. Cambridge: Cambridge University Press.
- CIAPUSCIO, Guiomar Elena (1998): «La terminología desde el punto de vista textual: selección, tratamiento y variación», *Organon* 26, pp. 43-65.
- CIAPUSCIO, Guiomar Elena (2003): *Textos especializados y terminología*. Barcelona: Institut Universitari de Lingüística Aplicada, Universitat Pompeu Fabra.
- CIAPUSCIO, Guiomar Elena y KUGUEL, Inés (2002): «Hacia una tipología del discurso especializado: aspectos teóricos y aplicados». En: GARCÍA, Joaquín y FUENTES, María Teresa (eds.): *Texto, terminología y traducción*. Salamanca: Almar, pp. 37-73.
- CIAPUSCIO, Guiomar Elena; KUGUEL, Inés y OTAÑI, Isabel (2005): «El conocimiento especializado: el texto de especialización y los criterios para su tipologización». En: CABRÉ, María Teresa y BACH, Carme (eds.): *Coneixement, llenguatge i discurs especialitzat*. Barcelona: Institut Universitari de Lingüística Aplicada. Universitat Pompeu Fabra, pp. 95-110.
- CONDAMINES, Anne (1995): «Terminology: New needs, new perspectives», *Terminology* 2(2), pp. 219-238.
- CONTRERAS IZQUIERDO, Narciso Miguel (2008): *El diccionario de lengua y el conocimiento especializado*. Jaén: Universidad de Jaén.
- CORBACHO SÁNCHEZ, Alfonso (2006): «Textos, tipos de texto y textos especializados», *Revista de filología de la Universidad de La Laguna* 24, pp. 77-90. Disponible en: <https://dialnet.unirioja.es/descarga/articulo/2100070.pdf>.
- CORDER, S. Pit (1973): *Introducing Applied Linguistics*. Harmondsworth: Penguin.
- CORNFIELD, Jerome; GORDON, Tavia y SMITH, Willie (1961): «Quantal response curves for experimentally uncontrolled variables», *Bulletin of the International Statistical Institute* 38, pp. 97-115.
- CORPAS, Gloria (1997): *Manual de fraseología española*. Madrid: Gredos.
- CORPAS, Gloria (1998): «Criterios generales de clasificación del universo fraseológico de las lenguas, con ejemplos tomados del español y del inglés». En: ALVAR, Manuel y CORPAS, Gloria (eds.): *Diccionarios, Frases, Palabras*. Málaga: Servicio de Publicaciones de la Universidad, pp. 157-187.
- CORPAS, Gloria y SEGHIRI, Míriam (2010): «Size Matters: A

- Quantitative Approach To Corpus Representativeness». En: RABADÁN, Rosa; GUZMÁN, Trinidad y FERNÁNDEZ, Marisa (ed.): *Lengua, traducción, recepción: en honor de Julio César Santoyo*. León: Universidad de León, pp. 112-146. Disponible en: http://www.lexytrad.es/assets/cl24_0.pdf.
- CORPAS, Gloria y SEGHIRI, Míriam (2009): «Virtual Corpora as Documentation Resources: Translating Travel Insurance Documents (English-Spanish)». En: BEEBY, Allison; RODRÍGUEZ INÉS, Patricia y SÁNCHEZ-GIJÓN, Pilar (Eds.): *Corpus Use and Translating*. Amsterdam/Philadelphia: John Benjamins, pp. 75-107. Disponible en: http://www.lexytrad.es/assets/c115_0.pdf.
- CORTÉS, Luis y CAMACHO, M.^a Matilde (2005): *Unidades de segmentación y marcadores del discurso*. Madrid: Arco Libros.
- COSERIU, Eugenio (1962): «Determinación y entorno». En: *Teoría del Lenguaje y lingüística general*. Madrid: Gredos, pp. 282-323.
- COSERIU, Eugenio (1992 [1982]): «La competencia lingüística general. (El saber elocutivo)». En: *Competencia lingüística*. Madrid: Gredos, pp. 106-180.
- COSERIU, Eugenio (1998): «Texto, valores y enseñanza». En: CASADO
- VELARDE, Manuel (ed.): *Lengua, literatura y valores*. Pamplona: Newbook, pp. 61-78.
- CRYSTAL, David (1980): *A First Dictionary of Linguistics and Phonetics*. London: Deutsch.
- CUENCA, María Josep y HILFERTY, Joseph (1999): *Introducción a la lingüística cognitiva*. Barcelona: Ariel.
- DA CUNHA, Iria; CABRÉ, María Teresa; SAN JUAN, Eric; SIERRA, Gerardo; TORRES-MORENO, Juan-Manuel y VIVALDI, Jorge (2011): «Automatic Specialized vs. Non-Specialized Sentence Differentiation». En: GELBUKH, Alexander (ed.): *Computational Linguistics and Intelligent Text Processing, Lecture Notes in Computer Science (LNCS)*. Berlin: Springer, pp. 266-276.
- DAVIES, Alan (1999): *An Introduction to Applied Linguistics. From Practice to Theory*. Edinburgh: Edinburgh University Press.
- DAVIES, Alan (2004): «Introduction to part I: Linguistics Applied (L-A)». En: DAVIES, Alan y ELDER, Catherine: *The Handbook of Applied Linguistics*. London: Blackwell Publishing, pp. 19-24.
- DAVIES, Alan y ELDER, Catherine (2004): *The Handbook of Applied Linguistics*. London: Blackwell Publishing.
- DÍAZ ROJO, José Antonio (2000): «El modelo ecolingüístico de

- comunicación especializada: investigación y divulgación», *Ibérica: Revista de la Asociación Europea de Lenguas para Fines Específicos (AELFE)* 2, pp. 69-84. Disponible en: <http://www.aelfe.org/documents/te xt2-Diaz.pdf>.
- DIÉGUEZ, Isabel (2006): «Análisis terminológico de anglicismos en el discurso económico semiespecializado del español de Chile». En: CABRÉ, María Teresa; ESTOPÀ, Rosa y TEBÉ, Carles (eds.) (2006): *La terminología en el siglo XXI. Contribución a la cultura de la paz, la diversidad y la sostenibilidad: actas del IX Simposio Iberoamericano de Terminología RITERM04*. Barcelona: Institut Universitari de Lingüística Aplicada. Universitat Pompeu Fabra, pp. 513-524.
- DIJK, Teun A. van (1980): *Texto y contexto: semántica y pragmática del discurso*. (J. Domingo Moyano, trad.). Madrid: Cátedra.
- DIJK, Teun A. van (1990): *La noticia como discurso: comprensión, estructura y producción de la información*. (G. Gal, trad.). Barcelona: Paidós.
- DIJK, Teun A. van (1992 [1983]): *La ciencia del texto*. Barcelona: Paidós.
- DOLEŽEL, Lubomír (1998): *Heterocosmica: Fiction and possible worlds*. Baltimore: Johns Hopkins University Press.
- DOMÈNECH, Meritxell (1998): *Unitats de coneixement i texts especialitzats: primera proposta d'anàlisi*. Trabajo de investigación no publicado. Institut Universitari de Lingüística Aplicada. Universitat Pompeu Fabra.
- DOMÈNECH, Ona (2006): *Textos especialitzats i variació vertical: la diversitat terminològica com a factor discriminant del nivell d'especialització d'un text*. Tesis doctoral. Barcelona: Institut Universitari de Lingüística Aplicada. Universitat Pompeu Fabra.
- DOMÈNECH, Ona (2007): «La noció de text especialitzat des de la perspectiva de la teoria comunicativa de la terminologia». En: LORENTE, Mercè; ESTOPÀ, Rosa; FREIXA, Judit; MARTÍ, Jaume y TEBÉ, Carles (eds.): *Estudis de lingüística i de lingüística aplicada en honor de M. Teresa Cabré Castellví*. Vol. 2. Barcelona: Institut Universitari de Lingüística Aplicada. Universitat Pompeu Fabra, pp. 241-254.
- DRAE – Real Academia Española (2014): *Diccionario de la lengua española*. Madrid: Real Academia Española. Disponible en línea en: <http://dle.rae.es/index.html>.
- DROZD, Lubomir y SEIBICKE, Wilfried (1973): *Deutsche Fach- und Wissenschaftssprache. Bestandsaufnahme. Theorie. Geschichte*. Wiesbaden: Brandstetter.

- DUBOIS, Jean (1973): *Dictionnaire de linguistique*. Paris: Larousse.
- EAGLES (1996): «Definitions», *EAGLES Preliminary recommendations on Corpus Typology EAG--TCWG--CTYP/P*. Disponible en: <http://www.ilc.cnr.it/EAGLES96/corpustyp/corpustyp.html>.
- ELLIS, Rod (1990): «Response to Gregg», *Applied Linguistics* 11(4), pp. 384-391.
DOI: <http://dx.doi.org/10.1093/app/lin/11.4.384>.
- ESTANY, Anna (2001): «La conceptualización de la realidad». En: CABRÉ, María Teresa y FELIU, Judit (eds.): *Terminología y cognición*. Barcelona: Institut Universitari de Lingüística Aplicada. Universitat Pompeu Fabra.
- ESTOPÀ, Rosa (1999): *Extracció de terminologia. Elements per a la construcció d'un SEACUSE (Sistema d'Extracció Automàtica de Candidats a Unitats de Significació Especialitzada)*. Tesis doctoral. Barcelona: Institut Universitari de Lingüística Aplicada. Universitat Pompeu Fabra. Disponible en: <http://www.tdx.cat/bitstream/handle/10803/7489/treb1de2.pdf;jsessionid=5B7CB700E2916C46585132EFFF6E3518.tdx2?sequence=1>.
- ESTOPÀ, Rosa (2014): «El reciclaje léxico: mecanismo de actualización terminológica recurrente entre los especialistas», *Debate Terminológico* 11, pp. 38-46. Disponible en: <http://seer.ufrgs.br/riterm/article/view/50977/31691>.
- FABER, Pamela (2009): «The cognitive shift in terminology and specialized translation», *MONTI: Monografías de traducción e interpretación* 1, pp. 107-134. Disponible en: http://rua.ua.es/dspace/bitstream/10045/13039/1/MonTI_01_10.pdf
- FASOLD, Ralph W. (1984): *Sociolinguistics of Language*. Oxford: Blackwell.
- FASOLD, Ralph W. (1996): *La sociolingüística de la sociedad: introducción a la sociolingüística*. Madrid: Visor.
- FILLMORE, Carles J. (1985): «Frames and the Semantics of Understanding», *Quademidi Semantica* 6, pp. 222-255.
- FILLMORE, Carles J. y ATKINS, Beryl (1992): «Towards a Frame-Based Lexicon: the Semantics of RISK and its Neighbors». En: LEHRER, Adrienne y KITTAY, Eva Feder (ed.): *Frames, Fields and Contrasts*. Hillsdale, New Jersey: Lawrence Erlbaum.
- FREIXA, Judit (2002a): «Reflexiones acerca de las causas de la variación denominativa en terminología». En: GUERRERO, Gloria y PÉREZ, Manuel Fernando (eds.): *Panorama actual de la terminología*. Granada: Comares, pp. 107-115.

- FREIXA, Judit (2002b): *La variació terminològica. Anàlisi de la variació denominativa en textos de diferent grau d'especialització de l'àrea de medi ambient*. Tesis doctoral. Barcelona: Institut Universitari de Lingüística Aplicada. Universitat Pompeu Fabra.
- FREIXA, Judit (2005): «Variación terminológica: ¿por qué y para qué?», *Meta: Journal Des Traducteurs / Meta: Translators' Journal*, 50(4). Disponible en: <http://www.erudit.org/revue/meta/2005/v50/n4/019917ar.pdf>.
- FREIXA, Judit (2013): «Otra vez sobre las causas de la variación denominativa», *Debate Terminológico* 9, pp. 38-46. Disponible en: <http://seer.ufrgs.br/riterm/article/viewFile/37170/24032>.
- FUCHS, Catherine (1994): *Paraphrase et énonciation*. Paris: Ophrys.
- FUENTES, Catalina (1993): «Conclusivos y reformulativos», *Verba* 20, pp. 171-198.
- FUERTES-OLIVERA, Pedro A. y TARP, Sven (2014): *Theory and Practice of Specialised Online Dictionaries*. Berlin; Boston: De Gruyter.
- GARCÉS GÓMEZ, María del Pilar (2006): «Las operaciones de reformulación». En VILLAYANDRE, Milka (ed.): *Actas del XXXV Simposio Internacional de la Sociedad Española de Lingüística*. León: Universidad de León, pp. 654-672.
- GARCÉS GÓMEZ, María del Pilar (2008): *La organización del discurso: marcadores de ordenación y de reformulación*. Madrid/Frankfurt am Main: Iberoamericana/Vervuert.
- GARCÍA NEGRONI, María Marta (2009): «Reformulación parafrástica y ethos discursivo en la escritura académica en español. Contraste entre escritura experta y escritura universitaria avanzada», *Letras de Hoje* 44(1), pp. 46-56. Disponible en: <http://revistaseletronicas.pucrs.br/ojs/index.php/fale/article/download/5652/4119>.
- GAUDIN, François (1993): «Socioterminologie: propos et propositions épistémologiques», *Le Language et l'Homme* XXVIII(4), pp. 247-258.
- GAUDIN, François (1995): «Champs, clotûres et domaines: des langues de la spécilaité à la culture scientifique», *Meta* 40(2), pp. 229-237.
- GAVRANEK, B. (1967): «Zadatxi literatúrno go iazikà i iegó cultura». En: KONDRAIXOV, N. A. (ed.): *Prajski lingvistítxeski krujok*. Moscú, pp. 338-377.
- GIL, José María (2007): *Lingüística generativa: un análisis histórico y epistemológico*. Santiago de Chile: RIL editores.

- GLÄSER, Rosemarie (1981, 1982): «The Problem of Style Classification in LSP (ESP)». En: HØEDT, Jørgen; LUNDQUIST, Lita; PICHT, Heribert y QVISTGAARD, Jacques (eds.): *Proceedings of the 3rd European Symposium on LSP*. Copenhagen: Erhversøkonomisk Forlag.
- GLÄSER, Rosemarie (1993): «A Multi-level Model for a Typology of LSP Genres», *Fachsprache International Journal of LSP*, 1-2(15), pp. 18-26.
- GÓMEZ DE ENTERRÍA, Josefa (2001): «La enseñanza-aprendizaje del español con fines específicos. Las lenguas de especialización. Su aplicación a la enseñanza del Español como Lengua Extranjera». En: GÓMEZ DE ENTERRÍA, Josefa (coord.): *La enseñanza-aprendizaje del español con fines específicos*. Madrid: Edinumen, pp. 7-18.
- GÓMEZ DE ENTERRÍA, Josefa (2007): «La enseñanza del español con fines específicos». En: LACORTE, Manel (coord.): *Lingüística aplicada del español*. Madrid: Arco Libros, pp. 149-181.
- GRABE, William (2002): «Applied linguistics: an emerging discipline for the twenty-first century». En: KAPLAN, Robert B. (ed.): *The Oxford Handbook of Applied Linguistics*. New York: Oxford University Press, pp. 3-12.
- GRABE, William (2010): «Applied linguistics: a twenty-first century discipline». En: KAPLAN, Robert B. (ed.): *The Oxford Handbook of Applied Linguistics*. Oxford: Oxford University Press, pp. 34-44.
- GRANGER, Sylviane y WYNNE, Martin (2000): *Optimising measures of lexical variation in EFL learner corpora*. En: KIRK, John M. (ed.): *Corpora galoreh*. Amsterdam: Rodopi, pp. 249-258.
- GUANTIVA, Ricardo; CABRÉ, María Teresa y CASTELLÀ, Josep M. (2008): «Clasificación de textos especializados a partir de su terminología», *Íkala. Revista de lenguaje y cultura* 13(19), pp. 15-39. Disponible en: <http://www.redalyc.org/html/2550/255020449002/>.
- GUERRERO, Gloria (1995): *Neologismos en el español actual*. Madrid: Arco/Libros.
- GUERRERO, Gloria (1999): «¿Tecnolectos, lenguajes (lenguas) específicos, especiales, especializados o de especialización?». En FERNÁNDEZ GONZÁLEZ, J. (ed.): *Lingüística para el siglo XXI. Actas del XXI Congreso de Lingüística General*. Salamanca: Universidad de Salamanca, pp. 879-887.
- GÜLICH, Elisabeth y KOTSCHI, Thomas (1987): «Les actes de reformulation dans la consultation "La dame de Caluire"». En: BANGE, Pierre (ed.): *L'analyse des interactions verbales*. Berne: Lang, pp. 15-81.

- GÜLICH, Elisabeth y KOTSCHI, Thomas (1995): «Discourse Production in Oral Communication». En: QUASTHOFF, Uta M. (ed.): *Aspects of Oral Communication*. Berlin: de Gruyter, pp. 30-66.
- GUTIÉRREZ, Bertha M. (1998): *La ciencia empieza en la palabra. Análisis e historia del lenguaje científico*. Barcelona: Península.
- HALLIDAY, Michael Alexander Kirkwood (1982 [1979]): *El lenguaje como semiótica social: la interpretación social del lenguaje y del significado* (J. Ferreiro Santana, trad.). México: Fondo de Cultura Económica.
- HALLIDAY, Michael Alexander Kirkwood (1985): *An Introduction to Functional Grammar*. London: Arnold.
- HARTMANN, Peter y RIESER, Hannes (eds.) (1974): *Angewandte Textlinguistik. Papiere zur Textlinguistik*. Hamburg: Buske.
- HARWEG, Roland (1968): *Pronomina und Textkonstitution*. München: W. Fink Verlag.
- HEINEMANN, Wolfgang y VIEHWEGER, Dieter (1991): *Textlinguistik: eine Einführung*. Tübingen: Niemeyer.
- HELLER, K. (1970): «Der Wortschatz unter dem Aspekt des Fachwortes. Versuch einer Systematik», *Wissenschaftliche Zeitschrift der Karl-Marx-Universität Leipzig. Gesellschafts- und Sprachwissenschaftliche Reihe* 19, pp. 531-544.
- HERNÁNDEZ, Roberto; FERNÁNDEZ, Carlos y BAPTISTA, Pilar: (2015): *Metodología de la investigación*. México: McGraw-Hill.
- HERRERA, Honesto; MARTÍNEZ, Rosario y AMENGUAL, Marian (2011): *Estadística aplicada a la investigación lingüística*. Madrid: Eos, D.L.
- HOFFMANN, Lothar (1990): *Fachtexte und Fachtextsorten*. Leipzig: Universität Leipzig.
- HOFFMANN, Lothar (1998 [1987]): *Llenguatges d'especialitat: selecció de textos* (BRUMME, Jenny, ed.). Barcelona: Institut Universitari de Lingüística Aplicada. Universitat Pompeu Fabra.
- HOFFMANN, Lothar (ed.) (1985): *Kommunikationsmittel Fachsprache: Eine Einführung*. Tübingen: Gunter Narr Verlag.
- HOSMER, David W. y LEMESHOW, Stanley (2000): *Applied Logistic Regression*. New York: Chichester, Wiley.
- HOWATT, Anthony P. R. (1984): *A history of English language teaching*. Oxford: Oxford University Press.
- HYMES, Dell H. (1996): «Acerca de la competencia comunicativa», *Forma y función* 9, pp. 13-37 (Traducción del original en PRIDE, John B. y HOLMES, Janet (eds.) (1972): *Sociolinguistics*).

- Harmondsworth: Penguin, pp. 269-293.
- ICMJE, International Committee of Medical Journal Editors (abril de 2010): *Uniform Requirements for Manuscripts Submitted to Biomedical Journals: Writing and Editing for Biomedical Publication*. Disponible en http://www.icmje.org/urm_full.pdf.
- ISENBERG, Horst (1987): «Cuestiones fundamentales de tipología textual». En: BERNARDEZ, Enrique (ed.): *Lingüística del texto*. Madrid: Arco/Libros.
- JACOBI, Daniel (1987): *Textes et images de la vulgarisation scientifique*. Berne: Peter Lang.
- JAKOBSON, Roman (1970): «La linguistique». En: *Tendances principales de la recherche dans les sciences sociales et humaines. Partie 1; Sciences sociales*. La Haya: Mouton, pp. 504-556.
- JAKOBSON, Roman (1983 [1958]): *Lingüística y poética*. (A. Gutiérrez Cabello, trad.) Madrid: Cátedra.
- JOHNSON, R. Burke y ONWUEGBUZIE, Anthony J. (2004): «Mixed methods research: A research paradigm whose time has come», *Educational Researcher* 33(7), pp. 14- 26.
- KABATEK, Johannes (2013): «¿Es posible una lingüística histórica basada en un corpus representativo?», *Ibero* 77, pp. 8-28.
- KALLMEYER, Werner; KLEIN, Wolfgang y MEYER-HERMANN, Reinhard *et al.* (1974): *Lektürekolleg zur Textlinguistik. Einführung* (vol. 1). *Reader* (vol. 2). Frankfurt: Athenäum-Fischer.
- KAPLAN, Robert B. (ed.) (2002): *The Oxford Handbook of Applied Linguistics*. New York: Oxford University Press.
- KAPLAN, Robert B. (ed.) (2010): *The Oxford Handbook of Applied Linguistics*. Oxford: Oxford University Press.
- KAPLAN, Robert B. y GRABE, William (2000): «Applied linguistics and the *Annual Review of Applied Linguistics*», *Annual Review of Applied Linguistics* 20 (*Applied linguistics as an emerging discipline*, GRABE, William; ed.), pp. 3-17. DOI: <http://dx.doi.org/10.1017/S0267190500299996>.
- KAPLAN, Robert B. y GRABE, William (1992): «Introduction». En: GRABE, William y KAPLAN, Robert B. (eds.): *Introduction to Applied Linguistics*. Reading, Mass. [etc.]: Addison-Wesley.
- KAPLAN, Robert B. y WIDDOWSON, Henry (1992): «Applied Linguistics: Overview». En: BRIGHT, William (ed.): *International Encyclopedia of*

- Linguistics*. New York: Oxford University Press.
- KENNEDY, Graeme (1998): *An Introduction to Corpus Linguistics*. London/New York: Longman.
- KOCOUREK, Rostislav (1991): *La langue française de technique et de la science. Vers une linguistique de la langue savante*. Wiesbaden: Oscar Branstetter.
- KRISTIANSEN, Gitte; RUIZ DE MENDOZA IBÁÑEZ, Francisco J. y FABER, Pamela (eds.) (2012): *Applications of Cognitive Linguistics [ACL]: Cognitive Linguistics View of Terminology and Specialized Language*. München, DEU: Walter de Gruyter.
- LABOV, William (1966): «Contraction, Deletion, and Inherent Variability of the English Copula», *Language* 45, pp. 715-762.
- LABOV, William (1972): *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.
- LACORTE, Manel (2007): «Lingüística aplicada del español. Consideraciones generales». En: LACORTE, Manel (coord.): *Lingüística aplicada del español*. Madrid: Arco/Libros, pp. 17-46.
- LAKOFF, George (1987): *Women, Fire and Dangerous Things*. Chicago: University of Chicago Press.
- LAUFER, Batia (2003): *The influence of L2 on L1 collocational knowledge and on L1 lexical diversity in free written expression*. En: COOK, Vivian (ed.): *Effects of the second language on the first*. Clevedon, UK: Multilingual Matters Ltd., pp. 19-31.
- LAUFFLER-LAURIAN, Anne Marie (1983): «Typologie des discours scientifiques: deux approches», *Études de Linguistique Appliquée* 51, pp. 8-21.
- LAUFFLER-LAURIAN, Anne Marie (1984): «Vulgarisation scientifique: formulation, reformulation, traduction», *Langue Française* 64, pp. 109-125.
- LEECH, Geoffrey (1992): «Corpora and theories of linguistics performance». En: SVARTVIK, Jan (ed.): *Directions in Corpus Linguistics (Proceedings of Nobel Symposium 82, Stockholm, 4-8 august, 1991)*. Berlin-New York: Mouton de Gruyter, pp. 105-147.
- LEWANDOWSKI, Theodor (1982): *Diccionario de Lingüística* (M. L. García-Denche Navarro y E. Bernárdez, trads.). Madrid: Cátedra.
- LÓPEZ RODRÍGUEZ, Clara Inés (2001): «El cotexto como restrictor semántico en textos especializados». En: MUÑOZ, Carmen (ed.): *Trabajos en Lingüística Aplicada*. Barcelona: AESLA, pp. 723-736.
- LÓPEZ SERENA, Araceli y BORREGUERO ZULOAGA, Margarita (2010): «Los marcadores discursivos y la

- variación lengua hablada vs. lengua escrita». En: LOUREDA LAMAS, Ó. y ACÍN VILLA, E. (coords.): *Los estudios sobre marcadores del discurso en español, hoy*. Madrid: Arco, pp. 325-405.
- LOPONEN, Mika (2009): «Creating a Semiotic Framework For the Translation of Fictional Cultures», *Chinese Semiotic Studies* 2(1), pp. 165-175.
- LOTMAN, J. (1990): *Universe of the mind: A semiotic theory of culture*. Bloomington: Indiana University Press.
- LOUREDA LAMAS, Óscar (2002): «Los tipos de texto». En: ROMERO GUALDA, María Victoria (coord.): *Lengua española y comunicación*. Barcelona: Ariel, pp. 155-176.
- LOUREDA LAMAS, Óscar (2003): *Introducción a la tipología textual*. Madrid: Arco/Libros.
- LUNDQUIST, Lita (2003): «How do we know what is known? A text- and psycholinguistic challenge to the study of texts for specific purposes». En: CABRÉ, María Teresa; FREIXA, Judit; TEBÉ, Carles (eds.): *Terminología y conocimiento especializado*. Barcelona: Institut Universitari de Lingüística Aplicada, Universitat Pompeu Fabra, pp. 17-38.
- MACKEY, William F. (1966): «Applied Linguistics: Its Meaning and Use», *English Language Teaching* 20, pp. 197-206.
- MACQUEEN, James B. (1967): «Some Methods for classification and Analysis of Multivariate Observations», *Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability*. Berkeley: University of California Press, pp. 281-297.
- MALVERN, David; RICHARDS, Brian; CHIPERE, Ngoni y DURÁN, Pilar (2004): *Lexical diversity and language development: quantification and assessment*. New York: Palgrave Macmillan.
- MANDELBROT, Benoît (1987 [1975]): *Los objetos fractales: forma, azar y dimensión*. Barcelona: Tusquets.
- MARSLAND, Stephen (2015): *Machine learning: an algorithmic perspective*. Boca Ratón: CRC.
- MARTÍN MINGORANCE, Leocadio (1998): *El modelo lexemático-funcional* (A. Marín Rubiales, ed.). Granada: Editorial Universidad de Granada.
- MARTÍN ZORRAQUINO, M.^a Antonia (1993): «Observaciones sobre propiedades atribuidas al habla femenina en el dominio hispánico». En PENNY, Ralph (ed.): *Actas del Primer Congreso Anglo-Hispano*, vol. 1. Madrid: Castalia, pp. 115-126.
- MARTÍN ZORRAQUINO, M.^a Antonia y PORTOLÉS, José (1999): «Los marcadores del discurso». En: BOSQUE, Ignacio y DEMONTE, Violeta (dirs.) (1999): *Gramática descriptiva de la lengua española*,

- vol 3. Madrid: Espasa Calpe, pp. 4051-4213.
- MARTÍNEZ LINARES, M.^a Antonia (2007): «Sobre la (morfo)sintaxis de las lenguas de especialidad». En: ALCARAZ VARÓ, Enrique; MATEO MARTÍNEZ, José y YUS RAMOS, Francisco (eds.): *Las lenguas profesionales y académicas*. Barcelona: Ariel, pp. 13-25.
- MEYER, Ingrid y MACKINTOSH, Kristen (2000): «When terms move into our everyday lives: an overview of de-terminologization», *Terminology* 6(1), pp. 111-138.
- MEYER, Ingrid; MACKINTOSH, Kristen y VARANTOLA, Krista (1997): «Exploring the reality of virtual: on the lexical implications of becoming a knowledge society», *Lexicology* 3(1), pp. 129-163.
- MEYER, Ingrid; MACKINTOSH, Kristen y VARANTOLA, Krista (1998): «From Virtual Sex to Virtual Dictionaries: On the Analysis and Description of a de-terminologized Word». En: FONTENELLE, Thierry; HILIGSMANN, Philippe; MOULIN, André y THEISSEN, Siegfried (eds.): *Euralex '98 proceedings*. Liège: Euralex, pp. 645-654. Disponible en: http://www.euralex.org/elx_proceedings/Euralex1998_2/Ingrid%20MEYER,%20Kristen%20MACKINTOSH%20&%20Krista%20VARANTOLA%20From%20Virtual%20Sex%20to%20Virtual%20Dictionaries%20On%20t.pdf.
- MIRANDA-GARCÍA, Antonio y CALLE, Javier (2005): «Yule's characteristic k revised», *Language Resources and Evaluation* 39, pp. 287-294.
- MITROFÁNOVA, O. A. (1985): *Nauchnii stil rechi: problemi obuchenia*. Moskvá: Russki Yazik.
- MONTALT, Vicent y SHUTTLEWORTH, Mark (2012): «Research in translation and knowledge mediation in medical and healthcare settings», *Linguistica Antverpiensia New Series – Themes in Translation Studies* 11, pp. 9-30.
- MONTERDE REY, Ana María (2002a): «La asignatura troncal de terminología en las Facultades de Traducción e Interpretación de España: técnicas motivacionales en su enseñanza». En: GUERRERO, Gloria y PÉREZ, Manuel Fernando (eds.): *Panorama actual de la terminología*. Granada: Comares, pp. 197-220.
- MONTERDE REY, Ana María (2002b): *Ejercicios de introducción a la terminología*. Las Palmas de Gran Canaria: Universidad de Las Palmas de Gran Canaria.
- MONTERDE REY, Ana María (2007): «Estudio terminológico en textos aeronáuticos de tres niveles de especialización», *Ibérica: Revista de la Asociación Europea de Lenguas para Fines Específicos (AELFE)* 13, pp. 39-64. Disponible en: <https://dialnet.unirioja.es/descarga/articulo/2474012.pdf>.

- MONTERO, Silvia y FABER, Pamela (2008): *Terminología para traductores e intérpretes*. Granada: Tragacanto.
- MORENO, Francisco (2008): *Principios de sociolingüística y sociología del lenguaje*. Barcelona: Ariel.
- NAZAR, Rogelio y CABRÉ, María Teresa (2011): «Un experimento de extracción de terminología utilizando algoritmos estadísticos supervisados», *Debate Terminológico* 7, pp. 36-55. Disponible en: <http://seer.ufrgs.br/riterm/article/download/20712/11980>.
- NKWENTI-AZEH, Blaise (1998): «Information Mediation: The Interface between Terminology and Translation». En: KENT, Allen (ed.): *Encyclopedia of Library and Information Science* 62(25). New York/Basel/Hong Kong: Marcel Dekker, Inc, pp. 157-170.
- OBNEO – Observatori de Neologia (2004): *Metodología del trabajo en neología: criterios, materiales y procesos*. Barcelona: Institut Universitari de Lingüística Aplicada. Universitat Pompeu Fabra. Disponible en: <http://repositori.upf.edu/bitstream/handle/10230/1307/04mon009.pdf?sequence=1>
- PARODI, Giovanni (2008): «Lingüística de corpus: una introducción al ámbito», *Revista de Lingüística Teórica y Aplicada* 46(1), pp. 93-119.
- PARODI, Giovanni (2010): *Lingüística de corpus: de la teoría a la empiria*. Madrid: Iberoamericana/Vervuert.
- PAYRATÓ, Lluís (1998): *De profesión, lingüista. Panorama de la lingüística aplicada*. Barcelona: Ariel.
- PÉREZ HERNÁNDEZ, Chantal (2002): «Explotación de los corpórea textuales informatizados para la creación de bases de datos terminológicas basadas en el conocimiento», *Estudios de Lingüística del español*, 18. Disponible en: <http://elies.rediris.es/elies18/>.
- PICHT, Heribert (1984): «Lengua general, lengua profesional y terminología». En: FELBER, Helmut y PICHT, Heribert: *Métodos de terminografía y principios de investigación terminológica*. Madrid: Instituto “Miguel de Cervantes”. CSIC.
- PONS BORDERÍA, Salvador (2004): *Conceptos y aplicaciones de la teoría de la relevancia*. Madrid: Arco, D.L.
- PORTOLÉS, José (2004): *Pragmática para hispanistas*. Madrid: Síntesis.
- POTTIER, Bernard (1972 [1967]): *Presentación de la Lingüística: fundamentos de una teoría* (A. Quilis, trad. del original *Présentation de la linguistique. Fondements d'une théorie*. Paris: Klincksieck). Madrid: Ediciones Alcalá.

- PRESTON, Dennis (2013): «The Influence of Regard on Language Variation and Change», *Journal of Pragmatics* 52, pp. 93-104.
- PRIETO VELASCO, Juan Antonio (2008): *Información gráfica y grados de especialización en el discurso científico: un estudio de corpus*. Tesis doctoral. Granada: Universidad de Granada. Disponible en: <http://hera.ugr.es/tesisugr/17491332.pdf>.
- PUSTEJOVSKY, James (1995): *The generative lexicon*. Cambridge: MIT Press.
- PUTNAM, Hilary (1975): «The meaning of meaning». En: PUTNAM, Hilary: *Mind, Language and Reality. Philosophical Papers*, vol. 2. Cambridge: Cambridge University Press, pp. 215-271.
- QUINLAN, John Ross (1993): *C4.5: Programs for Machine Learning*. San Mateo (California) : Morgan Kaufmann.
- REINHARDT, Werner (1966): «Produktive verbale Wortbildungstypen in der Fachsprache der Technik», *Wissenschaftliche Zeitschrift der Pädagogischen Hochschule Potsdam* 2, pp. 183-195.
- REYES, Graciela (2002): *Metapragmática. Lenguaje sobre lenguaje, ficciones y figuras*. Valladolid: Universidad de Valladolid.
- RIGGS, Fred. W. (1988): *The INTERCOCTA Manual. Towards an International Encyclopaedia of Social Science Terms*. Paris: UNESCO. UNESCO, Reports and Papers on the Social Sciences, 58.
- ROBERTO, John A.; MARTÍ, M. Antonia y SALAMÓ, Maria (2012): «Análisis de la riqueza léxica en el contexto de la clasificación de atributos demográficos latentes», *Procesamiento de Lenguaje Natural* 48, pp. 97-104. Disponible en: <http://journal.sepln.org/sepln/ojs/ojs/index.php/pln/article/download/4493/2661>.
- RODRÍGUEZ-TAPIA, Sergio (2015): «Estrategias de traducción inglés-español basadas en el análisis cuantitativo de procedimientos de reformulación formal y conceptual del texto semiespecializado», *Tonos digital: Revista electrónica de estudios filológicos* 29(2), pp. 1-34. Disponible en: <http://www.tonosdigital.com/ojs/index.php/tonos/article/viewFile/1330/805>.
- RODRÍGUEZ-TAPIA, Sergio (2016a): «Los textos especializados, semiespecializados y divulgativos: una propuesta de análisis cualitativo y de clasificación cuantitativa», *Signa: Revista de la Asociación Española de Semiótica* 25, pp. 987-1006. Disponible en: <https://dialnet.unirioja.es/descarga/articulo/5476806.pdf>.

- RODRÍGUEZ-TAPIA, Sergio (2016b): «Clasificación cuantitativa de los textos según su grado de especialización: parámetros para la elaboración de los índices de densidad terminológica y de reformulación de un corpus sobre insuficiencia cardíaca», *Anuario de Estudios Filológicos* 39, pp. 227-250. Disponible en <https://dialnet.unirioja.es/download/articulo/5854830.pdf>.
- RODRÍGUEZ-TAPIA, Sergio y CAMACHO-CAÑAMÓN, Julio (2016): «El aprendizaje automático aplicado a la identificación del grado de especialidad de los textos. Aproximación a un modelo metodológico», *III Jornadas de Iniciación a la Investigación Lingüística*, celebrado en Cáceres del 17 al 18 de noviembre de 2016.
- RODRÍGUEZ-TAPIA, Sergio y CAMACHO-CAÑAMÓN, Julio (2017): «Las ciencias de la computación como apoyo a la clasificación tipológica textual en terminología», *II Congreso Internacional Ciencia y Traducción*, celebrado en Córdoba del 5 al 7 de abril de 2017.
- RODRÍGUEZ-TAPIA, Sergio y CAMACHO-CAÑAMÓN, Julio (2018): «Los métodos de aprendizaje automático supervisado en la clasificación textual según el grado de especialización», *Tonos digital: Revista electrónica de estudios filológicos* 35, pp. 1-28. Disponible en: <http://www.tonosdigital.es/ojs/index.php/tonos/article/view/2009/1018>.
- RODRÍGUEZ-TAPIA, Sergio y MORENO PAZ, María del Carmen (en prensa): «El conocimiento ficcional como forma de acercamiento al conocimiento del mundo real: reclasificación de los tipos de conocimiento, caracterización y fundamentos para un enfoque lingüístico», *Hikma. Revista de estudios de traducción* 17.
- ROJO, Guillermo (2012): «El papel de los corpus en el estudio de la historia del español». En: *Actas del VIII Congreso Internacional de Historia de la Lengua Española*, vol. 1. Santiago de Compostela: Meubook, pp. 433-444.
- ROJO, Guillermo (2014): «Hispanic Corpus Linguistics». En: LACORTE, Manel (ed.): *The Routledge Handbook of Hispanic Applied Linguistics*. New York: Routledge, pp. 371-387.
- ROJO, Guillermo (2015): «Sobre los antecedentes de la lingüística de corpus». En: ÁLVAREZ MENÉNDEZ, Alfredo et al. (eds.): *Studium grammaticae. Homenaje al Profesor José Antonio Martínez*. Oviedo: Universidad de Oviedo, pp. 675-689.
- RONDEAU, Guy (1984): *Introduction à la Terminologie*. Québec: Gaëtan Morin.
- ROSCHE, Eleanor (1978): «Principles of Categorization». En: ROSCHE, Eleanor y LLOYD, Barbara B. (eds.): *Cognition and Categorization*.

- Hillsdale: Lawrence Erlbaum, pp. 27-48.
- ROSCH, Eleanor y MERVIS, Carolyn B. (1975): «Family resemblances: studies in the internal structure of categories», *Cognitive Psychology* 7, pp. 573- 605.
- ROSSARI, Corinne (1990): «Project pour une typologie des opérations de reformulation», *Cahiers de Linguistique Française* 11, pp. 345-359.
- RUIZ GURILLO, Leonor (1998): *La fraseología del español coloquial*. Barcelona: Ariel.
- SÁEZ, Aurora; SÁNCHEZ-MONEDERO, Javier; GUTIÉRREZ, Pedro Antonio y HERVÁS-MARTÍNEZ, César (2016): «Machine Learning Methods for Binary and Multiclass Classification of Melanoma Thickness From Dermoscopic Images», *IEEE transactions on medical imaging* 35(4), pp. 1036-1045.
- SAGER, Juan Carlos (1990): *A Practical Course in Terminology Processing*. Amsterdam/Philadelphia: John Benjamins.
- SAGER, Juan Carlos; DUNGWORTH, David y McDONALD, Peter F. (1980): *English special languages: principles and practice in science and technology*. Wiesbaden: Brandstetter.
- SAN VICENTE, Félix; BAZZOCCHI, Gloria y CAPANAGA, Pilar (eds.) (2017): *Oraliter. Formas de la comunicación presencial y a distancia*. Bologna: Bononia University Press.
- SÁNCHEZ PÉREZ, Aquilino y CANTOS GÓMEZ, Pascual (1998): «El ritmo incremental de palabras nuevas en los repertorios de textos. Estudio experimental y comparativo basado en dos corpus lingüísticos equivalentes de cuatro millones de palabras, de las lenguas inglesa y española y en cinco autores de ambas lenguas», *Atlantis* XIX (2), pp. 205-223.
- SANDIG, Barbara (1972): «Zur Differenzierung gebrauchssprachlicher Textsorten im Deutschen». En: GÜLICH, Elisabeth y RAIBLE, Wolfgang (eds.): *Textsorten. Differenzierungskriterien aus linguistischer Sicht*. Frankfurt: Athenäum, pp. 113-124.
- SANKOFF, David y LABERGE, Suzanne (1978): «The Linguistic market and the Statistical Explanation of Variability». En: SANKOFF, David (ed.): *Linguistic Variation: Models and Methods*. New York: Academic Press, 239-250.
- SAUSSURE, Ferdinand de (2007 [1916]): *Curso de lingüística general*. (A. Alonso, trad.) Buenos Aires: Losada.
- SCHMIDT, Wilhelm (1969): «Charakter und gesellschaftliche Bedeutung der Fachsprachen», *Sprachpflege* 19, pp. 10-21.

- SCHMITT, Norbert y CELCE-MURCIA, Marianne (2002): «An overview of applied linguistics». En: SCHMITT, Norbert (ed.): *An introduction to applied linguistics*. London: Arnold, pp. 1-16.
- SEARLE, John (1969): *Speech Acts: An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press.
- SEARLE, John (1995): *The Construction of Social Reality*. New York: Free Press.
- SEBASTIANI, Fabrizio (2002): «Machine learning in automated text categorization», *ACM computing surveys* 34(1), pp. 1-47. Disponible en: <http://nmis.isti.cnr.it/sebastiani/Publications/ACMCS02.pdf>.
- SHANNON, Claude Elwood (1948): «A Mathematical Theory of Communication», *The Bell System Technical Journal* 27, 379-423.
- SHANNON, Claude Elwood y WEAVER, Warren (1949): *The Mathematical Theory of Communication*. University of Illinois Press: Urbana.
- SINCLAIR, John (1992): «The automatic analysis of corpora». En: SVARTVIK, Jan (ed.): *Directions in Corpus Linguistics (Proceedings of Nobel Symposium 82, Stockholm, 4-8 august, 1991)*. Berlin-New York: Mouton de Gruyter, pp. 379-497.
- SPERBER, Dan y WILSON, Deirdre (1986): *Relevance*. Oxford: Blackwell.
- SPERBER, Dan y WILSON, Deirdre (1995): «Postface». En: *Relevance: communication and cognition*. Oxford: Basil Blackwell, 255-298.
- STREVEN, Peter (1992): «Applied Linguistics: History». En: BRIGHT, William (ed.): *International Encyclopedia of Linguistics*. New York: Oxford University Press.
- STUBBS, Michael (1996): *Text and Corpus Analysis. Computer-assisted Studies of Language and Culture*. Oxford: Blackwell Publishers.
- TEMMERMAN, Rita (2000): *Towards new ways of terminology description: the sociocognitive-approach*. Amsterdam/Philadelphia: John Benjamins.
- TEMMERMAN, Rita (2001): «Sociocognitive Terminology Theory». En: CABRÉ, María Teresa y FELIU, Judit (eds.): *Terminología y cognición*. Barcelona: Institut Universitari de Lingüística Aplicada. Universitat Pompeu Fabra, pp. 75-92.
- TEMMERMAN, Rita (2002): «Sociocognitive Terminology Theory». En: FABER, Pamela y JIMÉNEZ, Catalina (eds.): *Investigar en terminología*. Granada: Método, pp. 75-92.
- TEMPLIN, Mildred C. (1957): *Certain language skills in children: Their*

- development and interrelationships.* Wesport: Greenwood.
- TOGNINI-BONELLI, Elena (2001): *Corpus linguistics at work.* Amsterdam; Philadelphia: John Benjamins.
- TORRUELLA CASAÑAS, Joan (2017): *Lingüística de corpus: génesis y bases metodológicas de los corpus (históricos) para la investigación en lingüística.* Frankfurt am Main: Peter Lang.
- VÁZQUEZ Y DEL ÁRBOL, Esther (2006): *La redacción y traducción biomédica: inglés-español: un estudio basado en 200 textos.* Granada: Universidad de Granada.
- VEGA MORENO, Érika (2015): «Metodología», *La creatividad léxica en el lenguaje de la publicidad. Análisis lingüístico de los mecanismos de producción o interpretación neológica en los sectores de la estética y la peluquería.* Tesis doctoral. Cádiz: Universidad de Cádiz.
- VERSCHUEREN, Jef (1999 [2002]): *Para entender la Pragmática* (E. Baena y M. Lacorte, trads.). Madrid: Gredos.
- VERSCHUEREN, Jef (2000): «Notes on the role of metapragmatic awareness in language», *Pragmatics* 10(4), pp. 439-456.
- VICTORRI, Bernard (1996): «La construction dynamique du sens». En: PORTE, M. (dir.) *Passion des formes.* Sant-Cloud: Éditions de l'ENS de Fontenay-Saint-Cloud, pp. 733-747.
- VILARNOVO CAAMAÑO, Antonio y SÁNCHEZ SÁNCHEZ, José Francisco (1994): *Discurso, tipos de texto y comunicación.* Pamplona: EUNSA.
- VIVALDI, Jorge y RODRÍGUEZ, Horacio (2012): «Using Wikipedia for Domain Terms Extraction». En: GORNOSTAY, T. (ed.): *Proceedings of CHAT 2012: The 2nd Workshop on the Creation, Harmonization and Application of Terminology Resources: collocated with TKE 2012.* Linköping University Electronic Press. Linköpings Universitet, pp. 3-10.
- VIVALDI, Jorge (2001): *Extracción de candidatos a término mediante combinación de estrategias heterogéneas.* Tesis doctoral. Barcelona: Institut Universitari de Lingüística Aplicada. Universitat Pompeu Fabra.
- VIVANCO, Verónica (2006): *El español de la ciencia y la tecnología.* Madrid: Arco Libros.
- WALLACH, Hanna (2004): «Conditional random fields: An introduction». Technical Report MS-CIS-04-21. Pennsylvania: University of Pennsylvania.
- WALTER, Strother H. y DUNCAN, David B. (1967): «Estimation of the probability of an event as a function of several variables», *Biometrika* 54, pp. 167-179.

- WIDDOWSON, Henry (2000a): «Object language and the language subject: on the mediating role of applied linguistics», *Annual Review of Applied Linguistics*, 20 (*Applied linguistics as an emerging discipline*; GRABE, William; ed.), pp. 21-33. DOI: <http://dx.doi.org/10.1017/S0267190500200020>.
- WIDDOWSON, Henry (2000b): «On the limitations of linguistics applied», *Applied Linguistics* 21(1), pp. 3-25. DOI: <http://dx.doi.org/10.1093/app lin/21.1.3>.
- WIDDOWSON, Henry (2004): «A perspective on recent trends». En: HOWATT, Anthony P. R.: *A history of English Language Teaching*. Oxford: Oxford University Press, pp. 353-372.
- WIDDOWSON, Henry (2005): «Applied linguistics, interdisciplinary and disparate realities». En: BRUTHIAUX, Paul; ATKINSON, Dwight; EGGINTON William; GRABE, William y RAMANATHAN, Vaidehi (eds.): *Directions in Applied Linguistics: Essays in Honor of Robert B. Kaplan*. Clevedon, Avon, UK: Multilingual Matters, pp. 3-25.
- WITTEN, Ian H. y FRANK, Eibe (2005): *Data Mining: Practical machine learning tools and techniques*. Amsterdam: Morgan Kaufmann.
- WMFFRE, Iwan (2013): *Dynamic linguistics: Labov, Martinet, Jakobson and other precursors of the dynamic approach to language description*. New York: Peter Lang.
- WRAY, Alison (2000): «Formulaic Sequences in Second Language Teaching: Principle and Practice», *Applied Linguistics* 21(4), pp. 463-489.
- WÜSTER, Eugen (1985): *Einführung in die allgemeine Terminologielehre und terminologische Lexikographie*. Copenhagen: Fachsprachliches Zentrum.
- WÜSTER, Eugen (1998): *Introducción a la teoría general de la terminología y a la lexicografía terminológica* (CABRÉ, María Teresa; ed. y NOKERMAN, Anne-Cécile, trad.). Barcelona: Institut Universitari de Lingüística Aplicada. Universitat Pompeu Fabra.
- YULE, George U. (1944): *The statistical study of literary vocabulary*. Cambridge: Cambridge University Press.
- ZAMORANO AGUILAR, Alfonso (2012): «Teorías del caos y lingüística: aproximación caológica a la comunicación verbal humana», *Signa* 21, 679-705. DOI: <https://doi.org/10.5944/signa.vol21.2012.6323>.
- ZAMORANO AGUILAR, Alfonso (2013): «La terminología como disciplina: aproximación interpretativa a su evolución epistemológica y metodológica a través de la caología», *Moenia: Revista lucense de lingüística & literatura* 19,

pp. 25-43. Disponible en:
[http://www.usc.es/revistas/index.p
hp/moenia/article/download/1933/
1796.](http://www.usc.es/revistas/index.php/moenia/article/download/1933/1796)

VII. ÍNDICE DE ILUSTRACIONES, GRÁFICOS Y TABLAS

1. ÍNDICE DE ILUSTRACIONES, GRÁFICOS Y TABLAS

II. FUNDAMENTOS TEÓRICOS

2. La comunicación como objeto y problema

Ilustración 1: Funciones del lenguaje según Bühler (1979 [1934]: 48) 38

Tabla 1: Perfiles y subperfiles de tipos de interlocutores según su competencia cognitivo-comunicativa 49

3. El texto

Tabla 2: Relaciones opositivas entre los tipos de conversación (Loureda, 2002: 172) 62

4. La Terminología como campo científico

Tabla 3: Características principales de la TGT 66

Tabla 4: Características principales de la TSCT 70

Tabla 5: Características principales de la TbM 71

5. Fundamentos para analizar el término como objeto de estudio

Ilustración 2: Esquema básico unidad léxica especializada-unidad léxica no especializada 79

Ilustración 3: Modelo de Baldinger (1952) 82

Ilustración 4: Modelo de Reindhardt (1966: 185), en Arntz y Picht (1995: 32) 83

Ilustración 5: Modelo de Heller (1970: 532) en Arntz y Picht (1995: 31) 84

Tabla 6: Modelo de Hoffmann (1985) 84

6. El valor especializado (I)

Ilustración 6: Esquema de denominaciones en torno al valor especializado [elaboración propia] 99

Ilustración 7: Esquema de relación entre conocimiento general, conocimiento común y conocimiento especializado 103

Tabla 7: Características del conocimiento común y del conocimiento especializado (adaptado de Cabré et al., 2001: 174 y ss.) 103

7. Relaciones denominación-concepto-situación discursiva

Gráfico 1: Frecuencia absoluta de gripe A entre 2000 y 2015 como banalización efímera según CORPES XXI	114
Gráfico 2: Frecuencia absoluta de ADN entre 2000 y 2015 como banalización perenne según CORPES XXI	116
Gráfico 3: Frecuencia absoluta de wifi entre 2000 y 2015 como banalización perenne según CORPES XXI	117
Ilustración 8: Esquema de relaciones concepto-denominación-discurso	122
Tabla 8: Terminologización de la unidad léxica duro en lingüística	108
Tabla 9: Terminologización de la unidad léxica débil en ciencias de la computación	108
Tabla 10: Transterminologización de la unidad léxica virus en biología y ciencias de la computación	109
Tabla 11: Proceso de terminologización y banalización de la unidad ratón	111
Tabla 12: Banalización de la unidad léxica gripe A en el discurso común	113
Tabla 13: Banalización de la unidad léxica prima de riesgo en el discurso común	114
Tabla 14: Banalización de la unidad léxica ADN en el discurso común	116
Tabla 15: Banalización de la unidad léxica wifi en el discurso común	117
Tabla 16: Significados de trabajo en diferentes discursos especializados	119
Tabla 17: Tipos y subtipos de causas de variación denominativa según Freixa (2013)	120

8. El valor especializado (II): conclusión

Ilustración 9: Esquema de criterios de caracterización terminológica y criterios de terminologicidad	128
Ilustración 10: Ejes X, Z e Y en el principio de activación del valor especializado en el discurso común	133
Ilustración 11: Principio de activación del valor especializado en el discurso común	134
Tabla 18: Frecuencias absolutas de términos en CORPES XXI	131
Tabla 19: Posibles casos de activación del valor especializado según los tipos de texto y los emisores	132
Tabla 20: Leyenda para el principio de activación del valor especializado en el discurso común	133

Tabla 21: Principio de activación del valor especializado en el discurso especializado	134
Tabla 22: Principio de activación del valor especializado en el discurso común	134
Tabla 23: Contraste del principio de activación del valor especializado en el discurso especializado y en el discurso común	135

9. El texto especializado y el grado de especialización

Ilustración 12: Gradación entre polos opuestos según la TCT	145
Ilustración 13: Gradación en y entre polos opuestos según nuestro modelo	147
Tabla 24: Contraste de clasificación del grado de especialización entre la TCT y nuestra propuesta	146
Tabla 25: Características principales del modelo de análisis del texto de Ciapuscio y Kuguel (2002)	150

10. Los procedimientos de reformulación formal en terminología

Tabla 26: Algunos trabajos en español sobre los mecanismos de reformulación (Carmona, 2015: 288)	153
Tabla 27: Marcadores de reformulación parafrástica según el procedimiento que introduce la operación reformulativa (Bach et al., 2003; en Rodríguez-Tapia, 2015: 7)	154
Tabla 28: Marcadores de reformulación parafrástica según la función reformulativa (García Negroni, 2009: 51)	155
Tabla 29: Marcadores de reformulación no parafrástica según la función reformulativa (García Negroni, 2009: 51)	156
Tabla 30: Procedimientos de reformulación formal y conceptual según Rodríguez-Tapia (2015: 14; 2016b: 233)	159
Tabla 31: Herramienta de análisis de procedimientos de reformulación parafrástica	161

III. FUNDAMENTOS METODOLÓGICOS

2. La percepción sociocognitiva del término

Ilustración 14: Esquema del funcionamiento de la percepción	175
---	-----

3. Corpus para analizar el grado de especialización

Tabla 32: Esquema de selección de criterios para compilar el corpus de análisis	183
Tabla 33: Información eliminada de los textos del corpus	190

Tabla 34: Criterios de clasificación y tipos de corpus según Torruella (2017)	191
Tabla 35: Relación de número de textos por tipos textuales en el corpus	194
Tabla 36: Relación provisional de número de palabras y de formas léxicas por tipos textuales en el corpus antes del análisis estadístico	195
Tabla 37: Relación definitiva de número de palabras y de formas léxicas por tipos textuales en el corpus después del análisis estadístico	196

5. Variables de estudio

Tabla 38: Fórmula K de Yule para calcular el índice de diversidad léxica	201
Tabla 39: Fórmula original del índice de densidad terminológica	203
Tabla 40: Ejemplo de texto para cómputo de IDT	204
Tabla 41: Casos posibles en el cómputo de IDT	204
Tabla 42: Cálculo de IDT con número de formas léxicas totales y número de elementos léxicos especializados	206
Tabla 43: Cálculo de IDT con número de palabras totales y número de elementos léxicos especializados	206
Tabla 44: Fórmula del índice de densidad terminológica	206
Tabla 45: Fórmula del índice de reformulación en Rodríguez-Tapia (2016b)	207
Tabla 46: Fórmula del índice de reformulación	208
Tabla 47: Ejemplo de texto para cómputo de IR (1)	209
Tabla 48: Ejemplo de texto para cómputo de IR (2)	209
Tabla 49: Ejemplo del texto MED ART09 sobre procedimiento de reformulación que afecta a término	210
Tabla 50: Ejemplo del texto INF NOT01 sobre procedimiento de reformulación que afecta a término	210
Tabla 51: Ejemplo del texto INF BLOG13 sobre procedimiento de reformulación que afecta a término	210
Tabla 52: Ejemplo del texto MED TES02 sobre procedimiento de reformulación que afecta al término	211
Tabla 53: Ejemplo del texto MED ART09 sobre IR	211
Tabla 54: Estrategias de identificación de la relación emisor-destinatario en los textos del corpus	213

6. Grupos de informantes

Tabla 55: Resumen de grupos de informantes	219
Tabla 56: Ejemplo del texto INF FOR04 sobre selección y clasificación de términos por informantes	223

Tabla 57: Ejemplo del texto INF NOT11 sobre selección y clasificación de términos por informantes	224
Tabla 58: Resumen de información proporcionada por los grupos de informantes	225
Tabla 59: Resumen de información proporcionada por el investigador	227

7. Aplicación de la inteligencia artificial

Ilustración 15: Contribución de las ciencias de la computación a la reflexión teórica de la lingüística (Rodríguez-Tapia y Camacho-Cañamón, 2018)	229
Ilustración 16: Ejemplo de árbol de decisión (Rodríguez-Tapia y Camacho-Cañamón, 2018: 17)	230
Ilustración 17: Ejemplo de agrupamiento en diferente número de clusters	231
Ilustración 18: Esquema de validación cruzada [elaboración propia]	232
Ilustración 19: Esquema de validación cruzada de k iteraciones [elaboración propia]	232
Ilustración 20: Configuración de red bayesiana en WEKA	234
Ilustración 21: Configuración de regresión logística en WEKA	235
Ilustración 22: Configuración de árbol de decisión J48 en WEKA	236
Ilustración 23: Ejemplo de funcionamiento de k-means [elaboración propia]	237
Ilustración 24: Configuración de simple k-means en WEKA	238
Tabla 60: Ejemplo de tabla de decisión	230
Tabla 61: Fórmula del teorema de Bayes	233
Tabla 62: Fórmula del teorema de Bayes aplicado al corpus	233
Tabla 63: Ejemplo de la fórmula del teorema de Bayes para determinar la probabilidad de clasificación de un texto como especializado cuando el emisor es especialista	233
Tabla 64: Centroides aleatorios de partida para k-means en WEKA	239

IV. RESULTADOS Y DISCUSIÓN

1. Análisis correlacional descriptivo global

1.1. Análisis del TEMA

Gráfico 4: Datos globales de clase según tema	244
Gráfico 5: Datos de subemisor para participación en foro sobre informática	246
Gráfico 6: Datos globales de subemisor según tema	247
Gráfico 7: Datos globales de destinatario según tema	248

Gráfico 8: Superposición de datos porcentuales para subemisor y subdestinatario según temas	249
Gráfico 9: Comparación de número de textos por subemisor y subdestinatario según temas	249
Gráfico 10: Relación de datos globales para la correspondencia emisor especialista-tipo de subdestinatario según tema	250
Gráfico 11: Relación de datos globales para la correspondencia emisor instruido-tipo de subdestinatario según tema	251
Gráfico 12: Relación de datos globales para la correspondencia emisor lego-tipo de subdestinatario según tema	253
Gráfico 13: Relación de datos globales para la correspondencia emisor especialista-tipo de clase según tema	254
Gráfico 14: Relación de datos globales para la correspondencia emisor instruido-tipo de clase según tema	255
Gráfico 15: Relación de datos globales para la correspondencia emisor lego-tipo de clase según tema	257
Gráfico 16: Relación de datos globales para la correspondencia destinatario especialista-tipo de clase según tema	259
Gráfico 17: Relación de datos globales para la correspondencia destinatario instruido-tipo de clase según tema	260
Gráfico 18: Relación de datos globales para la correspondencia destinatario lego-tipo de clase según tema	264
Gráfico 19: Datos globales de tipos de FLE en el corpus	265
Gráfico 20: Datos globales de tipos de FLE según tema	266
Gráfico 21: Datos globales de tipos de PRF en el corpus	268
Gráfico 22: Datos globales de tipos de PRF según tema	269
Tabla 65: Promedio de datos numéricos para textos emitidos por legos e instruidos en participación en foro sobre informática	246
Tabla 66: Comparación de media para todos los temas para instruido semiespecialista y lego no formado según subemisor y	250
Tabla 67: Diferencias según temas para textos emitidos por instruido semiespecialista para texto no especializado	255
Tabla 68: Diferencias según temas para textos emitidos por instruido semiespecialista para texto semiespecializado	256
Tabla 69: Diferencias según temas para textos emitidos por instruido lego formado para texto semiespecializado	256
Tabla 70: Diferencias según temas para textos emitidos por instruido lego formado para texto no especializado	257

Tabla 71: Diferencias según temas para textos emitidos por lego no formado para texto no especializado	258
Tabla 72: Diferencias según temas para textos recibidos por destinatario especialista investigador para texto especializado	259
Tabla 73: Diferencias según temas para textos recibidos por destinatario instruido lego formado para texto semiespecializado	261
Tabla 74: Diferencias según temas para tipos de IDT y de IR en textos con destinatario instruido lego formado para texto semiespecializado	262
Tabla 75: Diferencias según temas para textos recibidos por destinatario instruido lego formado para texto no especializado	262
Tabla 76: Diferencias según temas para tipos de IDT y de IR en textos recibidos por destinatario instruido lego formado para texto no especializado	263
Tabla 77: Diferencias según temas para textos recibidos por destinatario lego no formado para texto no especializado	265
Tabla 78: Descripción de los tipos de FLE que distinguen los grupos de informantes	265
Tabla 79: Ejemplos de tipos de términos en derecho	267
Tabla 80: Ejemplos de tipos de términos en informática	267
Tabla 81: Ejemplos de tipos de términos en medicina	268
Tabla 82: Ejemplos de tipos de PRF en derecho	270
Tabla 83: Ejemplos de tipos de PRF en informática	270
Tabla 84: Ejemplos de tipos de PRF en medicina	270

1.2. Análisis de la CLASE

Gráfico 23: Datos globales de tipo de texto según clase	273
Gráfico 24: Datos globales de subemisor según clase	278
Gráfico 25: Datos globales de subdestinatario según clase	279
Gráfico 26: Superposición de datos porcentuales para subemisor y subdestinatario según clases	280
Gráfico 27: Relación de datos globales para la correspondencia emisor especialista-tipo de subdestinatario según clase	281
Gráfico 28: Relación de datos globales para la correspondencia emisor instruido-tipo de subdestinatario según clase	282
Gráfico 29: Relación de datos globales para la correspondencia emisor lego-tipo de subdestinatario según clase	283
Gráfico 30: Datos globales de tipos de FLE según clase	284
Gráfico 31: Suma de FLE a) según tema y clase de texto	285
Gráfico 32: Suma de FLE b) según tema y clase de texto	286

Gráfico 33: Suma de FLE c) según tema y clase de texto	286
Gráfico 34: Datos globales de tipos de PRF según clase	287
Gráfico 35: Suma de PRF a) según tema y clase de texto	287
Gráfico 36: Suma de PRF b) según tema y clase de texto	288
Gráfico 37: Suma de PRF c) según tema y clase de texto	288
Tabla 85: Distribución de tipos textuales según clases	273
Tabla 86: Similitudes entre artículo en revista científica y tesis doctoral especializados	274
Tabla 87: Similitudes entre blog, entrevista y revista o web de divulgación especializada especializados	275
Tabla 88: Similitudes entre blog, entrevista y revista o web de divulgación especializada semiespecializados y no especializados	275
Tabla 89: Similitudes entre artículo en revista científica y tesis doctoral semiespecializados	276
Tabla 90: Similitudes y diferencias entre participación en foro y revista o web de divulgación especializada en las clases de texto semiespecializado y de texto no especializado	277
Tabla 91: Diferencias porcentuales para subemisor y subdestinatario según tipo de usuario y clase	280
Tabla 92: Correspondencias más frecuentes entre subemisor, subdestinatario y tipos textuales según clases	283
Tabla 93: Correspondencias más frecuentes entre subemisor y subdestinatario según clases	284
1.3. Análisis del TIPO DE TEXTO	
Gráfico 38: Datos globales de subemisor según tipo de texto	290
Gráfico 39: Datos globales de subdestinatario según tipo de texto	291
Gráfico 40: Superposición de datos porcentuales para subemisor y subdestinatario (artículo en revista científica)	295
Gráfico 41: Superposición de datos porcentuales para subemisor y subdestinatario (blog)	295
Gráfico 42: Superposición de datos porcentuales para subemisor y subdestinatario (entrevista)	295
Gráfico 43: Superposición de datos porcentuales para subemisor y subdestinatario (noticia)	295
Gráfico 44: Superposición de datos porcentuales para subemisor y subdestinatario (participación en foro)	295

Gráfico 45: Superposición de datos porcentuales para subemisor y subdestinatario (revista o web de divulgación especializada)	295
Gráfico 46: Superposición de datos porcentuales para subemisor y subdestinatario (tesis doctoral)	296
Gráfico 47: Datos comparativos sobre relación subemisor especialista-subdestinatario según temas para el artículo en revista científica	296
Gráfico 48: Datos comparativos sobre relación subemisor instruido-subdestinatario según temas para la entrevista	297
Gráfico 49: Datos comparativos sobre relación subemisor instruido-subdestinatario según temas para la noticia	298
Gráfico 50: Datos comparativos sobre relación subemisor instruido-subdestinatario según temas para la participación en foro	299
Gráfico 51: Datos comparativos sobre relación subemisor especialista-subdestinatario según temas para la tesis doctoral	300
Gráfico 52: Tipos de FLE según temas para el artículo en revista científica	300
Gráfico 53: Tipos de FLE según temas para el blog	300
Gráfico 54: Tipos de FLE según temas para la entrevista	301
Gráfico 55: Tipos de FLE según temas para la noticia	301
Gráfico 56: Tipos de FLE según temas para la participación en foro	301
Gráfico 57: Tipos de FLE según temas para la revista o web de divulgación especializada	301
Gráfico 58: Tipos de FLE según temas para la tesis doctoral	301
Gráfico 59: Tipos de PRF según temas para el artículo en revista científica	305
Gráfico 60: Tipos de PRF según temas para el blog	305
Gráfico 61: Tipos de PRF según temas para la entrevista	305
Gráfico 62: Tipos de PRF según temas para la noticia	305
Gráfico 63: Tipos de Gráfico 65PRF según temas para la participación en foro	305
Gráfico 64: Tipos de PRF según temas para la revista o web de divulgación especializada	305
Gráfico 65: Tipos de PRF según temas para la tesis doctoral	305
Tabla 94: Diferencias según subdestinatario instruido lego formado y lego no formado para blog, entrevista, noticia, participación en foro y revista o web de divulgación especializada	292
Tabla 95: Diferencias de tipo de IDT y tipo de IR según subdestinatario instruido lego formado (LF) y lego no formado (LF) para blog, entrevista, noticia, participación en foro y revista o web de divulgación especializada	293
Tabla 96: Comparación de entrevista de derecho entre semiespecialista-lego formado (DER ENT07) y semiespecialista-lego no formado (DER ENT06)	298

Tabla 97: Ejemplo de participación en foro lego formado-lego formado sobre informática	299
Tabla 98: Ejemplo de FLE en blog sobre derecho	301
Tabla 99: Ejemplo de PRF por expansión en blog sobre informática	304
Tabla 100: Ejemplo de PRF por variación en artículo en revista científica sobre informática	304
Tabla 101: Ejemplo de PRF por variación en artículo en revista científica sobre informática	304
Tabla 102: Características más representativas para tipos textuales	307

1.4. Análisis del ÍNDICE DE DENSIDAD TERMINOLÓGICA

Gráfico 66: Promedio de IDT global y tipos según tema y diferencias con promedio de IDT global y tipos	308
Gráfico 67: Promedio de IDT global y tipos según clase y diferencias con promedio de IDT global y tipos	310
Gráfico 68: Promedio de IDT global y tipos según tipo de texto y diferencias con promedio de IDT global y tipos	311
Gráfico 69: Tipos de IDT según temas para el artículo en revista científica	312
Gráfico 70: Tipos de IDT según temas para el blog	312
Gráfico 71: Tipos de IDT según temas para la entrevista	313
Gráfico 72: Tipos de IDT según temas para la noticia	313
Gráfico 73: Tipos de IDT según temas para la participación en foro	313
Gráfico 74: Tipos de IDT según temas para la revista o web de divulgación especializada	313
Gráfico 75: Tipos de IDT según temas para la tesis doctoral	313
Gráfico 76: Promedio de IDT global y tipos según subemisor y diferencias con promedio de IDT global y tipos	314
Gráfico 77: Promedio de IDT global y tipos según subdestinatario y diferencias con promedio de IDT global y tipos	316
Gráfico 78: Relación de datos globales para la correspondencia subemisor-clase según IDT global y tipos	317
Gráfico 79: Relación de datos globales para la correspondencia subdestinatario-clase según IDT global y tipos	318
Gráfico 80: Relación de datos globales para la correspondencia emisor especialista-tipo de subdestinatario según IDT global y tipos	320
Gráfico 81: Relación de datos globales para la correspondencia emisor instruido-tipo de subdestinatario según IDT global y tipos	322
Gráfico 82: Relación de datos globales para la correspondencia emisor lego-tipo de subdestinatario según IDT global y tipos	322

Tabla 103: Ejemplo de texto con IDT a) elevado	309
Tabla 104: Ejemplo de texto con IDT b) elevado	309
Tabla 105: Ejemplo de texto con IDT c) elevado	309
Tabla 106: Ejemplo de IDT c) elevado en tesis doctoral sobre medicina	315
Tabla 107: Ejemplo de IDT a) elevado en tesis doctoral con destinatario especialista investigador	316
Tabla 108: Ejemplo de IDT a) elevado en blog con destinatario lego no formado	316
Tabla 109: Ejemplo de explicaciones sobre la terminología en una participación en foro no especializada sobre medicina	319
Tabla 110: Ejemplo de FLE a) en entrevista sobre medicina	321
Tabla 111: Ejemplo de FLE a) en entrevista sobre medicina	321

1.5. Análisis del ÍNDICE DE REFORMULACIÓN

Gráfico 83: Promedio de IR global y tipos según tema y diferencias con promedio de IR global y tipos	325
Gráfico 84: Superposición de datos normalizados de tipos de IR y tipos de PRF según tema	326
Gráfico 85: Promedio de IR global y tipos según clase y diferencias con promedio de IR global y tipos	327
Gráfico 86: Promedio de IR global y tipos según tipo de texto y diferencias con promedio de IR global y tipos	328
Gráfico 87: Tipos de IR según temas para el artículo en revista científica	329
Gráfico 88: Tipos de IR según temas para el blog	329
Gráfico 89: Tipos de IR según temas para la entrevista	329
Gráfico 90: Tipos de IR según temas para la noticia	329
Gráfico 91: Tipos de IR según temas para la participación en foro	329
Gráfico 92: Tipos de IR según temas para la revista o web de divulgación especializada	329
Gráfico 93: Tipos de IR según temas para la tesis doctoral	329
Gráfico 94: Promedio de IR global y tipos según subemisor y diferencias con promedio de IR global y tipos	330
Gráfico 95: Promedio de IR global y tipos según subdestinatario y diferencias con promedio de IR global y tipos	331
Gráfico 96: Relación de datos globales para la correspondencia subemisor-clase según IR global y tipos	334
Gráfico 97: Superposición de datos normalizados de IDT global e IR global para subemisor y clase	336

Gráfico 98: Superposición de datos normalizados de IDT global e IR a) para subemisor y clase	337
Gráfico 99: Superposición de datos normalizados de IDT global e IR b) para subemisor y clase	337
Gráfico 100: Superposición de datos normalizados de IDT global e IR c) para subemisor y clase	337
Gráfico 101: Relación de datos globales para la correspondencia subdestinatario-clase según IR global y tipos	339
Gráfico 102: Superposición de datos normalizados de IDT global e IR global para subdestinatario y clase	341
Gráfico 103: Superposición de datos normalizados de IDT global e IR a) para subdestinatario y clase	342
Gráfico 104: Superposición de datos normalizados de IDT global e IR b) para subdestinatario y clase	342
Gráfico 105: Superposición de datos normalizados de IDT global e IR c) para subdestinatario y clase	342
Gráfico 106: Relación de datos globales para la correspondencia emisor especialista-tipo de subdestinatario según IR global y tipos	344
Gráfico 107: Relación de datos globales para la correspondencia emisor instruido-tipo de subdestinatario según IR global y tipos	345
Gráfico 108: Relación de datos globales para la correspondencia emisor lego-tipo de subdestinatario según IR global y tipos	346
Gráfico 109: Superposición de datos normalizados de IDT global e IR global para tipo de texto y clase	347
Gráfico 110: Superposición de datos normalizados de IDT global e IR a) para tipo de texto y clase	348
Gráfico 111: Superposición de datos normalizados de IDT global e IR b) para tipo de texto y clase	348
Gráfico 112: Superposición de datos normalizados de IDT global e IR c) para subdestinatario y clase	348
Tabla 112: Ejemplo de reformulaciones en tesis doctoral sobre derecho con subemisor especialista investigador	331
Tabla 113: Ejemplo de reformulaciones en entrevista sobre informática con subdestinatario lego no formado	331
Tabla 114: Ejemplo de reformulaciones por expansión en artículo en revista científica sobre derecho con subdestinatario especialista investigador	333
Tabla 115: Ejemplo de reformulaciones por variación en tesis doctoral sobre informática con subdestinatario especialista investigador	333
Tabla 116: Relación entre subemisor, clase e IR	335

Tabla 117: Ejemplo de reformulación por reducción en artículo en revista científica sobre informática	344
Tabla 118: Ejemplo de reformulación por expansión de un término críptico en una participación en foro sobre medicina con destinatario lego no formado	346
Tabla 119: Ejemplo de reformulación por expansión en una participación en foro sobre medicina con emisor lego no formado	346
Tabla 120: Ejemplo de reformulación por expansión en una entrevista sobre medicina	347

1.6. Análisis de la VARIEDAD LÉXICA

Gráfico 113: Datos de promedio de TTR estandarizada según tema e IDT	351
Gráfico 114: Datos globales de promedio de TTR estandarizada según tema e IR	352
Gráfico 115: Datos globales de promedio de TTR estandarizada según clase de texto e IDT	352
Gráfico 116: Datos globales de promedio de TTR estandarizada según clase de texto e IR	354
Gráfico 117: Datos globales de promedio de TTR estandarizada según tema e IDT	354
Gráfico 118: Datos globales de promedio de TTR estandarizada según tipo de texto e IR	356
Tabla 121: Ejemplo de texto no especializado con TTR elevada	353
Tabla 122: Ejemplo de texto semiespecializado con TTR elevada	353
Tabla 123: Ejemplo de texto especializado con TTR reducida	353
Tabla 124: Ejemplo de blog con TTR e IDT elevados	355

1.7. Resumen de análisis correlacional descriptivo global

Tabla 125: Resumen global de valores más representativos para cada variable según tema y tipo de texto	357
--	-----

2. Análisis correlacional descriptivo según los informantes

2.1. Análisis de la CLASE

Gráfico 119: Datos de clase de texto según tipo de informante	359
Gráfico 120: Número de textos especializados según grupo de informantes de estudiantes	361
Gráfico 121: Número de textos especializados según grupo de informantes de usuarios medios	361
Gráfico 122: Número de textos especializados	362

Gráfico 123: Número de textos especializados según grupo de informantes de traductores	362
Gráfico 124: Número de textos semiespecializados según grupo de informantes de estudiantes	364
Gráfico 125: Número de textos semiespecializados según grupo de informantes de usuarios medios	364
Gráfico 126: Número de textos semiespecializados según investigador	364
Gráfico 127: Número de textos semiespecializados según grupo de informantes de traductores	364
Gráfico 128: Número de textos no especializados según grupo de informantes de estudiantes	366
Gráfico 129: Número de textos no especializados según grupo de informantes de usuarios medios	366
Gráfico 130: Número de textos no especializados según investigador	366
Gráfico 131: Número de textos no especializados según grupo de informantes de traductores	366
Tabla 126: Diferencias de clasificación de textos especializados entre grupo de informantes de estudiantes y de	359
Tabla 127: IDT global para textos especializados no coincidentes entre grupo de informantes de estudiantes y de	360
Tabla 128: IDT c) para textos especializados no coincidentes entre grupo de informantes de estudiantes y de	360
Tabla 129: Características según grupos de informantes para textos de noticia y participación en foro clasificados como textos especializados por el grupo de informantes de estudiantes	363
Tabla 130: Características de artículos en revista científica semiespecializados según informantes	365
Tabla 131: Tipos de IDT para artículos en revista científica semiespecializados exclusivos del grupo de estudiantes	365
2.2. Análisis del EMISOR y del SUBEMISOR	
Gráfico 132: Datos de emisor según tipo de informante	368
Gráfico 133: Datos de subemisor según tipo de informante	368
Gráfico 134: Número de textos de artículos en revista científica según subemisor según grupo de informantes de estudiantes	369
Gráfico 135: Número de textos de artículos en revista científica según subemisor según grupo de informantes de usuarios medios	369
Gráfico 136: Número de textos de artículos en revista científica según subemisor según investigador	370

Gráfico 137: Número de textos de artículos en revista científica según subemisor según grupo de informantes de traductores	370
Gráfico 138: Número de textos de tesis doctorales según subemisor según grupo de informantes de estudiantes	370
Gráfico 139: Número de textos de tesis doctorales según subemisor según grupo de informantes de usuarios medios	370
Gráfico 140: Número de textos de tesis doctorales según subemisor según investigador	370
Gráfico 141: Número de textos de tesis doctorales según subemisor según grupo de informantes de traductores	370
Gráfico 142: Número de textos de participaciones en foro según subemisor según grupo de informantes de estudiantes	370
Gráfico 143: Número de textos de participaciones en foro según subemisor según grupo de informantes de usuarios medios	
Gráfico 143: Número de textos de participaciones en foro según subemisor según grupo de informantes de usuarios medios	370
Gráfico 144: Número de textos de participaciones en foro según subemisor según investigador	371
Gráfico 145: Número de textos de participaciones en foro según subemisor según grupo de informantes de traductores	371
Gráfico 146: Número de textos de noticias según subemisor según grupo de informantes de estudiantes	371
Gráfico 147: Número de textos de noticias según subemisor según grupo de informantes de usuarios medios	371
Gráfico 148: Número de textos de noticias según subemisor según investigador	371
Gráfico 149: Número de textos de noticias según subemisor según grupo de informantes de traductores	371
Gráfico 150: Número de textos de blogs según subemisor según grupo de informantes de estudiantes	372
Gráfico 151: Número de textos de blogs según subemisor según grupo de informantes de usuarios medios	372
Gráfico 152: Número de textos de blogs según subemisor según investigador	372
Gráfico 153: Número de textos de blogs según subemisor según grupo de informantes de traductores	
Gráfico 153: Número de textos de blogs según subemisor según grupo de informantes de traductores	372
Gráfico 154: Número de textos de entrevistas según subemisor según grupo de informantes de estudiantes	372
Gráfico 155: Número de textos de entrevistas según subemisor según grupo de informantes de usuarios medios	372
Gráfico 156: Número de textos de entrevistas según subemisor según investigador	372

Gráfico 157: Número de textos de entrevistas según subemisor según grupo de informantes de traductores	372
Gráfico 158: Número de textos de revistas o web de divulgación especializada según subemisor según grupo de informantes de estudiantes	373
Gráfico 159: Número de textos de revistas o web de divulgación especializada según subemisor según grupo de informantes de usuarios medios	373
Gráfico 160: Número de textos de revistas o web de divulgación especializada según subemisor según investigador	373
Gráfico 161: Número de textos de revistas o web de divulgación especializada según subemisor según grupo de informantes de traductores	373

2.3. Análisis del DESTINATARIO y del SUB

Gráfico 162: Datos de destinatario según tipo de informante	374
Gráfico 163: Datos de subdestinatario según tipo de informante	374
Gráfico 164: Número de textos de artículos en revista científica según subdestinatario según grupo de informantes de estudiantes	375
Gráfico 165: Número de textos de artículos en revista científica según subdestinatario según grupo de informantes de usuarios medios	375
Gráfico 166: Número de textos de artículos en revista científica según subdestinatario según investigador	375
Gráfico 167: Número de textos de artículos en revista científica según subdestinatario según grupo de informantes de traductores	375
Gráfico 168: Número de textos de tesis doctorales según subdestinatario según grupo de informantes de estudiantes	375
Gráfico 169: Número de textos de tesis doctorales según subdestinatario según grupo de informantes de usuarios medios	375
Gráfico 170: Número de textos de tesis doctorales según subdestinatario según investigador	376
Gráfico 171: Número de textos de tesis doctorales según subdestinatario según grupo de informantes de traductores	376
Gráfico 172: Número de textos de participaciones en foro según subdestinatario según grupo de informantes de estudiantes	376
Gráfico 173: Número de textos de participaciones en foro según subdestinatario según grupo de informantes de usuarios medios	376
Gráfico 174: Número de textos de participaciones en foro según subdestinatario según investigador	376
Gráfico 175: Número de textos de participaciones en foro según subdestinatario según grupo de informantes de traductores	376
Gráfico 176: Número de textos de noticias según subdestinatario según grupo de informantes de estudiantes	377

Gráfico 177: Número de textos de noticias según subdestinatario según grupo de informantes de usuarios medios	377
Gráfico 178: Número de textos de noticias según subdestinatario según investigador	377
Gráfico 179: Número de textos de noticias según subdestinatario según grupo de informantes de traductores	377
Gráfico 180: Número de textos de entrevistas según subdestinatario según grupo de informantes de estudiantes	377
Gráfico 181: Número de textos de entrevistas según subdestinatario según grupo de informantes de usuarios medios	377
Gráfico 182: Número de textos de entrevistas según subdestinatario según investigador	378
Gráfico 183: Número de textos de entrevistas según subdestinatario según grupo de informantes de traductores	378
Gráfico 184: Número de textos de revistas o web de divulgación especializada según subdestinatario según grupo de informantes de estudiantes	378
Gráfico 185: Número de textos de revistas o web de divulgación especializada según subdestinatario según grupo de informantes de usuarios medios	378
Gráfico 186: Número de textos de revistas o web de divulgación especializada según subdestinatario según investigador	378
Gráfico 187: Número de textos de revistas o web de divulgación especializada según subdestinatario según grupo de informantes de traductores	378
Gráfico 188: Número de textos de blogs según subdestinatario según grupo de informantes de estudiantes	378
Gráfico 189: Número de textos de blogs según subdestinatario según grupo de informantes de usuarios medios	378
Gráfico 190: Número de textos de blogs según subdestinatario según investigador	379
Gráfico 191: Número de textos de blogs según subdestinatario según grupo de informantes de traductores	379

2.4. Análisis de las FORMAS LÉXICAS ESPECIALIZADAS

Gráfico 192: Datos de FLE a) según tipo de informante	379
Gráfico 193: Datos de FLE b) según tipo de informante	380
Gráfico 194: Datos de FLE c) según tipo de informante	382
Gráfico 195: Número de tipos de FLE según tipo de texto según tipo de informante	385
Tabla 132: Ejemplo de texto con FLE a) según grupo de estudiantes	379
Tabla 133: Ejemplo de texto con FLE a) según grupo de usuarios medios	380

Tabla 134: Ejemplo de texto con FLE b) según grupo de estudiantes	381
Tabla 135: Ejemplo de texto con FLE b) según grupo de usuarios medios	381
Tabla 136: Ejemplo de texto con FLE b) según grupo de estudiantes	381
Tabla 137: Ejemplo de texto con FLE b) según grupo de usuarios medios	382
Tabla 138: Ejemplo de texto con FLE c) según grupo de estudiantes	382
Tabla 139: Ejemplo de texto con FLE c) según grupo de usuarios medios	383
Tabla 140: Ejemplo de texto con FLE a) según	383
Tabla 141: Ejemplo de texto con FLE c) según grupo de traductores	383

2.5. Análisis del ÍNDICE DE DENSIDAD TERMINOLÓGICA

Gráfico 196: Promedio de IDT global según tipo de informante y diferencias con promedio de IDT global	386
Gráfico 197: Promedio de IDT a) según tipo de informante y diferencias con promedio de IDT a)	386
Gráfico 198: Promedio de IDT b) según tipo de informante y diferencias con promedio de IDT b)	387
Gráfico 199: Promedio de IDT c) según tipo de informante y diferencias con promedio de IDT c)	387
Gráfico 200: Promedio de tipos de IDT según informantes según temas y diferencia con promedios de IDT	387
Gráfico 201: Promedio de IDT global para artículo en revista científica según informantes y diferencia con promedio de IDT global	388
Gráfico 202: Promedio de IDT a) para artículo en revista científica según informantes y diferencia con promedio de IDT a)	388
Gráfico 203: Promedio de IDT b) para artículo en revista científica según informantes y diferencia con promedio de IDT b)	388
Gráfico 204: Promedio de IDT c) para artículo en revista científica según informantes y diferencia con promedio de IDT c)	388
Gráfico 205: Promedio de IDT global para tesis doctoral según informantes y diferencia con promedio de IDT global	389
Gráfico 206: Promedio de IDT a) para tesis doctoral según informantes y diferencia con promedio de IDT a)	389
Gráfico 207: Promedio de IDT b) para tesis doctoral según informantes y diferencia con promedio de IDT b)	389
Gráfico 208: Promedio de IDT c) para tesis doctoral según informantes y diferencia con promedio de IDT c)	389
Gráfico 209: Promedio de IDT global para blog según informantes y diferencia con promedio de IDT global	389

Gráfico 210: Promedio de IDT a) para blog según informantes y diferencia con promedio de IDT a)	389
Gráfico 211: Promedio de IDT b) para blog según informantes y diferencia con promedio de IDT b)	390
Gráfico 212: Promedio de IDT c) para blog según informantes y diferencia con promedio de IDT c)	390
Gráfico 213: Promedio de IDT global para entrevista según informantes y diferencia con promedio de IDT global	390
Gráfico 214: Promedio de IDT a) para entrevista según informantes y diferencia con promedio de IDT a)	390
Gráfico 215: Promedio de IDT b) para entrevista según informantes y diferencia con promedio de IDT b)	390
Gráfico 216: Promedio de IDT c) para entrevista según informantes y diferencia con promedio de IDT c)	390
Gráfico 217: Promedio de IDT global para noticia según informantes y diferencia con promedio de IDT global	391
Gráfico 218: Promedio de IDT a) para noticia según informantes y diferencia con promedio de IDT a)	391
Gráfico 219: Promedio de IDT b) para noticia según informantes y diferencia con promedio de IDT b)	391
Gráfico 220: Promedio de IDT c) para noticia según informantes y diferencia con promedio de IDT c)	391
Gráfico 221: Promedio de IDT global para participación en foro según informantes y diferencia con promedio de IDT global	391
Gráfico 222: Promedio de IDT a) para participación en foro según informantes y diferencia con promedio de IDT a)	391
Gráfico 223: Promedio de IDT b) para participación en foro según informantes y diferencia con promedio de IDT b)	392
Gráfico 224: Promedio de IDT c) para participación en foro según informantes y diferencia con promedio de IDT c)	392
Gráfico 225: Promedio de IDT global para revista o web de divulgación especializada según informantes y diferencia con promedio de IDT global	392
Gráfico 226: Promedio de IDT a) para revista o web de divulgación especializada según informantes y diferencia con promedio de IDT a)	392
Gráfico 227: Promedio de IDT b) para revista o web de divulgación especializada según informantes y diferencia con promedio de IDT b)	392
Gráfico 228: Promedio de IDT c) para revista o web de divulgación especializada según informantes y diferencia con promedio de IDT c)	392
Gráfico 229: Promedio de tipos de IDT según informantes según clases y diferencia con promedios de IDT	393

Gráfico 230: Promedio de IDT global para subemisor especialista investigador según informantes y diferencia con promedio de IDT global	394
Gráfico 231: Promedio de IDT a) para subemisor especialista investigador según informantes y diferencia con promedio de IDT a)	394
Gráfico 232: Promedio de IDT b) para subemisor especialista investigador según informantes y diferencia con promedio de IDT b)	394
Gráfico 233: Promedio de IDT c) para subemisor especialista investigador según informantes y diferencia con promedio de IDT c)	394
Gráfico 234: Promedio de IDT global para subemisor especialista no investigador según informantes y diferencia con promedio de IDT global	395
Gráfico 235: Promedio de IDT a) para subemisor especialista no investigador según informantes y diferencia con promedio de IDT a)	395
Gráfico 236: Promedio de IDT b) para subemisor especialista no investigador según informantes y diferencia con promedio de IDT b)	395
Gráfico 237: Promedio de IDT c) para subemisor especialista no investigador según informantes y diferencia con promedio de IDT c)	395
Gráfico 238: Promedio de IDT global para subemisor instruido semiespecialista según informantes y diferencia con promedio de IDT global	395
Gráfico 239: Promedio de IDT a) para subemisor instruido semiespecialista según informantes y diferencia con promedio de IDT a)	395
Gráfico 240: Promedio de IDT b) para subemisor instruido semiespecialista según informantes y diferencia con promedio de IDT b)	396
Gráfico 241: Promedio de IDT c) para subemisor instruido semiespecialista según informantes y diferencia con promedio de IDT c)	396
Gráfico 242: Promedio de IDT global para subemisor instruido aprendiz de especialista según informantes y diferencia con promedio de IDT global	396
Gráfico 243: Promedio de IDT a) para subemisor instruido aprendiz de especialista según informantes y diferencia con promedio de IDT a)	396
Gráfico 244: Promedio de IDT b) para subemisor instruido aprendiz de especialista según informantes y diferencia con promedio de IDT b)	396
Gráfico 245: Promedio de IDT c) para subemisor instruido aprendiz de especialista según informantes y diferencia con promedio de IDT c)	396
Gráfico 246: Promedio de IDT global para subemisor instruido lego formado según informantes y diferencia con promedio de IDT global	397
Gráfico 247: Promedio de IDT a) para subemisor instruido lego formado según informantes y diferencia con promedio de IDT a)	397
Gráfico 248: Promedio de IDT b) para subemisor instruido lego formado según informantes y diferencia con promedio de IDT b)	397
Gráfico 249: Promedio de IDT c) para subemisor instruido lego formado según informantes y diferencia con promedio de IDT c)	397

Gráfico 250: Promedio de IDT global para subemisor lego no formado según informantes y diferencia con promedio de IDT global	397
Gráfico 251: Promedio de IDT a) para subemisor lego no formado según informantes y diferencia con promedio de IDT a)	397
Gráfico 252: Promedio de IDT b) para subemisor lego no formado según informantes y diferencia con promedio de IDT b)	398
Gráfico 253: Promedio de IDT c) para subemisor lego no formado según informantes y diferencia con promedio de IDT c)	398
Gráfico 254: Promedio de IDT global para subdestinatario especialista investigador según informantes y diferencia con promedio de IDT global	398
Gráfico 255: Promedio de IDT a) para subdestinatario especialista investigador según informantes y diferencia con promedio de IDT a)	398
Gráfico 256: Promedio de IDT b) para subdestinatario especialista investigador según informantes y diferencia con promedio de IDT b)	399
Gráfico 257: Promedio de IDT c) para subdestinatario especialista investigador según informantes y diferencia con promedio de IDT c)	399
Gráfico 258: Promedio de IDT global para subdestinatario especialista no investigador según informantes y diferencia con promedio de IDT global	399
Gráfico 259: Promedio de IDT a) para subdestinatario especialista no investigador según informantes y diferencia con promedio de IDT a)	399
Gráfico 260: Promedio de IDT b) para subdestinatario especialista no investigador según informantes y diferencia con promedio de IDT b)	399
Gráfico 261: Promedio de IDT c) para subdestinatario especialista no investigador según informantes y diferencia con promedio de IDT c)	399
Gráfico 262: Promedio de IDT global para subdestinatario instruido semiespecialista según informantes y diferencia con promedio de IDT global	400
Gráfico 263: Promedio de IDT a) para subdestinatario instruido semiespecialista según informantes y diferencia con promedio de IDT a)	400
Gráfico 264: Promedio de IDT b) para subdestinatario instruido semiespecialista según informantes y diferencia con promedio de IDT b)	400
Gráfico 265: Promedio de IDT c) para subdestinatario instruido semiespecialista según informantes y diferencia con promedio de IDT c)	400
Gráfico 266: Promedio de IDT global para subdestinatario instruido aprendiz de especialista según informantes y diferencia con promedio de IDT global	400
Gráfico 267: Promedio de IDT a) para subdestinatario instruido aprendiz de especialista según informantes y diferencia con promedio de IDT a)	400
Gráfico 268: Promedio de IDT b) para subdestinatario instruido aprendiz de especialista según informantes y diferencia con promedio de IDT b)	401
Gráfico 269: Promedio de IDT c) para subdestinatario instruido aprendiz de especialista según informantes y diferencia con promedio de IDT c)	401

Gráfico 270: Promedio de IDT global para subdestinatario instruido lego formado según informantes y diferencia con promedio de IDT global	401
Gráfico 271: Promedio de IDT a) para subdestinatario instruido lego formado según informantes y diferencia con promedio de IDT a)	401
Gráfico 272: Promedio de IDT b) para subdestinatario instruido lego formado según informantes y diferencia con promedio de IDT b)	401
Gráfico 273: Promedio de IDT c) para subdestinatario instruido lego formado según informantes y diferencia con promedio de IDT c)	401
Gráfico 274: Promedio de IDT global para subdestinatario lego no formado según informantes y diferencia con promedio de IDT global	402
Gráfico 275: Promedio de IDT a) para subdestinatario lego no formado según informantes y diferencia con promedio de IDT a)	402
Gráfico 276: Promedio de IDT b) para subdestinatario lego no formado según informantes y diferencia con promedio de IDT b)	402
Gráfico 277: Promedio de IDT c) para subdestinatario lego no formado según informantes y diferencia con promedio de IDT c)	402

2.7. Resumen de análisis correlacional descriptivo según los informantes

Tabla 142: Resumen global de valores más representativos para cada variable según tema y tipo de texto según grupo de estudiantes	405
Tabla 143: Resumen global de valores más representativos para cada variable según tema y tipo de texto según grupo de usuarios medios	406
Tabla 144: Resumen global de valores más representativos para cada variable según tema y tipo de texto según grupo de	407
Tabla 145: Resumen global de valores más representativos para cada variable según tema y tipo de texto según grupo de traductores	408

3. Análisis correlacional a partir de aprendizaje automático

3.1. Probabilidades de clasificación en clases según variables

Ilustración 25: Distribución de patrones por TTR según algoritmo de red bayesiana	411
Ilustración 26: Distribución de patrones por IDT a) según algoritmo de red bayesiana	412
Tabla 146: Probabilidades de clasificación de clase según datos globales	413
Tabla 147: Probabilidades de clase según emisor y destinatario en el corpus	414
Tabla 148: Probabilidades de clasificación de clase según datos de estudiantes	415
Tabla 149: Probabilidades de clasificación de clase según datos de usuarios medios	416

Tabla 150: Probabilidades de clasificación de clase según datos de traductores	417
Tabla 151: Probabilidades de clasificación de clase según datos del	418
Tabla 152: Diferencias para menor y mayor probabilidad de tipo textual según clase según informantes	419
Tabla 153: Diferencias para menor y mayor probabilidad de clase según subemisor según informantes	419

3.2. Relevancia de los atributos

Tabla 154: Relevancia de los valores de los atributos en la discriminación del texto especializado y no especializado según datos globales	421
Tabla 155: Contraste de la relevancia de los valores de los atributos en la discriminación del texto especializado y no especializado según datos globales	422
Tabla 156: Porcentaje de textos especializados sobre el total de textos emitidos por instruido aprendiz de especialista o instruido semiespecialista	422
Tabla 157: Relevancia de los valores de los atributos en la discriminación del texto especializado y no especializado según informantes estudiantes	423
Tabla 158: Relevancia de los valores de los atributos en la discriminación del texto especializado y no especializado según informantes usuarios medios	424
Tabla 159: Relevancia de los valores de los atributos en la discriminación del texto especializado y no especializado según informantes traductores	425
Tabla 160: Relevancia de los valores de los atributos en la discriminación del texto especializado y no especializado según	426
Tabla 161: Contraste de la relevancia de los valores de los atributos en la discriminación del texto especializado y no especializado según informantes estudiantes	427
Tabla 162: Contraste de la relevancia de los valores de los atributos en la discriminación del texto especializado y no especializado según informantes usuarios medios	427
Tabla 163: Contraste de la relevancia de los valores de los atributos en la discriminación del texto especializado y no especializado según informantes traductores	427
Tabla 164: Contraste de la relevancia de los valores de los atributos en la discriminación del texto especializado y no especializado según	428

3.3. Árboles de decisión

Ilustración 27: Árbol de decisión de clases según atributos [0]	430
Ilustración 28: Árbol de decisión de clases según atributos [1]	431
Ilustración 29: Árbol de decisión de clases según atributos [2]	432
Ilustración 30: Árbol de decisión de clases según atributos [3]	433

Ilustración 31: Árbol de decisión de clases según atributos [4]	434
Ilustración 32: Árbol de decisión de clases según atributos [5]	435
Ilustración 33: Árbol de decisión de clases según atributos [6]	436
Ilustración 34: Árbol de decisión de clases según atributos [7]	437

3.4. Grupos identificados según el aprendizaje automático

Ilustración 35: Agrupamiento de patrones por clase según algoritmo k-means	441
Ilustración 36: Agrupamiento de patrones por tema según algoritmo k-means	441
Ilustración 37: Agrupamiento de patrones por TTR estandarizada según algoritmo k-means	442
Ilustración 38: Agrupamiento de patrones por tipo de texto según algoritmo k-means	442
Ilustración 39: Agrupamiento de patrones por IDT global según algoritmo k-means	443
Ilustración 40: Agrupamiento de patrones por IDT a) según algoritmo k-means	443
Ilustración 41: Agrupamiento de patrones por IDT b) según algoritmo k-means	444
Ilustración 42: Agrupamiento de patrones por IDT c) según algoritmo k-means	444
Ilustración 43: Agrupamiento de patrones por IR global según algoritmo k-means	445
Ilustración 44: Agrupamiento de patrones por IR a) según algoritmo k-means	445
Ilustración 45: Agrupamiento de patrones por IR b) según algoritmo k-means	446
Ilustración 46: Agrupamiento de patrones por IR c) según algoritmo k-means	446
Ilustración 47: Agrupamiento de patrones por emisor según algoritmo k-means	447
Ilustración 48: Agrupamiento de patrones por destinatario según algoritmo k-means	447
Ilustración 49: Agrupamiento de patrones por subemisor según algoritmo k-means	448
Ilustración 50: Agrupamiento de patrones por subdestinatario según algoritmo k-means	449
Tabla 165: Centroides para los tres grupos de clasificación según el algoritmo k-means	439
Tabla 166: Contraste de porcentajes de representación de TE, TSE y TNE según el algoritmo k-means y los datos del corpus	439
Tabla 167: Probabilidades de clasificación de emisor y destinatario según clase según datos globales	447
Tabla 168: Probabilidades de clasificación de subemisor y subdestinatario según clase según datos globales	449

V. SUMMARIZATION AND CONCLUSIONS

2. Conclusions

Ilustración 51: Gradación en y entre polos opuestos según nuestro model 458