

VARIACIÓN LÉXICA Y TEXTO: ANÁLISIS CUALITATIVO DEL GRADO DE FORMALIDAD EN CONTEXTO CON AQUAD

Fernández Smith, Gérard

*Universidad de Cádiz
Facultad de Filosofía y Letras
Avda. Gómez Ulla, nº 1, 11003 Cádiz (España)
Tel: (+34) 956015579
e-mail: gerard.fernandez@uca.es*

Escoriza Morera, Luis

*Universidad de Cádiz
Facultad de Filosofía y Letras
Avda. Gómez Ulla, nº 1, 11003 Cádiz (España)
Tel: (+34) 956015587
e-mail: luis.escoriza@uca.es*

(Recibido septiembre 2004; aceptado septiembre 2004)

BIBLID [1133-682X (2004) 12; 73-89]

Resumen

En este trabajo presentamos las pautas y los resultados de una investigación en la que hemos pretendido mostrar la relación que existe entre las variantes léxicas y los estándares textuales en que se emplean. Nuestra hipótesis previa establece que determinados textos, que podríamos calificar de estándares en cuanto que se caracterizan porque manifiestan un determinado grado de formalidad, favorecen la selección de variantes léxicas de acuerdo con un principio comunicativo de adecuación. Dicho principio es observable como un factor que contribuye a la competencia comunicativa de los hablantes de una lengua dada. Esta investigación se ha basado en un conjunto de encuestas en las que los sujetos debían seleccionar variantes léxicas en parejas de unidades, diferenciadas *a priori* como 'formales' o 'informales'. Para su procesamiento y posterior interpretación nos hemos servido de la herramienta AQUAD, un programa informático de análisis de datos cualitativos.

Palabras clave: variación léxica, texto, contexto, grado de formalidad, adecuación, análisis cualitativo y cuantitativo, AQUAD.

Abstract

In the following paper results are set forward of an investigation intended to demonstrate that there is a relation between lexical variants and the textual forms in which they are used. Our previous hypothesis is that certain texts, considered standards because they register a certain degree of formality, favour the selection of lexical variants in order to accomplish a communicative principle of appropriateness. Such principle can be seen as a contribution to the communicative competence of speakers. This research project has been based in a series of surveys in which individuals had to choose the appropriate variants within lexical pairs previously labeled as 'formal' or 'informal'. In order to process and interpret this data we have used the tool AQUAD, a computer program for qualitative data analysis.

Keywords: lexical variation, text, context, degree of formality, appropriateness, qualitative and quantitative analysis, AQUAD.

Résumé

Dans cette étude nous présentons les étapes et les résultats de notre recherche qui prétend montrer la relation existante entre les variantes lexicales et les standards textuels dans lesquels elles sont employées. Notre hypothèse préalable établit que certains textes, que nous pouvons qualifier comme standards étant donné qu'ils manifestent un certain degré de formalité, favorisent la sélection de variantes lexicales conformément à un principe communicatif d'adéquation. Ce principe est considéré comme un facteur qui contribue à la compétence communicative des locuteurs d'une langue donnée. Notre recherche s'est fondée sur un ensemble d'enquêtes dans lesquelles les sujets devaient sélectionner des variantes lexicales par paires d'unités, différenciées *a priori* comme *formelles* ou *informelles*. Nous avons employé L'AQUAD (un programme informatique d'analyse de données qualitatives) pour interpréter les résultats des enquêtes.

Mots-clés: Variante lexicale, texte, contexte, degré de formalité, adéquation, analyse qualitative et quantitative, AQUAD.

Sumario

0. Introducción. 1. Léxico y variación. 2. Texto y contexto. 3. Aspectos metodológicos. 3.1. Diseño de la prueba y recogida de datos. 3.2. Procesamiento de los datos. Codificación con AQUAD. 4. Análisis cualitativo y cuantitativo e interpretación de los datos. 5. Conclusiones. 6. Referencias.

0. Introducción

El estudio de elementos léxicos y textos ha constituido un motivo de frecuente preocupación para los estudiosos del lenguaje desde las primeras reflexiones en torno al mismo. Además de numerosas investigaciones que de manera aislada han contribuido a ahondar en la caracterización tanto de unidades léxicas como de unidades textuales, no son pocos los autores que, al igual que en esta investigación, han analizado la mutua influencia entre la elección de palabras y la construcción de textos. En las siguientes páginas intentaremos abordar algunas consideraciones de interés surgidas en el estudio de léxico y texto para, una vez expuestos los principales aspectos metodológicos desarrollados en nuestra investigación, mostrar algunos de los resultados y conclusiones más relevantes en nuestro particular acercamiento al funcionamiento de esas dos unidades de análisis lingüístico.

1. Léxico y variación

A pesar de los problemas con los que podemos encontrarnos para describir el concepto de palabra y de las múltiples teorías (en las que no entraremos aquí, por no ser ese el objeto de este trabajo) surgidas fundamentalmente en el ámbito de la morfología y la lexicología para sustentar la existencia de dicha unidad en las distintas lenguas naturales, no es difícil encontrar en multitud de lenguas una serie considerable de estudios, a los que ampara nuestra intuición como hablantes, en torno al léxico. Ese interés en el léxico, plasmado, por ejemplo, en la lexicografía, de la que existen testimonios en casi todos los sistemas lingüísticos relativamente cercanos a la existencia de escritura, engloba aspectos tan diversos como la delimitación o *enumeración* de las diferentes palabras de una lengua, mediante la confección de listas; la *definición* de esas palabras, a través de la extracción de sus rasgos distintivos; la búsqueda de *equivalencia* entre palabras de diferentes lenguas; la explicación de la *organización* y relación de las palabras de una misma lengua (mediante el establecimiento

to de campos o el análisis de las relaciones léxicas), o la caracterización del *uso* que de esas palabras realiza cualquier comunidad hablante. En torno a esta última idea, la preocupación acerca de cuándo y por qué se usan unas palabras y no otras, se articulan algunas de las opiniones vertidas en este trabajo.

El hecho de que el hombre sea un animal social y el lenguaje humano uno de los rasgos que caracterizan la especificidad de la especie humana, deben bastar para comprender el interés, desde las primeras reflexiones en torno al lenguaje, en vincular las lenguas a las sociedades que las hablan. Más allá de esa relación de carácter general entre lengua y sociedad, que ha aportado diversos trabajos de indudable interés para la lingüística, desde las últimas décadas del siglo pasado, y ya en el seno de una lingüística constituida como ciencia y considerada como tal, se menciona a la sociolingüística como disciplina lingüística preocupada por los aspectos sociales de las lenguas. Entre la enorme multitud de estudios, trabajos y líneas de investigación situados en el seno de esa disciplina sociolingüística sobresalen, a nuestro juicio, dos grandes campos de investigación: el *contacto lingüístico e intercultural*, consecuencia del multilingüismo o diversidad lingüística, esto es, de la existencia de diferentes sistemas en el seno del lenguaje humano y las relaciones entre los pueblos que los hablan, y el denominado *variacionismo*, que se preocupa de la existencia de una variación intralingüística, dentro de cada sistema, motivada por su riqueza y por su propio carácter, insistimos, de instrumento social.

La sociolingüística variacionista parte de la idea, común en el ámbito de la lingüística, de la existencia de una serie de unidades lingüísticas (invariantes) que subyacen a multitud de realizaciones de las mismas (variantes), las cuales difieren algo de su modelo (eso las caracteriza) pero nunca lo suficiente para dejar de ser identificadas con el mismo. A partir de ese hecho la sociolingüística analizaría qué factores (fundamentalmente sociales) son los que provocan o motivan la aparición de esas distintas realizaciones o variantes. Es lo que ocurre con claridad en el ámbito fonético-fonológico, en el que la sociolingüística ha tratado y trata de investigar si cualquier realización (variante) de cualquier fonema (invariante) se halla vinculada en una comunidad de habla determinada a aspectos no estrictamente lingüísticos como la edad de sus hablantes (puede ser una marca de la generación joven o una realización característica de los mayores, en desuso), el sexo, la procedencia de un segmento de población, la pertenencia a un estrato sociocultural concreto, etc.

Si aplicamos este modelo a niveles superiores de análisis lingüístico, como el léxico, nos encontramos con que carecemos de inventarios de las unidades o invariantes de cualquier lengua tan precisos como los sistemas fonológicos, habida cuenta de la especificidad de las unidades fonológicas, carentes de significado pero distinguidoras del mismo. En el nivel léxico encontramos distintas opiniones, desde aquellos que niegan la variación tal y como la definíamos más arriba, ya que cada unidad léxica comporta un matiz significativo diferente y todas han de ser consideradas en ese sentido invariantes, hasta aquellos otros (entre los que nos incluimos) que consideran necesario manejar otro concepto de variación para las unidades significativas y entonces establecer series de elementos léxicos (nos hemos situado ahora en ese nivel de análisis, recuérdese) que o bien comparten su expresión (total o parcialmente) y difieren en su contenido, o bien (nos centraremos en ellos en esta investigación)

comparten al menos parte de su contenido y difieren en su expresión (piénsese en casos como los que trataremos después “empezar”/“comenzar” o “regalo”/“presente”). La utilidad de manejar dicha concepción viene dada por la información que la sociolingüística variacionista, como veíamos más arriba, puede darnos para verificar si la utilización de uno u otro de los elementos de una serie está directamente condicionada por factores de índole social, lo que supone un mejor conocimiento del uso de las lenguas.

De entre los muchos factores sociales (algunos de ellos ya citados) que la sociolingüística ha estudiado para caracterizar los diferentes usos lingüísticos hay varios que bien de forma separada bien de forma conjunta competirían a lo que podríamos denominar el grado de formalidad de nuestro uso. Es intuitivo que en todas las lenguas, a pesar de las múltiples opiniones existentes en lo que concierne a su delimitación y denominación, existen diferentes grados de formalidad que irían desde un discurso más formal hasta uno más espontáneo. Desde la sociolingüística, y no sólo desde ella, también desde otras disciplinas como la etnografía de la comunicación, por ejemplo, se han citado diferentes rasgos que condicionan la formalidad del discurso¹. Podría pensarse que ese grado de formalidad (configurador de lo que algunos autores denominan estilos, otros registros) afecta a las unidades léxicas, de forma que para una comunidad de habla un determinado elemento léxico se asocia a un uso más formal o más espontáneo. Es también posible pensar que esa diferencia en la caracterización de la formalidad de los elementos puede ser uno de los factores que nos ayude a distinguir en ciertos casos entre series de elementos como los mencionados más arriba, de expresión diferente y contenido común, y que la metodología sociolingüística puede prestar una gran ayuda para ese tipo de indagaciones.

2. Texto y contexto

Consideremos ahora más detalladamente algunos de los rasgos que hemos mencionado como condicionadores de la formalidad del discurso. Sin duda, todos y cada uno de ellos conforman lo que actualmente denominamos *comunicación*, y su importancia ha sido puesta de manifiesto por especialistas de muy diversas disciplinas. Claro que éstos han resaltado aquellos aspectos que convenían en cada caso a los fenómenos lingüísticos estudiados. Puesto que pretendemos centrar nuestra atención en el contexto y en el tópico de discurso (tema), adoptaremos el punto de vista proporcionado por la lingüística textual y discursiva, conjunto de disciplinas que, como la sociolingüística, abordan el estudio del **uso** desde una perspectiva comunicativa. Puesto que la noción de contexto resulta todavía demasiado amplia para los fines que aquí nos hemos propuesto, habremos de entenderlo como el contexto lingüístico pertinente, en la medida en que es manifestación de tipos estandarizados de contextos verbales mediatos, llamados *contextos temáticos* (Coseriu 1955-56: 49).

El contexto y también el estilo o registro que éste determina, pueden observarse así, en cierto modo, como el *marco / fondo* de la comunicación, una de cuyas *figuras* viene repre-

¹ F. Moreno Fernández (1998: 102) recoge algunos de los más comunes como el propio hablante, la audiencia, el tema y el tipo de comunicación, o el momento y el tipo de actividad.

sentada por las variantes léxicas que los hablantes seleccionan realmente. Desde este punto de vista, los factores implicados en la estratificación social que los sociolingüistas llevan a cabo en sus investigaciones van más allá de su incidencia puntual en las variantes efectivamente seleccionadas por los hablantes de una lengua dada. Necesariamente dichos factores deben ser relacionados, como sabemos, con la competencia lingüística de los hablantes, lo que supone, entre otras habilidades, un conocimiento más o menos amplio de las formas textuales más adecuadas a las situaciones de comunicación y a los tópicos de discurso concretos, de ahí que para muchos autores deba hablarse más correctamente de *competencia comunicativa* o incluso de *competencia sociolingüística*.

Ahora bien, particularmente en la lingüística del texto y corrientes relacionadas, puede inferirse una noción alrededor de la cual gravita una parte nada desdeñable de las teorías formuladas, ubicadas, por tanto, en un *enfoque comunicativo* del lenguaje. Se trata de la *adecuación*², concepto que, sin haber sido mencionado directamente y desde luego en ningún caso delimitado, resulta, en nuestra opinión, un factor fundamental a la hora de calibrar el grado de formalidad de un texto y sus implicaciones en la selección de unidades léxicas pertinentes. En cierto modo, este concepto representa la versión actual del antiguo problema de los estilos y registros. La adecuación subyace frecuentemente a diferentes propuestas teóricas en el seno de la lingüística textual y se revela como un hecho de difícil aprehensión para teorías textuales concretas. Probablemente, se ha dado por sentado que lo que había de interesar a la lingüística del texto eran los textos indisolublemente unidos a sus contextos, de manera que la discontinuidad entre textos concretos y situaciones de comunicación concretas era salvada, bien mediante el recurso al concepto de «coherencia», con todas sus extensiones y matizaciones (*coherencia global, coherencia pragmática, coherencia de contenido*, etc.), bien mediante la adscripción de valores estilísticos a los textos, como los resultantes de procedimientos como la ironía, el humor o la metáfora, entre otros. La adecuación es, pues, una variable sociocultural y psicológica, un patrón si se prefiere, que tiene la propiedad de ser condición necesaria, aunque no suficiente, para producir e interpretar textos en contexto. Asimismo, la estructura del léxico empleado para designar las entradas enciclopédicas que configuran cada tópico de discurso particular coadyuba en iguales condiciones a este mismo fin³. Pues bien, el grado de formalidad con-

² Es importante resaltar que no nos referimos a la *adecuación* tal como se define ésta entre los principios reguladores de la textualidad en el modelo de De Beaugrande y Dressler (1981), sino más bien en la línea de Blakemore (1988: 280).

³ Desde luego, la teoría en lingüística del texto da muestras suficientes del relevante papel que desempeña la organización del léxico en los textos. Incluso prescindiendo de marchamos como *tipo de texto, forma textual, género discursivo, secuencia* y otros que se han propuesto a la hora de tratar de establecer tipologías textuales, el léxico de un texto de organiza hasta tal punto que, con frecuencia, las conexiones entre sus elementos trascienden los límites de la semántica lingüística, como es el caso, según cuál sea el repertorio escogido, del mecanismo de la cohesión léxica o el de las relaciones semánticas entre lexemas, de los que en ocasiones no puede decirse más que los elementos léxicos se asocian de alguna manera. Claro está que, en el caso del texto, nos hallamos ante niveles del significar distintos, ya que tratamos con significados actualizados, esto es, con *sentidos*. De ahí que nos decanemos cada vez más por una concepción de la sustancia semántica conformada y realizada, el sentido, como resultado de las mismas operaciones cognitivas generales que nos permiten percibir, discriminar y jerarquizar el mundo que nos rodea y, por tanto, también el lenguaje que usamos para referirnos a él.

stituye uno de los índices más evidentes de la adecuación textual, y como tal será tratado a lo largo de este trabajo.

Existe una relación indiscutible entre el estándar textual y el léxico empleado para la producción de textos reales⁴. Pero, ¿qué es determinante y qué determinado? Sin duda, y la intuición así lo avala, el plan global del hablante y, correlativamente, las expectativas del receptor o receptores del acto de comunicación presuponen un estado de cosas o un *sistema de referencia*, que determina la selección de las variantes léxicas más apropiadas para los fines comunicativos en los que se emplean. Desde ese punto de vista, estándares textuales y variantes léxicas, así como otros muchos aspectos lingüísticos, sirven a una misma finalidad comunicativa en cada acto de comunicación. De manera análoga a los que consideran toda unidad léxica como invariante, cabría pensar con respecto a los textos, en cuanto que actos de habla (complejos) únicos e irrepetibles. En este sentido, siguiendo un principio de identidad según el cual los textos no pueden sino ser iguales a sí mismos, Trujillo (1996: 25), por ejemplo, considera inviable una tipología de textos. Pero, en nuestra opinión, esta cuestión no afecta para nada a la existencia de estilos o registros adecuados a contextos y también a situaciones de comunicación. Unos y otras se caracterizan por ser particulares, únicos e irrepetibles en el momento en que se producen pero, a su vez, presentan elementos constantes cuyas posibilidades de actualización (potencial) se presuponen o son esperables de manera que proporcionan configuraciones reales accesibles a todos y cada uno de los participantes en la actividad del hablar. Además, es constatable que los estilos / registros trascienden los límites impuestos por la tipología de texto, ya que diferentes tipos de textos, sujetos a finalidades comunicativas distintas, pueden presentar un determinado grado de formalidad según el contexto en que se producen. Por tanto, los textos puede muy bien seguir resistiéndose a la tipologización, pero sin duda obedecen, como ya hemos señalado, a unos patrones socioculturales y psicológicos que permiten a los individuos de una comunidad de habla producir e interpretar adecuadamente textos en contexto. Se sigue de lo dicho que algunos textos resultan más adecuados que otros en determinados contextos de comunicación.

Algunas de las variables que determinan el grado de formalidad de un texto o estándar textual son el tópico de discurso, la situación social de los interlocutores o de los potenciales receptores de la comunicación, el canal de difusión, etc. Más específicamente, Werlich (1976: 274) define el estilo formal como una variedad referencial de la *forma textual* cuyas marcas reflejan el respeto que el oyente muestra hacia su destinatario. Entre otros, cita como convencionalismos propios de este estilo su empleo dirigido a un público amplio o a un des-

⁴ Los tres aspectos fundamentales según los cuales el conjunto de las unidades léxicas de un texto inciden en la estructuración y organización de éste son, en nuestra opinión, el *tópico de discurso*, el *registro / estilo* (es decir, la adecuación) y la *semántica del texto* (relacionada con la informatividad). Más aún, todos ellos son imposibles de analizar si no es en relación con el contexto, desde una teoría pragmática del texto o *pragmática textual*, puesto que prueban el hecho ya aceptado de que el texto que se comprende no es igual al que se ha producido o, dicho de otro modo, que lo que como intérpretes obtenemos es una versión aproximada del contenido y la intención del productor. De ahí la importancia del conocimiento compartido, de la existencia de un estado de cosas que es accesible al hablante y al oyente, y de principios generales como el de la relevancia.

tinario desconocido. Son marcas frecuentes del mismo el uso del impersonal o de la 1ª persona del plural, las selecciones léxicas complejas y precisas, etc. En general, cabe señalar que, dadas sus múltiples finalidades comunicativas, los textos escritos tienden a favorecer la presencia de un grado más alto de formalidad. En este sentido, el productor del texto cuenta con un lapso de tiempo mayor para actualizar su plan previo, lo que supone un mayor cuidado a la hora de seleccionar las estructuras más apropiadas, además de que este canal de difusión puede actualizarse por medio de géneros discursivos cuyos mensajes se dirigen frecuentemente a un destinatario indeterminado, o a muchos de ellos que, en todo caso, no están presentes.

Pues bien, en última instancia, hemos pretendido en esta investigación poner en correlación aquellos factores sociolingüísticos que determinan el uso de variantes léxicas con los textos-contexto en que éstas se emplean, de acuerdo con las finalidades comunicativas que los textos imponen. Para ello, hemos empleado, en primer lugar, dos textos pertenecientes a géneros discursivos distintos, una noticia periodística y una carta al director, caracterizados como formales de acuerdo, entre otros factores, con los destinatarios de la comunicación, en ambos casos un público indeterminado. Además, se ha facilitado a los sujetos encuestados un tercer texto, una carta personal que es variante a su vez del mismo género discursivo, el epistolar, al que pertenecía el segundo texto citado, pero caracterizado en este caso como informal, por estar dirigido a un hipotético amigo.

3. Aspectos metodológicos

Una vez abordados algunos de los presupuestos teóricos en los que fundamentamos nuestro trabajo y antes de presentar los principales resultados obtenidos en nuestra investigación, nos detendremos en la explicación de algunas cuestiones de carácter metodológico. Concretamente, haremos referencia a los procesos del diseño de la prueba, recogida de datos y procesamiento de los mismos, y a la caracterización de la herramienta informática empleada, el programa AQUAD.

3.1. Diseño de la prueba y recogida de datos

Para nuestra investigación seleccionamos cuarenta elementos léxicos del español, divididos en veinte parejas y caracterizados por poseer una expresión diferente y un contenido parcialmente común: *marido / esposo, borracho / ebrio, atar / amarrar, tirar / jalar, amigo / colega, cura / sacerdote, camisa / blusa, bonito / mono, periódico / diario, estudiar / empollar, empezar / comenzar, trabajar / currar, dinero / pasta, ayuda / auxilio, dañino / perjudicial, afecto / cariño, regalo / presente, enfadarse / mosquearse, embarazada / preñada y terminar / finalizar*. La existencia de un nexo de unión entre los miembros de cada pareja venía avalada por su tratamiento en el *Diccionario de la Lengua Española* de la Real Academia Española. De los veinte grupos mencionados, cinco contenían definiciones muy similares e incluso idénticas en el citado diccionario; en tres casos encontramos definiciones circulares en las que en la definición de cada palabra se mencionaba la otra, y en los doce restantes una de las palabras aparecía reflejada en la definición de la otra, en siete ocasiones,

además, con marca de uso que la diferenciaba (*familiar* en tres ocasiones, *coloquial* en dos y *popular* y *despectivo* y *entre estudiantes* en una), hecho éste que corroboraba nuestra intuición acerca de la posibilidad de hallar diferencias en el uso de estos elementos merced a su diferente grado de formalidad al tiempo que evidencia la falta de uniformidad y rigurosidad en el sistema de etiquetas relativas a marcas de uso en el diccionario académico.

Para la recogida de muestras diseñamos una encuesta compuesta de tres textos de carácter escrito, ficticios, entre los que intentábamos establecer una diferencia en cuanto a la formalidad de su uso, y en los que aparecían cada una de las veinte parejas para que el informante eligiese en todas ellas cuál de las dos posibilidades le parecía más idónea para cada texto, mostrara su indiferencia ante las mismas o incluso propusiera una tercera posibilidad. En este sentido, tal y como se recoge en el segundo punto de este artículo, se utilizaron, en ese orden, una noticia periodística, una carta al director y una carta informal.

El trabajo de campo consistió en la realización de las encuestas de manera presencial a una muestra de 54 informantes nacidos y con la mayor parte de su vida al menos desarrollada en la comunidad de habla de Cádiz. Para la elección de esos informantes, seleccionamos una muestra poblacional estratificada por cuotas de asignación fija compuesta de cincuenta y cuatro informantes, atendiendo a su sexo (hombres frente a mujeres), edad (hablantes de entre 14 y 30 años, de entre 31 y 50 años y de más de 50 años) y nivel de estudios (sin estudios, con estudios medios como máximos alcanzados y con estudios superiores como máximos alcanzados).

3.2. Procesamiento de los datos. Codificación con AQUAD

En lo que se refiere al procesamiento de los datos, hemos hecho uso de la herramienta AQUAD (Analysis of QUALitative Data)⁵. Este programa, desarrollado por Huber para el análisis de datos cualitativos, proporciona a los investigadores del ámbito de las ciencias sociales un medio para la comprobación de hipótesis previas basadas en datos subjetivos. Particularmente, se ha empleado con notable éxito en los campos de la pedagogía y la psicología, pero no es habitual su empleo en lingüística.

De entre las diferentes motivaciones que pueden impulsar una investigación en el ámbito de las ciencias sociales, el programa se adapta perfectamente a aquélla que trata de establecer vínculos entre los conceptos que subyacen a los datos analizados. Pues bien, el proceso según el cual se delimitan los segmentos relevantes de un texto y se establecen relaciones entre los datos reales y los conceptos a los que éstos se refieren (esto es, definir categorías mediante códigos) es lo que habitualmente se denomina en metodología de la investigación *codificación*. Así, el programa AQUAD ofrece diferentes tipos de códigos que con-

⁵ Concretamente, hemos trabajado con la versión 5.7 del programa, si bien puede encontrarse ya en el mercado la nueva versión 6.0 que introduce sensibles mejoras, como la posibilidad de procesar, no ya sólo textos escritos, sino también archivos de sonido (*.wav, *.mp3), imagen (*.jpg) y video (*.avi). Puede hallarse más información en <http://www.aquad.de>.

tribuyen a una codificación útil de los datos mediante categorías conceptuales. Los distintos tipos de códigos son los siguientes: *conceptuales*, *socio-demográficos*, *numéricos*, *de control*, *secuenciales* o *de vínculos* y *de hablante* (cf. Huber, Fernández Smith, Lorenzo Quiles & Herrera Torres 2001: 87-90).

En la codificación de las categorías estudiadas en este trabajo, hemos empleado, principalmente, códigos socio-demográficos, también llamados *códigos de perfil*, y códigos conceptuales. Los primeros nos han servido para llevar a cabo la estratificación social de la muestra, dado que son la clase de códigos que sirven para determinar perfiles para cada uno de los archivos al completo. En este sentido, estos “códigos singulares” se caracterizan porque son excluyentes, de tal manera que un texto codificado como perteneciente a un sujeto ‘mujer’ no puede serlo a la vez perteneciente a un sujeto ‘hombre’. Así pues, estos códigos particularizan, a la vez que agrupan, los textos en función de unas características que les son propias. En cuanto a los segundos, los códigos conceptuales, se corresponden con las categorías asignadas a las unidades léxicas objeto de nuestro análisis. Los códigos conceptuales sirven con eficacia al proceso de conceptualización e interpretación de los datos reales. Resultan, por otra parte, fundamentales a la hora de establecer posteriormente vínculos entre las categorías que sirvan para validar hipótesis previas. El programa AQUAD lleva a cabo estas acciones mediante el módulo de implicaciones lógicas, basado en el álgebra booleana, que convierte configuraciones constantes de códigos en valores binarios de tablas de verdad. Dichos valores pueden ser combinados después mediante conectores lógicos a partir de diferentes criterios de comparación. En relación con estos tipos de códigos que, recordémoslo, son sólo dos de las seis opciones que ofrece AQUAD, Medina Díaz & Marín Sánchez (1998: 384), por ejemplo, distinguen entre las que denominan *unidades de base gramatical* y *unidades de base no gramatical*. Las primeras comprenden la producción verbal real del sujeto encuestado (nuestros códigos conceptuales), mientras que las segundas representan características propias del sujeto hablante o del documento en su totalidad (códigos de perfil en nuestra investigación).

En síntesis, los parámetros que hemos tenido en cuenta a la hora de realizar la codificación de las encuestas son los siguientes: por una parte, para los códigos sociodemográficos hemos utilizado los tradicionales factores sociolingüísticos de la edad ([/gener_1], [/gener_2] y [/gener_3]), el sexo ([/masc] y [/fem]) y el estrato sociocultural, definido por el nivel de estudios ([/estudios_0], [/estudios_1] y [/estudios_2]). Dados estos parámetros, la codificación mediante códigos de perfil arroja una matriz de 2*3*3 y, puesto que contamos con tres ejemplos para cada posibilidad, ello supone un balance total de 54 individuos encuestados. Por otra parte, para los códigos conceptuales hemos diseñado una propuesta basada en la concurrencia de tres factores, a saber: a) marcas propias de la unidad léxica, no formal ([tf-]) frente a formal ([tf+]), indiferencia ([indifer]), o bien otra ([otra]); b) marcas de aparición en un contexto formal ([cf+]) frente a un contexto no formal ([cf-]), y c) como resultado de la combinación de los anteriores, adecuación ([tf+cf+] y [tf-cf-]) o inadecuación ([tf+cf-] y [tf-cf+]) de su uso. En este caso, hemos trabajado sobre la base de veinte parejas léxicas distribuidas en tres textos, de manera que los sujetos encuestados contaban con dos oportunidades de realizar una operación de selección para cada par léxico, es decir,

cuarenta respuestas reales sujetas a la codificación mediante códigos conceptuales. Si multiplicamos estas cuarenta respuestas reales por el número de encuestados obtenemos un total de 2160 respuestas que, divididas entre cada uno de los factores sociolingüísticos citados, arrojan totales de 1080 respuestas para hombres y mujeres, 720 para cada uno de los tres grupos de edad y otras 720 para cada uno de los tres estratos socioculturales definidos por el nivel de estudios.

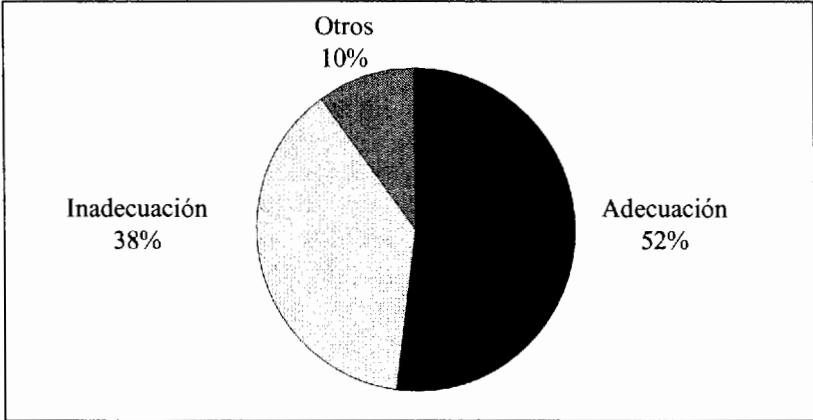
Desde luego, debe señalarse que el programa proporciona muchas otras opciones relativas a la forma en que aparecen y se organizan los códigos en un texto, que facilitan el establecimiento de vínculos más complejos. Así, por ejemplo, AQUAD puede trabajar con estructuras de códigos anidados, códigos superpuestos, códigos múltiples, secuencias de códigos y secuencias repetidas de códigos. Puesto que, por sus especiales características, este tipo de estructuras no se presentan en la muestra, la presente investigación no explota en toda su extensión las posibilidades, ni de AQUAD, ni del análisis cualitativo aplicado a parcelas concretas de la lingüística, pero supone, en nuestra opinión un buen ejemplo de la ayuda que la herramienta puede prestar a los lingüistas.

4. Análisis cualitativo y cuantitativo e interpretación de los datos

En las siguientes líneas mostraremos algunos de los resultados más relevantes obtenidos en nuestra investigación de entre las muchas posibilidades de análisis consideradas. En primer lugar, el fenómeno que, de manera general, despertó nuestro interés desde el comienzo es lo que aquí hemos dado en denominar *adecuación* o *inadecuación*, esto es, el hecho de que haya correspondencia entre palabras y textos bien formales bien informales. Dicho de otro modo, hablamos de *adecuación*: a) en los casos en que aparece la variante léxica considerada formal en nuestro trabajo en el texto considerado formal, y b) en los casos en que aparece la variante léxica considerada informal en el texto considerado informal.

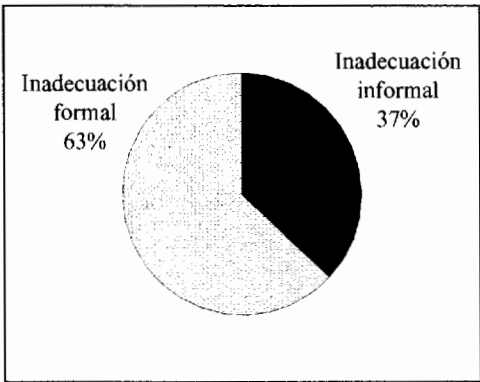
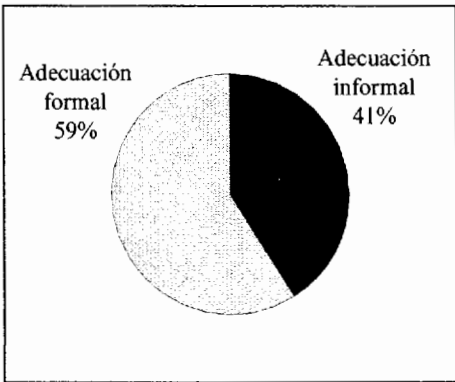
Los casos contrarios (variante léxica formal en texto informal o variante léxica informal en texto formal) se consideran ejemplos de *inadecuación*. Los datos totales obtenidos evidencian que los hablantes asociaron palabras y textos en más de la mitad de los casos. Concretamente, de un total de 2160 realizaciones en 1122 se produjo adecuación y en 822 ésta no se dio⁶.

⁶ Ha de tenerse en cuenta, además, que para el establecimiento de casos totales se computan no sólo los casos de *adecuación* o *inadecuación*, sino también todas aquellas ocasiones (216 en total) en las que el informante se mostraba indiferente en el uso de las dos variantes propuestas o incluso proporcionaba una tercera alternativa no presente en nuestra prueba.



Este primer resultado confirma, al menos de forma parcial, la validez de nuestra hipótesis de partida en cuanto a la relación entre el grado de formalidad de elementos léxicos y textuales.

Un análisis más minucioso de la (in)adecuación nos permite establecer una diferenciación según el carácter de textos y palabras. En este sentido, en los ejemplos de *adecuación* podemos distinguir: a) casos en los que aparece la palabra formal en el texto formal, en los que hablaríamos de *adecuación formal*, y b) casos en los que aparece la palabra informal en el texto informal, en los que hablaríamos de *adecuación informal*. De igual manera, en los casos de *inadecuación* se puede diferenciar entre: a) casos en los que aparece la palabra formal en el texto informal, en los que hablaríamos de *inadecuación formal*, y b) casos en los que aparece la palabra informal en el texto formal, en los que hablaríamos de *inadecuación informal*. Los resultados obtenidos al respecto se reflejan en las siguientes gráficas:



Es evidente el predominio de elementos léxicos formales tanto en los casos en los que estos eran esperables como en los que nuestra intuición como hablantes no presagiaba su

aparición⁷. De esta manera parece mostrarse una preferencia generalizada por tal tipo de variantes léxicas, las más formales, que se constituyen en los “elementos no marcados” de las parejas analizadas. Como posibles factores que nos ayuden a entender este hecho cabe pensar en el carácter escrito de los textos seleccionados (más proclives que los discursos orales a un mayor grado de formalidad, como se mencionaba en el punto 2 de este artículo), en una mayor amplitud de contenido de los elementos léxicos considerados formales en nuestra investigación que sus pares etiquetados como informales, o incluso en una tendencia natural de los hablantes seleccionados que, conscientes del diferente grado de formalidad de las alternativas propuestas en cada caso y del prestigio asociado a ellas, evitan en mayor medida la presencia “incorrecta” de elementos menos formales (y como tales menos prestigiosos) en todos los textos en general.

Esta última idea se refuerza tras el análisis de cada una de las parejas de variantes léxicas estudiadas, ya que el predominio general de elementos formales se debe en gran medida a la abrumadora presencia del elemento formal de 7 de las parejas de variantes: *tirar* (80.8%) / *jalar* (19.1%), *amigo* (81.1%) / *colega* (18.8%), *estudiar* (87.3%) / *empollar* (12.6%), *trabajar* (84.5%) / *currar* (15.4%), *dinero* (88.3) / *pasta* (11.6%), *bonito* (94.8%) / *mono* (5.1%) y *enfadarse* (82.4%) / *mosquearse* (17.5%)⁸, en las que parece más evidente aún que en las demás el carácter informal de una de las alternativas (*jalar*, *colega*, *empollar*, *currar*, *pasta*, *mono* y *mosquearse*), la no utilizada en la mayoría de ocasiones, ni siquiera en el texto menos formal, la carta personal. Curiosamente, y corroborando esta intuición, 6 de esos 7 elementos son los únicos de los 40 analizados etiquetados con marca de uso, de informalidad, en el *Diccionario de la Lengua Española* de la R.A.E., en el que se tacha de coloquial a *colega* y *currar*, de familiar a *mono* y *jalar*, de popular a *pasta* y de propio de estudiantes y despectivo a *empollar*⁹. Diferencias similares y opuestas, significativas en el uso general a favor de la variante informal, sólo se dan en 3 parejas presentes en nuestra prueba: *embarazada* (96.2%) / *preñada* (3.7%), *borracho* (80.9%) / *ebrio* (19%) y *regalo* (95.1%) / *presente* (4.8%), en las que, a nuestro juicio, el elemento formal (*preñada*, *ebrio* y *presente*) alcanza un mayor grado que en otras parejas hasta el punto de hacer que el considerado informal funcione como elemento neutro en cuanto a ese rasgo.

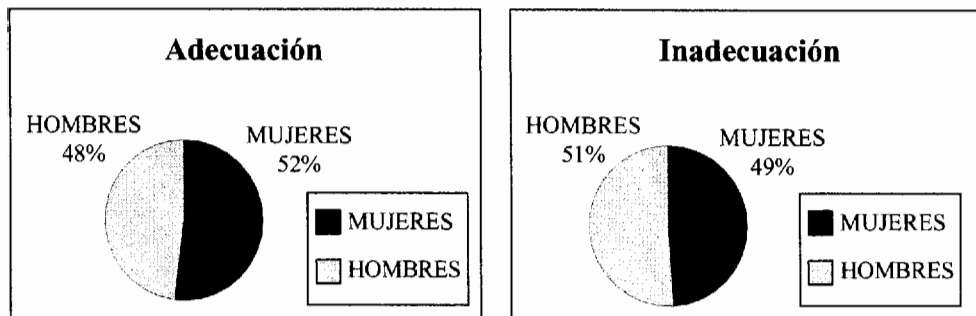
Los factores sociales tomados en cuenta en nuestra investigación (sexo, edad y nivel de estudios), además de perseguir una mayor representatividad de la muestra analizada con respecto a la comunidad de habla de la que se extrajo, proporcionan información de carácter sociolingüístico acerca de los fenómenos estudiados. Mostraremos a continuación algunas ideas generales acerca de la información proporcionada en este trabajo. El factor sexu-

⁷ De hecho las apariciones de elementos léxicos formales alcanzan un total de 1182 casos frente a los 762 de los informales.

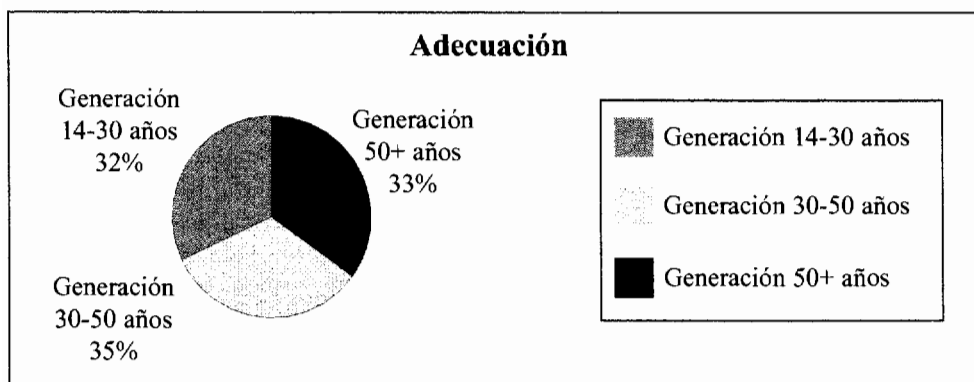
⁸ Los datos aportados expresan el porcentaje general de aparición de cada una de las variantes en el conjunto de la prueba.

⁹ Sólo faltaría en este caso un etiquetado parecido para el elemento *mosquearse*. Si bien el sistema de marcas de uso del diccionario académico dista bastante de ser un marco de referencia exhaustivo, puede ser considerado un apoyo para refrendar alguna de nuestras ideas por la información que proporciona, por incompleta que sea.

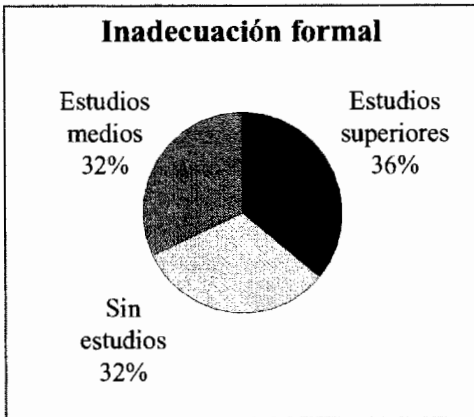
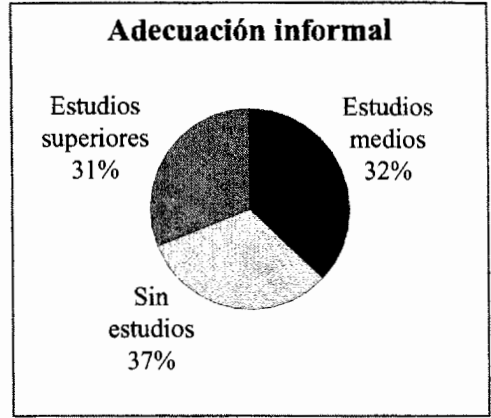
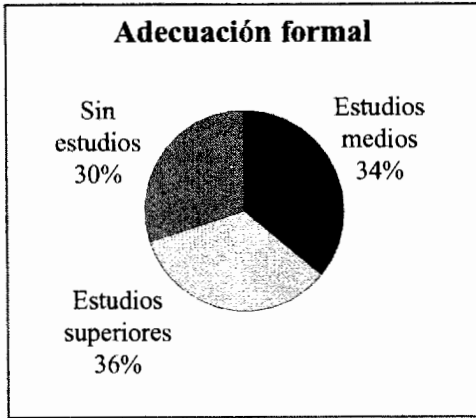
al mostró en líneas generales una mayor adecuación, leve eso sí, en las mujeres de la muestra:



Este hecho parece refrendar, si bien, repetimos, se produce de forma moderada y no muy significativa desde el punto de vista estadístico, aquellas teorías surgidas en el marco de la sociolingüística que evidencian la preferencia de las mujeres en comunidades occidentales como la nuestra por las formas estándares, ya sea por una cuestión de prestigio o de conservadurismo (*cf.*, al respecto, por ejemplo, la completa monografía de López Morales 2004: 122-131). La existencia de tres estratos generacionales en la muestra de población analizada se reveló como el factor menos importante *per se*, de manera aislada, más allá de su utilidad para confiar en la representatividad de la muestra de informantes seleccionada al recogerse igual número de testimonios de diferentes fracciones de edad. El único hecho destacado es, quizás, el mayor grado de adecuación de la generación intermedia (hablantes de entre 31 y 50 años) por encima de las generaciones jóvenes y adulta, como se refleja en el siguiente gráfico:

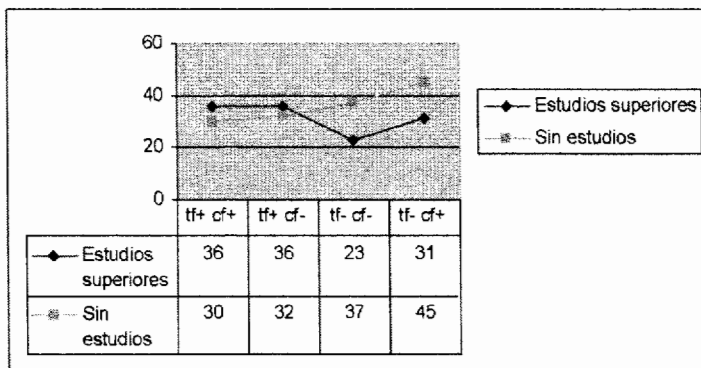


Por el contrario, la consideración en nuestra investigación del nivel de estudios de los informantes (divididos, recuérdese, en hablantes sin estudios, hablantes con estudios medios como máximo alcanzados y hablantes con estudios superiores), mostraba una relación directa entre dicho nivel de estudios y la formalidad expresada en textos y elementos léxicos:



En los informantes con estudios universitarios, los más altos de la muestra, se daba una preferencia mayor por los elementos léxicos más formales, ya fuera en textos formales (como se recoge en el primer gráfico, el de adecuación formal) o informales (tal y como se aprecia en el tercer gráfico, el que hace referencia a la inadecuación formal), al tiempo que se promocionaba en menor medida las palabras informales, tanto en unos textos como en otros (obsérvense los gráficos segundo y cuarto, correspondientes a la adecuación no formal y la inadecuación no formal). En el extremo opuesto, la población sin estudios reflejaba el comportamiento contrario, ya que estos hablantes eran los más reacios a elegir los elementos formales y los que más seleccionaban los informales, independientemente de su adecuación textual¹⁰:

¹⁰ Recuérdese que los códigos utilizados han de entenderse de la siguiente manera: tf+cf+ (palabra formal usada en contexto formal), tf+cf- (palabra formal usada en contexto informal), tf-cf- (palabra informal usada en contexto informal) y tf-cf+ (palabra informal usada en contexto formal).



El cuadro anterior refleja las tendencias cruzadas de ambos grupos, separados por su nivel de estudios, en los que parece quedar clara la influencia de la acción educativa en la elección entre variantes separadas por el grado de formalidad.

Más allá del análisis de los factores sociales mencionados de manera aislada, nuestro estudio investigó la posible correlación de los mismos. En este sentido, y corroborando los datos aportados en cuanto a sexo y nivel de estudios, el análisis de los distintos factores cruzados desveló la existencia de dos grupos antagónicos en cuanto a su comportamiento ante la formalidad de léxico y texto: los hombres sin estudios y las mujeres con estudios superiores. En la siguiente tabla puede apreciarse cómo los hombres sin estudios son los que, en general, menos usan elementos léxicos formales y más utilizan palabras informales mientras que las mujeres con estudios superiores hacen justo lo contrario, aportan los valores más altos en la producción de léxico formal y más bajo en la de informal:

	Sin estudios		Estudios medios		Estudios superiores	
	Hombres	Mujeres	Hombres	Mujeres	Hombres	Mujeres
tf+cf-	83	86	90	78	95	89
tf+cf+	96	102	115	108	110	130
tf-cf-	86	87	64	83	69	72
tf-cf+	71	65	39	58	40	28

Dicha actitud en el uso del léxico recuerda algunos patrones de comportamiento destacados por la sociolingüística variacionista en el estudio, sobre todo, de la variación fonética. Así, podría pensarse que el grupo constituido por las mujeres de estudios superiores se acomodan con mayor facilidad en las formas de prestigio (las variantes léxicas formales en nuestra investigación) por el acceso a dichas formas mediante una educación superior y la habitual preferencia femenina por las formas estándares ya comentada, mientras que es fácil adivinar que los hombres sin estudios, por menor contacto educativo con estas formas y por

el prestigio encubierto que confieren, prefieran las formas menos prestigiosas (las variantes léxicas informales, en este caso).

5. Conclusiones

Finalmente, señalaremos algunas de las conclusiones que, creemos, pueden extraerse de nuestra investigación. Son, sin duda, conclusiones que proyectos similares pero de mayor envergadura, que se realizaran a mayor escala en cuanto a su número de informantes y de comunidades de habla, podrían demostrar más fehacientemente. El desarrollo de esta investigación nos permite, por tanto, extraer algunas conclusiones en lo que respecta al objeto principal de estudio de este trabajo, la adecuación de elementos léxicos y textuales; a determinadas afirmaciones en torno al estudio de factores sociales en la investigación lingüística, y a aspectos de índole metodológica en estudios como el aquí presentado.

En primer lugar, desde un punto de vista más general, que afecta tanto a la epistemología lingüística como a sus métodos de investigación, consideramos que estudios como éste certifican la validez de una perspectiva interdisciplinaria desde la cual deben abordarse ciertas cuestiones relativas a la manera en que los seres humanos se comunican verbalmente, esto es, a una pretendida *lingüística del hablar*. Así, mediante la estrecha colaboración entre disciplinas o, si se quiere, perspectivas diferentes dentro de la lingüística, como son las de la sociolingüística y la lingüística textual y discursiva, este trabajo contribuye, en alguna medida, al establecimiento de bases para una incipiente teoría de la adecuación, uno de cuyos índices más evidentes es el grado de formalidad.

Por otra parte, pueden extraerse conclusiones más específicas que atañen a los puntos de vista mencionados. En primer lugar, el análisis de elementos léxicos y textos del español parece concluir la existencia, corroborada por nuestra intuición como hablantes de la comunidad de habla analizada, de diferencias en el grado de formalidad tanto en unidades léxicas de expresión diferente y contenido común como en textos diferenciados por la distinta naturaleza de diversos aspectos comunicativos implicados en su realización como las características de emisor y receptor o el tema del mismo. Dichas diferencias en el grado de formalidad se manifiestan en la adecuación en el uso de elementos léxicos y textos, bien formales bien informales, lo que confirma, además, la presencia de dichos elementos léxicos como un rasgo más caracterizador de la formalidad de los textos. El análisis de factores sociales en nuestra investigación (concretamente sexo, edad y nivel de estudios) abunda en algunas de las conclusiones obtenidas por la sociolingüística, sobre todo en el estudio de diferentes niveles de análisis lingüístico como el fonético-fonológico. Apreciamos una mayor preferencia de las mujeres de la muestra por las formas estándares, las variantes léxicas formales (unido al mayor uso en los hombres de variantes informales, que proporcionan un prestigio encubierto), al tiempo que el ascenso en el nivel de estudios se correlaciona proporcionalmente con una mayor utilización de elementos formales, debido a la influencia de la acción educativa en el uso de la lengua. Así, parecen corroborarse, en un nivel de análisis, el léxico, distinto al que habitualmente se ha manejado con más éxito, determinados principios generales establecidos por la sociolingüística variacionista. Asimismo, este trabajo podría dotar de nuevas variables a la teoría de los tipos textuales con las que se caracteri-

zaría más rigurosamente la sistematicidad de sus propuestas, contribuyendo, por tanto, a la descripción de *invariantes textuales* (cf. Escoriza Morera 2002: 127-132).

Desde el punto de vista metodológico, nuestra investigación garantiza la utilidad de herramientas informáticas como el programa AQUAD en el estudio cualitativo de elementos lingüísticos y advierte de la necesidad de desarrollo de tal tipo de instrumentos para un mejor conocimiento del uso de los sistemas lingüísticos.

Aunando las diversas perspectivas que aquí se vislumbran, cabe establecer como reflexión final la reafirmación del hecho conocido de la complejidad de los actos de comunicación. Los diversos procedimientos lingüísticos que posibilitan la actividad verbal funcionan de manera solidaria en diferentes niveles de concreción, pero todos ellos están sujetos a un comportamiento teleológico regido por fórmulas convencionales, psicológica y socioculturalmente definidas, que realizan los propósitos comunicativos para los que actuamos verbalmente.

6. Referencias

- BEAUGRANDE, R.-A. de & W. U. DRESSLER (1981, 1997), *Introducción a la lingüística del texto*. Versión española de S. Bonilla, Barcelona: Ariel.
- BLAKEMORE, D. (1988, 1992), "La organización del discurso", F. J. Newmeyer (comp.), *Panorama de la lingüística moderna de la Universidad de Cambridge. IV El lenguaje: contexto socio-cultural*. Edición supervisada por L. Eguren, Madrid: Visor, pp. 275-327.
- COSERIU, E. (1955-56), "Determinación y entorno. Dos problemas de una lingüística del hablar", *Romanistisches Jahrbuch*, 7, pp. 29-54.
- ESCORIZA MORERA, L. (2002), *La variación lingüística. Propuesta de delimitación de variantes en el nivel léxico*. Tesis doctoral, Cádiz: Universidad de Cádiz.
- HUBER, G. L., FERNÁNDEZ SMITH, G., LORENZO QUILES, O. & L. HERRERA TORRES (2002), *Análisis de datos cualitativos con AQUAD Cinco para Windows*, Granada: Grupo Editorial Universitario.
- LÓPEZ MORALES, H. (2004), *Sociolingüística*, Madrid: Gredos.
- MEDINA DÍAZ, F. J. & M. MARÍN SÁNCHEZ (1998), "Análisis de contenido. Reflexiones sobre el uso de categorías", Ruiz Castellanos, A., Viñez Sánchez, A. & J. Sáez Durán (eds.), *Retórica y texto*, Cádiz: Universidad de Cádiz, pp. 383-387.
- MORENO FERNÁNDEZ, F. (1998), *Principios de sociolingüística y sociología del lenguaje*, Barcelona: Ariel.
- REAL ACADEMIA ESPAÑOLA (1992, 21ª ed.), *Diccionario de la Lengua Española*, Madrid: Espasa-Calpe.
- TRUJILLO, R. (1996), *Principios de semántica textual. Los fundamentos semánticos del análisis lingüístico*, Madrid: Arco/Libros.
- WERLICH, E. (1976), *A Text Grammar of English*, Heidelberg: Quelle & Meyer.