

DÍAZ RODRÍGUEZ, LOURDES

DEPARTAMENTO DE TRADUCCIÓN Y FILOLOGÍA, UNIVERSIDAD POMPEU FABRA, BARCELONA

CLIC-UNIVERSIDAD DE BARCELONA \*1\*2

## GRAMÁTICAS NUCLEARES DE RAÍZ PRAGMÁTICA (*CORE GRAMMARS*) BASADAS EN CORPUS ORALES DE ESPAÑOL NATIVO Y NO NATIVO. A PROPÓSITO DE LA INVESTIGACIÓN DE LA PRODUCCIÓN ORAL POR *ESCENARIOS* PROPUESTA EN LOS MODELOS DEL CONSEJO DE EUROPA Y EL *MCE*

### BIODATA

Lourdes Díaz es profesora de lengua española del Departamento de Traducción y Filología de la Universidad Pompeu Fabra, y de Gramática pedagógica y Adquisición de Español como L2/LE en diversos cursos de posgrado: UIMP (desde 1991), Universidad de Barcelona / UB Virtual (desde 1990), Universidad de Zaragoza (desde 1998). Colabora también con la EOIBD, el Instituto Cervantes y universidades y centros extranjeros.

Investiga la adquisición de segundas lenguas en el aula en equipos de investigación financiada (que ha dirigido en seis ocasiones), en colaboración con universidades nacionales y extranjeras, que han dado origen a numerosos artículos, tesis de máster y de doctorado. Es investigadora del CLIC-UB (Centre de Llengües i Computació). Le preocupa la aplicación de los resultados a la docencia en ELE, que ejerce ininterrumpidamente desde 1988. Ha sido organizadora y directora de cursos de Extranjeros y de Formación de profesores de ELE en la Universidad Internacional Menéndez Pelayo (Santander 2007-2008), directora del *Máster de Formación de Profesores de ELE* UIMP-Instituto Cervantes en Santander (2008-2012) y vicerrectora de Lenguas e internacionalización de la misma universidad (2011-2013). Actualmente es vicedecana de Relaciones internacionales y movilidad en la Facultad de Traducción de la UPF.

Entre sus publicaciones sobre ELE se encuentran: *La Destreza escrita* (2003) con M. Aymerich (Edelsa, Madrid); con Martínez y Redó: *Guía rápida de contenidos lingüísticos para el MCE* (Printulibro; Barcelona, 2007); *La adquisición del español: estudio de casos* (Barcelona, 2007); *Guía de contenidos lingüísticos para ELE* (Octaedro, Barcelona 2011) también con Martínez y Redó; *Lenguas de especialidad y su enseñanza* (Monográfico de marcoELE, 2010) y *Estudios sobre aprendizaje y adquisición de lenguas y sus*

<sup>1</sup> Quiero agradecer a Mónica Caballero, compañera en Clic y doctoranda, su generosidad al compartir los datos de este microcorpus para realizar la *toy-core-grammar*, este esbozo nuclear. También quiero dejar constancia de mi gratitud hacia M. Taulé y M. A. Martí (Clic) y A. Ortega y B. North (Eaquals), por sus útiles puntos de vista acerca de las gramáticas nativas y no nativas, en general y las *core* en particular.

<sup>2</sup> Esta investigación se enmarca en el proyecto DIANA (TIN-2012-38603-C02-02), financiado por el MINECO, cuyos investigadores principales son: Paolo Rosso (UPV) y M. A. Martí (UB).

*contextos* (marcoELE 2012). Ha publicado colaboraciones en volúmenes especializados en adquisición de L2 o en metodología para la enseñanza de lenguas: en *The role of formal features in Second Language Acquisition*, editado por J. Liceras - H. Zobl y H. Goodluck, (L. Erlbaum, 2008); en Rodríguez Halffter, E. y Á. Ortega: *Desarrollo de competencias en las destrezas de mediación* (Aulas de Verano, MEPSyD, 2008); en Ruiz Fajardo, G.: *Methodological Developments in Teaching Spanish as a Second and Foreign Language* (Cambridge, Cambridge Scholars, 2012). La investigación más reciente de su equipo se centra en el uso de corpus nativos y no nativos para la descripción pragmagramatical y de la competencia multilingüe.

#### RESUMEN

Desde que se produjo la eclosión de la pragmática en las descripciones lingüísticas para los profesores de español como lengua extranjera (ELE), estos están cada vez más familiarizados con los modelos de lengua surgidos de *corpus orales* y escritos de nativos y no nativos, frente literarios o a los creados artificialmente para el trabajo en el aula. Los actuales profesores ya se sumergen sin sorpresa en las transcripciones y encuestas anotadas y analizadas por lingüistas y pragmáticos, cuyos símbolos y convenciones apenas si resultan ya disuasorios (ver: Ruiz Fajardo 2012). Desde principios de los 90, los trabajos del equipo de E.A. Briz (1996 en adelante) que contienen datos orales, muestras transcritas y descripción de la lengua hablada o coloquial en Valencia son habituales en publicaciones sobre ELE. Y no solo el corpus ValEsCo sino algunos otros, como los trabajos de C. Fuentes (1996 y ss.) sobre el habla de Sevilla, los de Luis Cortés y su equipo, en Almería<sup>3</sup>, por citar algunos ejemplos relevantes de repertorios de discurso oral espontáneo nativo de español sobre los que operar abriendo el foco.

Datos y método han permitido descubrir una lengua viva y han cambiado notablemente el panorama de la enseñanza de lenguas primeras y, por extensión, segundas. Esto se ha hecho no solo a través de una renovación en el lenguaje descriptivo o la metodología de trabajo, sino también a través de la visualización de los elementos discursivos propios del objeto de estudio. El descubrimiento del papel de las *partículas, enlaces, marcadores, atenuadores, etc.* y su valor en el habla informal, o la distinta distribución y combinatoria de un elemento léxico en el discurso nativo son ejemplos de ello en sus publicaciones (ver las webs de: Ilse, *Oralia*, ValEsCo, etc. Pendiente de desarrollo paralelo está, no obstante, describir la producción no nativa: cuál es la restricción semántica o sintáctica que suponen las gramáticas y lexicones de los aprendices en función de los contextos y combinaciones lingüísticas, o las características de la lengua en la interacción en situación no solo experimental, por poner algunos ejemplos. (Ver: Díaz, L., 2010, *Lenguas de especialidad y su enseñanza, marcoELE*, para algunos ejemplos; Ruiz Fajardo 2012; etc.).

Quede constancia del valor de estas miradas como una ventana *de facto* sobre la lengua real de los nativos, sobre el uso (pragmático) real y el encarecimiento de la urgencia de descripciones y explicaciones similares pero para la L2. El reto, por tanto, está servido. Un punto de partida válido lo constituye la aplicación de los métodos de trabajo con corpus a nuevas situaciones de interacción específicas –escenarios– en esta ocasión ya entre nativos y no nativos o exclusivamente entre no nativos, a través de las que empezar a atisbar cómo pueden ser las gramáticas nucleares (*core grammars*) de esas situaciones con vistas a que constituyen el modelo específico para las aulas de lenguas extranjeras.

**PALABRAS CLAVE / KEY WORDS:** pragmagramática, escenarios y MCER, gramática nuclear, gramática basada en corpus, comparación corpus L1/L2, pragmatic grammar, CEFR escenarios, core grammar, corpus driven grammar, L1/L2 corpus comparison

---

<sup>3</sup> <http://www.grupoilse.org/> tiene un corpus de 108 entrevistas semidirigidas y las publicaciones resultantes, en ORALIA (revista centrada en la oralidad) u otras.

## 1. INTRODUCCIÓN

Como docentes investigadores en lingüística aplicada a la enseñanza de lenguas extranjeras, valoramos muy positivamente el cambio producido en los modelos de descripción de la producción oral nativa y no nativa. También, y muy especialmente, en el ámbito de la enseñanza de lenguas segundas, más incipiente, donde desde fines de los 70 emerge la lengua oral coloquial como objeto prioritario en los trabajos del Consejo de Europa, desde donde se erige en pilar fundamental para los descriptores destinados a L2 y como fuente para extraer repertorios de modelos de estrategias y situaciones de uso en la comunicación no nativa.

Valoramos también el cambio de perspectiva en las herramientas y enfoques descriptivos, como los cada vez más pragmagramaticales que han permitido abordar la caracterización del coloquio (la etiqueta es importante) desde la extracción de los datos (entrevistas secretas en muchos casos) hasta el análisis de los mismos para desvelar el uso espontáneo de la lengua. En 1996, al hilo de la presentación de su *Español coloquial. Esbozo de pragmagramática*, el profesor E. A. Briz, afirmaba tajantemente que la conversación es el ámbito más apropiado para el estudio del uso de la lengua en la comunicación y que ése era el análisis pertinente y pendiente (de ahí el carácter de *Esbozo* de la obra). Es obvio que veinte años más tarde, las cosas se han movido sustancialmente. Hay ya muchísimos estudios, descripciones<sup>4</sup> y datos abundantes de corpus para sustentarlos.

---

<sup>4</sup> Diana Bravo y E.A. Briz (Eds.) 2004. *Pragmática sociocultural: estudios*. Barcelona: Ariel. Se trata de una recopilación de trabajos que constituyen dos muestras de lo esbozado en 1998 por el equipo. Ver más en:

Paralelos a los estudios de corpus de L1, se han recogido y analizado también algunos corpus ya específicamente para español como L2/LE (Sacodeyl, CORAL-ELE, COLAm, CEDEL2<sup>5</sup>, etc.). También las propuestas de los expertos en estándares y niveles han avanzado en la integración de los datos empíricos como respaldo y guía de sus propuestas de niveles, comportamientos (Can Do), etc., como muestra la velocidad en la producción de documentos de referencia del Consejo de Europa y sus agencias. Equals- British Council, por ejemplo, han lanzado entre 20010 y 2011 un trabajo aglutinador de descriptores de niveles (que incluye los intermedios A1+, A2+, etc.),

---

<http://www.valesco.es/web/miembros/brizbiblio.php>. Asimismo, Ruiz Fajardo, G. (2012), en *Methodological developments in teaching Spanish as a Second and Foreign Language* (cap. 3) y en Díaz, L. (Coord.) *Lenguas de especialidad y su enseñanza*. 2011(<http://marcoele.com/television-en-el-aula>), constituye un ejemplo de aplicación de la metodología al análisis de *input* real procedente de la red.

<sup>5</sup> SACODEYL forma parte de un proyecto europeo (<http://www.um.es/sacodeyl>). Se trata de un corpus basado en entrevistas estructuradas, grabadas en vídeo, llevadas a cabo a estudiantes de entre 13 y 18 años. Las entrevistas se han transcrito y anotado de forma amplia, con el objetivo de que sean útiles en la enseñanza de lenguas. Son producto de un proyecto EU Socrates-Minerva "System Aided Compilation and Open Distribution of European Youth Language", liderado por la Universidad de Murcia.

COLAm ([http://www.colam.org/om\\_prosj-espanol.html](http://www.colam.org/om_prosj-espanol.html)) es una base de datos de habla informal de jóvenes y adolescentes hispanohablantes, transcrita y acoplada al sonido accesible en Internet para la investigación. Empezó en el año 2001. Ha sido subvencionado por la Facultad de Humanidades de la Universidad de Bergen y por la fundación Meltzer de la misma universidad. Recoge muestras de habla de Madrid y de distintas ciudades de habla española y sigue el mismo patrón que el corpus COLT (para el inglés). C-ORAL-rom (Moreno y Urresti 2004) también procede de un proyecto europeo de corpus de lenguas románicas, pensado para la explotación didáctica

CEDEL2 es un corpus del español escrito de aprendices de ELE desarrollado por la UAM y la Universidad de Granada:

<http://www.uam.es/proyectosinv/woslac/cedel2.htm>.

funciones, *gramáticas nucleares* que amplían el alcance del *Marco* (para el inglés). También proponen sistemas de trabajo, como el uso de *escenarios* (ver Díaz, 2013 en este número), para vertebrar e incluir todos estos aspectos descriptivos vinculados al contexto y las necesidades comunicativas específicas sin disonancias.

A falta de un desarrollo de propuestas aplicadas de escenarios y sus explotaciones lingüísticas específicas para el español no nativo y para el aula –basadas en datos empíricos amplios y más allá de las intuiciones–, presentamos aquí una propuesta de gramática nuclear B1 resultado de la puesta en práctica de un escenario (“Organizar y un viaje en tren”). Nos centramos, específicamente, en la *compra del billete* (interacción oral), para tener un microescenario y ensayar una microgramática nuclear. La reducción se justifica porque pretendemos fundamentar la propuesta en datos cien por ciento empíricos. Usamos tanto datos de corpus extraídos de la red, herramientas de análisis de carácter gratuito accesibles en la red (no específicas como las de algunos de los proyectos citados<sup>6</sup>) como datos de corpus específicamente creados para la investigación (simulaciones y grabaciones secretas). Analizamos todo ello con una selección de herramientas accesibles y fáciles de usar y apuntamos algunas implicaciones didácticas.

## 2. LOS INVENTARIOS NUCLEARES DESDE LA PERSPECTIVA DEL MARCO COMÚN EUROPEO DE REFERENCIA Y SUS APLICACIONES: MICRO-GRAMÁTICAS NUCLEARES

Es bien sabido que entre los desarrollos del *MCER*, directamente implicados en la enseñanza y la formación de profesorado, ha

---

<sup>6</sup> Clic trabaja con analizadores propios, etiquetadores, etc. más refinados que los que usamos aquí para la elaboración de esta *toy core grammar*.

estado el suministro de recursos y guías orientadoras sobre los contenidos lingüísticos y la forma de abordarlos (ausentes del documento *Marco*). Para el inglés, Brian North y su equipo, dentro de EAQUALS, han desarrollado un documento destinado a la formación de profesores<sup>7</sup>, en el que recogen y sistematizan los contenidos lingüísticos por niveles y proponen una metodología de trabajo por *escenarios*, para distribuir dichos contenidos por destrezas y plantear la forma de abordarlos en el aula. Básicamente, el trabajo por escenarios permite diferenciar, mediante los CAN DO, qué contenidos y habilidades se presentan como pasivos (input) o como activos (*output* esperado), por decirlo de una forma harto simplificada para cada nivel / propuesta / programación. También permite visualizar los exponentes lingüísticos como un *continuum* (no como actores de reparto obligado en uno u otro nivel y de forma excluyente) lo mismo que las funciones, léxico y exponentes. Este enfoque naturaliza la reutilización de contenidos (como *continuum*) y pone el acento en la necesidad de destacar y trabajar significativamente para cada secuencia del escenario un subconjunto de ellos, altamente contextualizado y centrado en unas necesidades (comunicativas) reales y fácilmente reconocibles, flexibles. Para guiar la planificación del profesor y apoyar el descubrimiento autónomo del aprendiz (autoaprendizaje), los investigadores de EAQUALS y los grupos de trabajo del CE actualizan continuamente el *English Profile Project*<sup>8</sup>. En su metodología de

---

<sup>7</sup> *Core inventory for General English*. British Council-Eaquals, 2010.

<sup>8</sup> <http://www.englishprofile.org> Los objetivos, según la propia web, son: “to create a ‘profile’ or set of Reference Level Descriptions for English. These will provide detailed information about what learners ‘can do’ in English at each of the six levels of the Common European Framework of Reference (CEFR), offering a clear benchmark for progress for English language learners. This site gives the background to the English Profile programme and provides access to our resources, including the English Vocabulary Profile, which is a searchable listing of the words and phrases typically produced by English language learners at each level of the CEFR. It provides information at word and sense level, as well as examples of the word in use by real learners, extremely useful for anyone in the ELT community”.

trabajo establecen los siguientes pasos (North, Ortega y Sheehan 2010: 9<sup>9</sup>):

1. Recogida de datos y análisis (muestras, descriptores del MCER, sílabos y manuales): permiten compilar las coincidencias en lo que se considera central (*core*) y menos central para hacer el inventario.
2. Los comités que son, a la vez, agencias examinadoras (Cambridge ESOL, Trinity, etc.) dan también *input* de puntos débiles en el aprendizaje. Se enriquece el inventario con las muestras de alumnos.
3. Una vez recogido el inventario, se escriben los exponentes y se revisan y discuten por los equipos de expertos internacionales.
4. Se identifican los tipos de textos fuente para cada uno de los niveles (géneros textuales ideales como *input* para cada función / exponente).
5. Para que los contenidos tengan UNA RELACIÓN EVIDENTE Y EXPLÍCITA CON LAS NECESIDADES REALES del mundo, SE INTEGRAN LOS OBJETIVOS DEL CURRÍCULO, los de la clase y el resultado del trabajo descrito de 1-4 EN UN CONJUNTO DE ESCENARIOS que constituyen el resultado del consenso sobre actividades y contenidos para cada nivel (Can Do, funciones, exponentes, estrategias, tipos textuales, léxico, áreas temáticas-culturales).

Llegados a los *escenarios* –cuya práctica hemos propuesto en otro lugar<sup>10</sup>– la pregunta pertinente que se harán algunos lectores es ¿qué aportan, o aportarán, desde el punto de vista descriptivo a las descripciones comunicativas al uso?

---

<sup>9</sup> North, B.-A. Ortega, S. Sheehan 2010. Core inventory for General English. British Council- Eaquls, Consejo de Europa. Descargable en:

<http://englishagenda.britishcouncil.org/sites/ec/files/books-british-council-eaquals-core-inventory.pdf>

<sup>10</sup> Ver: Díaz (2013) en este mismo número:

<http://marcoele.com/escenarios-comunicativos>

Claramente, un vistazo a los documentos en que se recogen los elementos lingüísticos (para el inglés) no parece llevarnos muy lejos de las descripciones de sílabos y manuales existentes. No obstante, la insistencia del documento en la significatividad y relación con la realidad comunicativa espolea la curiosidad para ver qué tipos de datos proporciona la puesta en práctica del *escenario* como fuente de evidencia empírica previa a la didactización.

Dar el paso de implementar el *escenario* y usarlo como fuente permite minimizar alguna de las críticas (razonables) que algunos autores de relieve han hecho a las propuestas de Inventarios.<sup>11</sup> En especial permite hacer frente a las dudas acerca de para qué se podrían usar y para qué les servirían a los profesores (en formación o no).

Es cierto que cuando se pone el acento en las “necesidades específicas” y en la “utilidad real” de cualquier práctica –y en el caso de los escenarios, esa es su razón de ser– no solo se trata de valorar la pertinencia de la tarea propuesta, sino también de descubrir bajo el foco de los datos de la *puesta en escena real* cuáles son los puntos realmente relevantes para trabajar. Puntos que pueden ir desde el uso específico de léxico en una situación, frecuencia de formas de cortesía (en función del interlocutor, tipo de intercambio: transacción, intercambio de información, etc.), de estructuras sintácticas o tipos de construcción (sintagmas preposicionales, interrogativas indirectas), etc. hasta una distribución de cualquiera de esos elementos que pueda resultar *caracterizadoras* de una situación.

---

<sup>11</sup> Ver, por ejemplo, la muy razonable de Thornbury, S. 2011.

<http://scottthornbury.wordpress.com/2011/02/06/c-is-for-core-inventory>

Vamos a ilustrar el funcionamiento. Pongamos un ejemplo: si oímos un fragmento de enunciado como este: “*con salida a las 21: 30*”, ¿qué nos sugiere? ¿en qué pensamos? Si no evocamos de manera inmediata una situación como la de “tomar un tren” o bien si nos parece forzado evocarla solo porque nos hayamos enfrentado con ese *SPrep* o sintagma preposicional, probemos con este otro fragmento un poco más largo de enunciado: “*que efectuará su salida por la vía X*”, ¿resulta entonces evidente qué escenario activamos? ¿Solo por “vía” o por la coincidencia de una cierta forma de decir (registro), uso de “efectuar” y futuro? ¿Un poco por la suma de todo? Vayamos un poco más lejos: ¿a qué *tarea* en concreto de la vida real nos remite? ¿A una de producción oral en la que nos vemos *activamente* implicados? ¿A una de comprensión oral? ¿De comprensión lectora? (La respuesta buena es la penúltima).

Para los lectores que con el primer ejemplo tuvieron una intuición acertada: ¿qué elemento les dio la clave a ustedes? Seguramente, el peculiar uso de *salida*, con sus acompañantes o entorno: “con salida” y “a las X”. Los ejemplos antes propuestos son muy distintos, pese a que hay elementos que se repiten, a estos otros: “te espero a la *salida* a las nueve y media”, “hay que iluminar mejor la *salida*”, “esta calle no tiene *salida*”, etc. Es más: si buscamos en un corpus como el de Leeds “salida”, la ocurrencia en contexto “viajar en tren” aparecerá tarde o “muy abajo” en la lista de frecuencias, si es que lo hace (ver infra)<sup>12</sup>, por lo que la consideraremos acepción específica. Constataremos, además, que los resultados de las consultas no incluyen textos semejantes tipológicamente al ejemplo (que corresponde más bien información por megafonía en las estaciones o al discurso de un taquillero profesional). En los corpus

---

<sup>12</sup> “Salida” buscado en <http://corpus.leeds.ac.uk/internet.html> muestra que no hay ninguna ocurrencia relativa a este ámbito en las 100 primeras ocurrencias.

mencionados, la compra de billetes o secuencias de relato o descripción sobre ella no son frecuentes tampoco. Creemos que es a este tipo de proceso al que se refiere la “especificidad” que aduce AEQUALS para justificar LA NECESIDAD DE IMPLEMENTACIÓN de las *core Grammars* EN ESCENARIOS ESPECÍFICOS e insistir en la propuesta de tareas que sean *significativas* en ellos.

En los apartados siguientes presentaremos algunos ejemplos donde se ponga de manifiesto cómo operar con las simulaciones de escenarios para extraer información y datos empíricos relevantes que sustenten las tareas en el aula y adentrarse en los usos específicos que requiere la situación. Esto es: cómo transitar de los puntos 1 a 4 de la propuesta de EAQUALS pero con datos empíricos de respaldo.

### 3. LOS CORPUS DE BASE: LA COMPETENCIA NATIVA Y NO NATIVA JUNTAS Y EN CONTRASTE

La mayoría de los trabajos sobre corpus e incluso la mayoría de los propios corpus aquí referenciados parten explícita y restrictivamente de muestras nativas o no nativas, normalmente de forma excluyente. Asimismo, también suelen diferenciarse en función de la metodología empleada en la recogida de datos, esto es, en el procedimiento de elicitación de los mismos: entrevistas guiadas y semiguías o, por el contrario, producción espontánea y, en ocasiones, grabaciones secretas<sup>13</sup>.

---

<sup>13</sup> *InfoTren-Basurde* (un corpus sobre llamadas a información de RENFE) es un ejemplo de estos últimos. Bonafonte, A y E. Lleida (1999), de la Universidad Politécnica de Cataluña, han publicado diversos trabajos sobre su tratamiento, si bien no con fines didácticos sino para desarrollar aplicaciones en ingeniería. Ver: Caballero, M. et al. 2012 en *marcoELE*: <http://marcoele.com/monograficos/estudios-sobre-aprendizaje-y-adquisicion>.

Las ventajas de los corpus nativos para la obtención de *modelos meta* para el aprendizaje no necesita argumentarse. No obstante, quizá sí resulte necesario destacar que las investigaciones que proporcionan datos de *nativos y no nativos* ante las *mismas situaciones de comunicación* son particularmente valiosas. En primer lugar, desde el punto de vista de la *producción* en sí. La competencia nativa junto a las distintas etapas de la no nativa permite establecer un *continuum* y disponer de un patrón de comparación de naturaleza empírica. En segundo lugar, desde el punto de vista de las estrategias del discurso, el planteamiento y resolución del problema comunicativo (como por ejemplo: ser vendedor de billetes y no cliente; o ser cliente y tener un interlocutor profesional al otro lado) también condiciona que se obtengan patrones diferentes. Estudiar ambos aspectos juntos y por separado, en la arena compartida de un *escenario* común, da una información particularmente rica acerca de qué aspectos de la competencia se movilizan en cada caso (tanto en los sujetos nativos como en los no nativos) y hacer hipótesis sobre cuál es el peso de la "natividad", por un lado, o de la familiaridad con la situación, por el otro; además del papel de la L1 en los no nativos. Tener duplicados los corpus de acuerdo con estas consideraciones permite estudiar la variabilidad prevista para cada escenario en cuestión y dejar expedito el camino para diseñar herramientas para tratarla didácticamente.

Este es el enfoque adoptado en el equipo de Clic para *FerrovieELE*, el corpus oral circunscrito al ámbito *viajar en tren*. En el curso de la investigación llevada a cabo, por tanto, se han usado situaciones simuladas (*role-play*) y reales (grabaciones secretas), cuya comparación permite una extracción de patrones diferenciada. El corpus de la muestra está formado por los siguientes sujetos, residentes todos en España en el momento de la obtención de los datos:

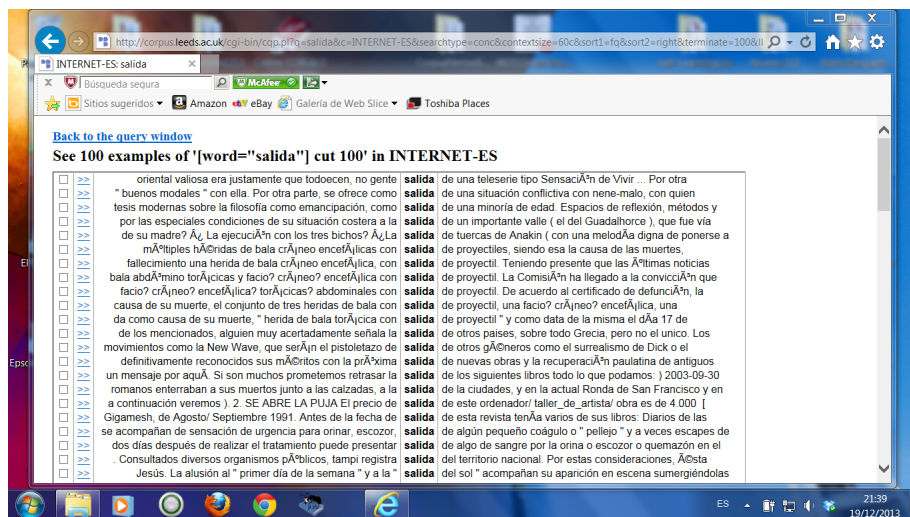
|                                  | NÚMERO DE DIÁLOGOS | NÚMERO DE INTERLOCUTORES | TIPO DE DISCURSO                                       | CARACTERÍSTICAS DE LOS INTERLOCUTORES  | FECHA DE GRABACIÓN |
|----------------------------------|--------------------|--------------------------|--|--|--------------------|
| <i>InfoTren-N(nativos)</i>       | 6                  | 12                       | Conversación telefónica espontánea                     | Nativos, diferentes edades, niveles socioculturales, ambos sexos.              | 1997               |
| <i>Ferrovie-ELE A(prendices)</i> | 6                  | 12                       | Diálogo semicontrolado semiguñado ( <i>role-play</i> ) | No nativos, estadounidenses, jóvenes universitarios, ambos sexos y anglófonos. | 2012               |
| <i>Ferrovie-ELE N(nativos)</i>   | 6                  | 12                       | Diálogo semicontrolado ( <i>role-play</i> )            | Nativos, diferentes edades, niveles socioculturales, ambos sexos.              | 2012               |
| TOTAL                            | 18                 | 36                       |  |  |                    |

De estos, se ha hecho un sub-corpus con el 50% de ellos, con el que llevar a cabo los análisis de este trabajo. El objetivo de la reducción es que permite aplicar con más facilidad las herramientas libres y localizar y presentar de forma no farragosa resultados interesantes para la descripción que podrán ampliarse y tratarse más refinadamente con herramientas más complejas y tratamiento estadístico en una fase posterior.

### 3. 1. HERRAMIENTAS LIBRES

Para el cálculo de frecuencias de palabras y las concordancias de los tres corpus mencionados, hemos recurrido a los programas libres de

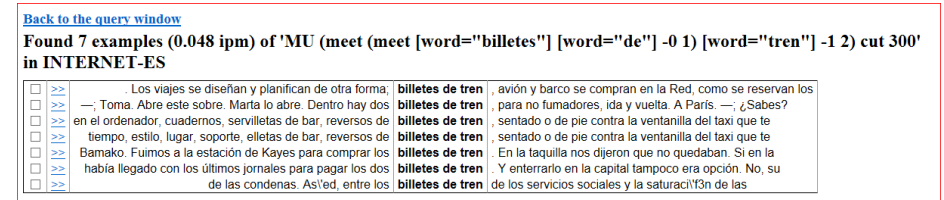
*Antcor*<sup>14</sup>. Asimismo, para establecer la idoneidad de las palabras elegidas, hemos hecho una cata en otro corpus abierto de materiales en la red (LEEDS<sup>15</sup>). Este proporciona una interfaz de consulta con otra herramienta buscador que permite obtener concordancias y colocaciones en diversas lenguas. Las búsquedas en este último, sirven para confirmar la distancia/especificidad del género “compra/transacción en interacción cara a cara” en este dominio específico. El ejemplo de “salida” y el de “billetes de tren” en este corpus presentan el siguiente aspecto:



CAPTURA DE CABECERA DE LEEDS para “salida” (recorte con las 22 primeras ocurrencias)

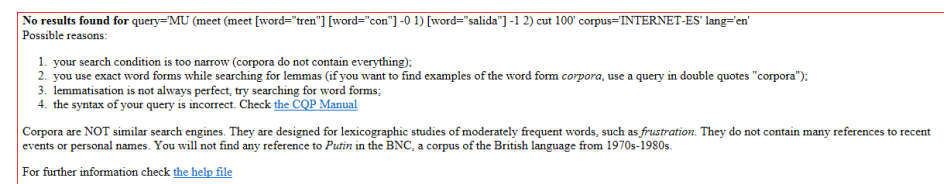
<sup>14</sup> *ANTCONC* es un programa de libre acceso desarrollado por el profesor L. Anthony. ActCon 3.2.4w puede usarse en distintas plataformas o sistemas (Windows, Linux, etc.). [www.antlab.sci.waseda.ac.jp/links](http://www.antlab.sci.waseda.ac.jp/links) permite ver las aplicaciones y publicaciones relacionadas con el software y el modelo de trabajo.

<sup>15</sup> <http://corpus.leeds.ac.uk/internet.html>

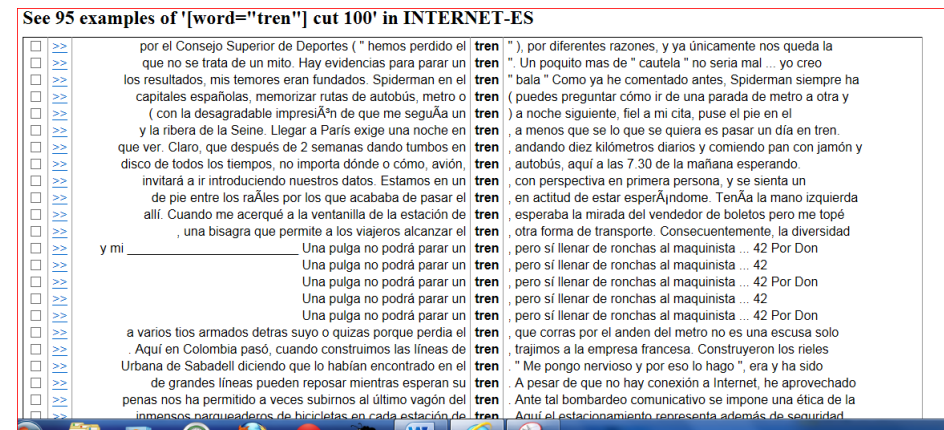


captura de cabecera de leeds para “billetes de tren” (recorte con el total de ocurrencias / primeras ocurrencias sobre 300 de la petición).

Si ponemos solo “tren con salida”, basándonos en lo que observamos en los corpus de nativos, el resultado es este:



Y si ponemos únicamente “tren”, este otro:





En concreto, vemos, el corpus de Leeds permite llevar a cabo una búsqueda de ocurrencias (como la que vimos al inicio para “salida”) sin restricción de dominio, situación o escenario posible. La observación de los resultados en este paso pone de manifiesto el carácter “específico” o “especializado” del léxico de nuestro escenario y justifica el trabajo detallado sobre el mismo que hacemos a continuación. En consonancia con la propuesta de EAQUALS, si estuviéramos en un contexto de formación de profesores o con alumnos con nivel de competencia suficiente, esa consulta en corpus abierto nos permitiría establecer la diferencia entre léxico necesario, “específico” y léxico “usual”. Veríamos cómo se ponen de manifiesto el campo, tenor y modo<sup>16</sup>: el campo (científico, médico, ocio, cultura, blogs personales, literatura) y modo del texto en que se encuentra la palabra, así como el registro y nivel de lengua. Estaríamos, entonces, preparados para observar la variabilidad y comportamiento de ese léxico objeto de atención en una interacción a su vez contextualizada, en contraposición con los corpus reales orales procedentes de entrevistas. En el caso de los corpus de no nativos, el grado de natividad sería *capturable* a través de la constatación de una variabilidad mayor en unos sujetos que en otros, en un tipo o modalidad de texto o en otro. La propia evidencia empírica constituiría un input de segundo nivel sobre el que el aprendiz realizaría operaciones de análisis metalingüístico debidamente soportadas por los datos.

---

<sup>16</sup> M. A. K. Halliday, A. McIntosh y P. Stevens (1964)  
[http://cvc.cervantes.es/ensenanza/biblioteca\\_ele/diccio\\_ele/diccionario/registro.htm](http://cvc.cervantes.es/ensenanza/biblioteca_ele/diccio_ele/diccionario/registro.htm)

## 4. ANÁLISIS DE DATOS: ALGUNAS LÍNEAS ILUSTRATIVAS DE LAS DIFERENCIAS

Para estudiar las diferencias en la *minicore grammar* o minigramática nuclear de la tarea “compra de billete de tren en interacción cara a cara”, hemos seleccionado una serie de palabras clave relevantes en la transacción y de alta frecuencia en los corpus orales FerroviELE, en las producciones orales. Se trata de las palabras: “Información”, “reserva”, “compra”, “pago, pagar”, “precio” y “tren”, cuya recurrencia en las funciones que permiten llevar a cabo la tarea las hace candidatos idóneos. La frecuencia, contexto de aparición, agrupación habitual y turno/ función en que aparecen podrán obtenerse semiautomáticamente con *Antcor*. *Antcor* nos permitirá, en una primera cata, y sin necesidad de tratar los datos (etiquetarlos, etc.) lo siguiente:

- Obtener la posición de las palabras en la lista de palabras de cada uno de los corpus, de forma que veamos si la FRECUENCIA difiere entre nativos y no nativos, o entre los nativos también.
- Obtener EL PATRÓN DE CONSTRUCCIÓN en que aparece la palabra clave y compararlo entre los grupos.
- Obtener información sobre el turno / función en que aparece, para compararla entre grupos.

### 4.1. FRECUENCIAS COMPARADAS

A modo de ejemplo, en relación con la FRECUENCIA de aparición de una palabra buscada en uno de los tres corpus FerroviELE, los resultados presentarían este perfil:

#### NATIVOS EN SIMULACIÓN / FRECUENCIAS

Total No. of Word types: 374

Total No. of Word tokens: 1216

Las diez palabras plenas más frecuentes en este corpus son:

*Tren* (posición 12, 14 ocurrencias)

*Ida* (posición 18, 11 oc.)

*Tiene* (posic. 18, 11 oc.)

*Usted* (posic. 19, 11 oc.)

*Días* (posic. 20, 10 oc.)

*Tarjeta* (posic. 23, 10 oc.)

*Viajar* (posic. 24, 10 oc.)

*AVE* (posic. 25, 9 oc.)

*Billete* (posic. 27, 9 oc.)

*Preferente* y *Business* (posic. 31 y 33 con 9 y 8 ocurrencias respectivamente).

Quedará como rasgo característico de este grupo la presencia del “usted” explícito, ausente en los otros dos.

#### NATIVOS SIN SIMULACIÓN / FRECUENCIAS

Total No. of Word types: 383

Total No. of Word tokens: 1567

Las diez palabras plenas más frecuentes en este corpus son:

*Sí/No* (posición 11 y 13, con 28 y 23 ocurrencias)

*Mañana* (posic. 15 con 18 oc.)

*Vale* (posic. 19 con 16 oc.)

*Vuelta* (posic. 20 con 16 oc.)

*Hay* (posic. 21 con 14 oc.)

*Ida* (posic. 22 con 14 oc.)

*Sale* (posic. 24 con 13 oc.)

*Noche* (posic. 31 con 11 oc.)

*Talgo* (posic. 33 con 11 oc.)

*Billete* (posic. 46 con 7 oc.)

Hay una vacilación recogida (“Hum”) en la posición 43 con 8 ocurrencias, aunque este corpus no lo ha recogido ni anotado Clic-UB y no tenemos información sobre cómo se trataron estos rasgos. Llamamos la atención sobre ello porque es un rasgo compartido con los no nativos.

*Pasaje* y *Estación* se encuentran en las posiciones 260 y 310 respectivamente, con una sola ocurrencia. La distancia respecto a los datos de la búsqueda en el corpus de Leeds es evidente.

#### NO NATIVOS EN SIMULACIÓN / FRECUENCIAS

Total No. of Word types: 268

Total No. of Word tokens: 1109

De entrada, el total de tipos de palabras es 100 puntos más bajo que cualquiera de los dos corpus nativos. En cuanto a las ocurrencias (*tokens*), están más cerca del otro grupo en simulación o *role-play*.

Las diez palabras plenas más frecuentes en este grupo son:

*Eh, Hum, Uhm* (entre las cinco primeras en frecuencia, con 84 ocurrencias). ¡Revelador!

*Tren* (posic. 11, 20 ocurrencias)

*Ida* (posic. 13 con 18 oc.)

*Hay* (posic. 16 con 15)

“*sí/para*” (posiciones 20 y 21 con 28 oc.)

*Turista* (posic. 22 con 14 oc.)

*Quiero* (posic. 24 con 13 oc.)

*Billete* (posic. 26 con 11 oc.)

*Cama* (posic. 27 con 11 oc.)

*Vale* (posic. 30 con 10 oc.)

*Tengo* (posic. 33 con 9 oc.)

*Individual* y *mañana* (posic. 36 y 37 con 8 oc.)

*Directo, noche* y *barato* están entre la 43 y la 49, con 7 y 6 ocurrencias.

*Sentado, comprar, cuesta, opciones, preferente, tiene, restaurante, AVE, cena, coche, desayuno, es, está, informarme o llegada* entran entre el 50 y el 98, lo que los aleja de lo que ocurre con los nativos.

*Tarjeta y viajar* aparecen rebasada la posición 100 de la lista.

*Información* en la 127 (con 2 ocurrencias). *Billetes* en la 165 (con 1 ocurrencia).

## 4.2. AGRUPACIONES DE PALABRAS (*CLUSTERS*): RESULTADOS POR GRUPOS PARA *TREN*

En relación con los *clusters* o agrupaciones de palabras, para ver patrones, buscamos los que contienen la palabra *TREN* en los tres corpus y observamos que el sintagma donde aparece presenta los siguientes patrones, con diferencias entre operador (taquillero) y cliente.

*Nativos no simulados*: usan *tren* con el determinante “un” (para presentar la entidad o sustantivo con carácter indeterminado) de forma consistente por parte del rol del operador o taquillero. La indeterminación también se expresa mediante la ausencia de determinantes (sustantivos escuetos) para las preguntas del cliente. La otra agrupación que emerge es la del grupo “el primer *tren*” también para las preguntas del cliente, fundamentalmente.

*Nativos simulados*: usan *tren* con una variedad amplia de determinantes, tanto el operador como el cliente. Los determinantes que predominan son “el / ese / este” (repertorio muy amplio) y su presencia responde a factores de coherencia discursiva, anáforas y control de la información dentro del turno o en turno distinto. Este comportamiento difiere tanto del grupo no nativo como del no simulado nativo.

Los datos que siguen ilustran este patrón para los determinantes que concurren con *TREN*:

### CLUSTER\_\_TREN\_\_ NATIVOS NO SIMULADO

Total No. of Cluster types: 17

Total No. of Cluster tokens: 17

|    |   |  |
|----|---|--|
| 1  | 1 | a ver; <b>hay un tren</b>                  |
| 2  | 1 | C1: #1 Quisiera preguntarle, un tren       |
| 3  | 1 | <b>el primer tren</b>                      |
| 4  | 1 | <b>En qué tren</b>                         |
| 5  | 1 | <b>es un Euromed, es un tren</b>           |
| 6  | 1 | es un tren                                 |
| 7  | 1 | Euromed, es un tren                        |
| 8  | 1 | <b>hay un tren</b>                         |
| 9  | 1 | interesaría saber el <b>primer tren</b>    |
| 10 | 1 | me interesaría saber el <b>primer tren</b> |
| 11 | 1 | O1: ¿En qué tren                           |
| 12 | 1 | preguntarle, un tren                       |
| 13 | 1 | Quisiera preguntarle, un tren              |
| 14 | 1 | saber el <b>primer tren</b>                |
| 15 | 1 | un Euromed, es un tren                     |
| 16 | 1 | vamos a ver; hay un tren                   |
| 17 | 1 | ver; hay un tren                           |

### CLUSTER\_\_TREN\_\_ NATIVOS

Total No. of Cluster types: 199

Total No. of Cluster tokens: 210

|    |   |                                    |
|----|---|------------------------------------|
| 1  | 5 | un tren                            |
| 2  | 3 | el tren                            |
| 3  | 2 | de tren                            |
| 4  | 2 | El tren                            |
| 5  | 2 | es un tren                         |
| 6  | 2 | O4: El tren                        |
| 7  | 2 | tren Alaris                        |
| 8  | 1 | Ah, <b>ese tren</b>                |
| 9  | 1 | Ah, <b>ese tren</b> me             |
| 10 | 1 | Ah, <b>ese tren</b> me parece      |
| 11 | 1 | Ah, <b>ese tren</b> me parece bien |

|    |   |  |
|----|---|--|
| 12 | 1 | <b>billete de tren</b>                 |
| 13 | 1 | billete de tren Barcelona              |
| 14 | 1 | billete de tren Barcelona-Sevilla      |
| 15 | 1 | billete de tren Barcelona-Sevilla para |
| 16 | 1 | C4: Ah, <b>ese tren</b>                |
| 17 | 1 | C4: Ah, <b>ese tren</b> me             |
| 18 | 1 | C4: Ah, <b>ese tren</b> me parece      |
| 19 | 1 | C4: ¿restaurante tiene <b>el tren</b>  |
| 20 | 1 | C4: ¿supongo que <b>el tren</b>        |
| 21 | 1 | C4: ¿supongo que <b>el tren</b> estará |
| 22 | 1 | <b>de tren</b> Barcelona               |
| 23 | 1 | de tren Barcelona-Sevilla              |
| 24 | 1 | de tren Barcelona-Sevilla para         |
| 25 | 1 | de tren Barcelona-Sevilla para mañana  |
| 26 | 1 | de tren prefiere                       |
| 27 | 1 | de tren prefiere, Alaris               |
| 28 | 1 | de tren prefiere, Alaris o             |
| 29 | 1 | de tren prefiere, Alaris o AVE         |
| 30 | 1 | <b>de un billete de tren</b>           |
| 31 | 1 | de un billete de tren Barcelona        |
| 32 | 1 | <b>de un tren</b>                      |
| 33 | 1 | de un tren con                         |

#### CLUSTER TREN NO NATIVOS

1. **tren de Barcelona a Granada -> Rank 1 Freq 2**
2. Ah, prefiere el tren de- → Rank 2, freq 1.....[A partir de este, todos freq 1]
3. Ah, ¿puedo informarme del tren
4. C3: Ah, prefiere el tren
5. C5: Yo quiero tren directo
6. C5: ¿podría informarme del tren
7. del tren Barcelona a Granada
8. del tren de Barcelona a
9. directo tren hum que salida
10. eh del tren Barcelona-Sevilla
11. eh el tren incluida cena
12. el primer tren eh directo
13. el segundo tren directo Linares
14. el tren de Ave y

15. el tren de Barcelona a
16. el tren incluida cena y
17. El tren salida a las
18. el tren tiene muchas prestaciones
19. enlace. Hum el primer tren
20. es un tren con enla
21. esta tren eh salió a
22. este tren eh tiene climatación
23. Guinares y el segundo tren
24. hora es el primer tren
25. Hum el primer tren eh
26. hum el tren de Ave
27. Hum tengo un directo tren

### 4.3. LÉXICO Y COOCURRENCIAS: CONCORDANCIAS Y DIFERENCIAS SEGÚN GRUPO

En relación con las palabras que habíamos considerado en la introducción del apartado 4, elegidas a partir del repertorio funcional (el rotulado de las funciones), estas son las diferencias halladas.

#### 4.3.1. Información

##### NO NATIVOS

Volver y comprar después C3: pero gracias por la **información**. O3: De nada, O3: hasta luego. C3: Hasta luego  
Te y no tengo mucho dinero C5: hum quiero un **eh información** sobre los trenes diferentes. O5: O

##### NATIVOS

1 C2: Hola buenos días. C2: Por favor, **desearía información** sobre los horarios de trenes para viajar a Sevilla

#### NATIVOS NO SIMULADO

1. O1: **Información de Renfe**, O1: buenos días. C1: Buenos días,
2. O1: **Información de Renfe**, O1: buenos días. C1: Hola, C1: buen
3. O1: **Información en Renfe**, O1: buenas tardes. C1: Hola, buenas
4. O1: **Información Renfe**, O1: buenas tardes. C1: Hola, buenas tar
5. O1: **Información Renfe**, O1: buenas tardes. C1: Hola, buenas tar
6. O1: **Información internacional**, O1: buenas tardes. C1:

#### 4.3.2. Reserva

##### NO NATIVOS

Reserva: CERO ocurrencias. Alternativa → **COMPRAR** → 4 ocasiones;  
frec: 69 / 270 tokens

##### NATIVOS

- 1 los datos y le realizamos en este mismo momento la **reserva oportuna**. C4:  
Pues, gracias señorita. C4: Vamo  
UB\_2.txt
- 2 Le doy los datos. O4: De acuerdo Sr. Hervás, **su reserva está conforme**. O4:  
Muchas gracias y buen viaje.  
UB\_2.txt
- 3 ¿game? C4: Buenos días, C4: **necesito hacer una reserva de un billete** de tren  
Barcelona-Sevilla para mañana  
UB\_2.txt
- 4 prensa. C4: Pues entonces, por favor, **hágame la reserva oportuna**. O4:  
¿Qué forma de pago prefiere: tarje  
UB\_2.txt
- 5 no hay problema. O4: Ahora le daré **su código de reserva para** que pueda  
recoger su billete en la estación.  
UB\_2.txt
- 6 a, O4: buenos días. C4: Mire, **quería hacer una reserva**, eh para... un... tren  
de Barcelona a Sevilla.  
O4 UB\_2.txt

#### NATIVOS NO SIMULADO

C1: ¿hoy lo **puedo reservar**? ¿No hay problemas de plazas de momento? O1: ¿Pa

#### 4.3.3. Pago

##### NO NATIVOS

Pago → CERO ocurrencias. Alternativa → **PAGAR**

Total No. of Cluster Types: 2

Total No. of Cluster Tokens: 2

- |   |   |                                 |
|---|---|---------------------------------|
| 1 | 1 | ah, ¿ <b>vas a pagar</b>        |
| 2 | 1 | hum ¿ <b>cómo quieres pagar</b> |

##### NATIVOS

Total No. of Cluster Types: 3

Total No. of Cluster Tokens: 3

- |   |   |   |
|---|---|---|
| 1 | 1 | <b>forma de pago prefiere</b> : tarjeta, metálico |
| 2 | 1 | O4: ¿Qué <b>forma de pago prefiere</b>            |
| 3 | 1 | Qué <b>forma de pago prefiere</b> : tarjeta       |

##### NATIVOS NO SIMULADO

PAGO → CERO ocurrencias. Alternativa → *importe*

1: Y el precio del el **importe** del billete, de ida solo. O1: ¿Me ha dicho para

PAGO → CERO ocurrencias. Alternativa → *precio*

C1: Vale, C1: ¿y el **precio** del billete? O1: A ver por ocho mil pesetas le  
1: De la mañana C1: y ¿el billete de éste qué **precio** tiene? O1: #1 ¿En qué lo  
tasamos, en asiento? #  
nce. C1: Veintidós quince. O1: Sí. C1: ¿Y qué **precio** tiene éste, por favor? O1:  
¿En asiento? C1: Hum  
las siete en punto, no? O1: Hum, hum. C1: Y el **precio** del el importe del billete,  
de ida solo. O1: ¿



3 C4: necesito hacer una reserva de **un billete de tren Barcelona-Sevilla** para mañana. O4: ¿Qué tipo de UB\_2.txt  
4 Barcelona-Sevilla para mañana. O4: ¿Qué **tipo de tren prefiere**, Alaris o AVE? C4: Hum, ¿qué diferencia UB\_2.txt  
5 Mira, mire, hum, precisamente dispone usted de **un tren con salida** a las quince y cincuenta. O4: ¿Lo ne UB\_2.txt  
6 En business, por supuesto; C4: ¿supongo que **el tren estará** climatizado? O4: El tren está climatizado UB\_2.txt  
7 ¿supongo que el tren estará climatizado? O4: **El tren está climatizado** y, además, se le ofrecerá una me UB\_2.txt  
8 Mire, quería hacer una reserva, eh para... un... **tren de Barcelona** a Sevilla. O4: Muy bien, O4: eh ¿ UB\_2.txt  
9 mañana pero no demasiado temprano. O4: **Tenemos un tren a las** ocho treinta que llega ahí a las nueve, a l UB\_2.txt  
10 cuánta, cuánto tarda! O4: Sí, O4: porque **es un tren Alaris**. C4: Es un Alaris. O4: Sí, tarda diez h UB\_2.txt  
11 pero luego también a las ocho treinta **tenemos un tren que llega** a las trece cincuenta y siete minutos, UB\_2.txt  
12 rda cinco horas y veintisiete minutos; éste **es un tren AVE**. C4: Éste me interesa mucho más. O4: Perfec UB\_2.txt  
13 o, no, no. C4: Vale, C4: ¿restaurante **tiene el tren?** O4: El restaurante... tiene, sí, tiene servicio UB\_2.txt  
14 , sí. # C4: #2 Vale, esto lo entiendo. # O4: **El tren tiene** servicio de internet, O4: sí sí por supu UB\_2.txt

#### NATIVOS NO SIMULADO

1 O1: Hola. C1: #1 Quisiera preguntarle, **un tren que** sale para dirección Murcia. # O1:  
2 C1: Sí eso. O1: #1 A ver, vamos a ver; **hay un tren que sale** de Barcelona-Alicante por la mañana, un  
3 De Barcelona a Sevilla? C1: Exacto. O1: ¿En **qué tren? C1: Sale**, por lo que tengo aquí puesto, de Sants  
4 O1: Mire, me interesaría saber el primer **tren que sale** dirección Alicante, el sábado día veinticinco  
5 ¿Ocho mil? # O1: Éste es un Euromed, **es un tren de alta velocidad**, O1: ¿eh? C1: Ah, vale, vale

#### 4.5. SÍNTESIS

A la vista de los datos analizados, se pone de manifiesto que hay rasgos coincidentes en los corpus en una relación dos a dos: por un lado, entre los simulados (nativos y no nativos); por el otro, entre los nativos.

Estos resultados coinciden con lo observado en Caballero-Díaz-Taulé (2012), y apuntan a la necesidad de profundizar en el estudio de las coincidencias y peculiaridades entre los dos grupos de sujetos y situaciones. Asimismo, apuntan a la pertinencia de considerar en detalle los aspectos de índole más lingüística de los datos y en ampliar la aplicación de la metodología a nuevos escenarios para desarrollar más microgramáticas nucleares.

Desde el punto de vista práctico, se pone de manifiesto que los *outputs* obtenidos de los análisis de los corpus y los corpus mismos constituyen un material muy útil para llevar al aula casi tal cual (tanto por formato como por la presentación de colocaciones y concordancias). El tipo de reflexión que puede generarse en el aula, es compatible y se puede reforzar con el uso de otros recursos

electrónicos afines, como diccionarios con estructura semántica y multilingüe (EuroWordnet, por ejemplo).

Sin ánimo de resultar exhaustivos, creemos que las búsquedas realizadas permiten dibujar las líneas maestras descriptivas y de trabajo del escenario de una forma que refleja el uso real y los conocimientos no solo de los nativos sino de los no nativos competentes en esa tarea. También permiten capturar la variabilidad e iluminan zonas que hubieran quedado oscuras sin un enfoque como el que hemos aplicado.

Creemos que la propuesta presentada permite satisfacer, además, las recomendaciones de buenas prácticas que hacen North et al. (2010: 20), que presentamos adaptadas y con las que cerramos la reflexión.

| DO / HAZ  | DON'T / NO HAGAS  |
|---|---|
| Identifica conscientemente las necesidades del aprendiz respecto al mundo real y haz derivar el contenido didáctico de contextos genuinamente procedentes del mundo real. | No enseñes los puntos difíciles de la lengua y la gramática "porque están ahí".   |
| Haz lluvias de ideas y elabora tus propios escenarios basados en el MCER con tus colegas y alumnos.   | No construyas tus actividades pedagógicas o de aula como escenarios del mundo real (en el sentido de que no son autoelocuentes).                                  |
| Complementa los puntos aportados por el Inventario con las cosas que sepas que necesitan tus estudiantes en tu contexto.  | No pienses que tienes que enseñar solo lo que aparece en el Inventario ni tampoco que has de enseñar todo lo que está en el Inventario (elige según tu contexto). |

|   |   |
|---|---|
| Adapta los exponentes a tu propio contexto cambiando nombres propios, de lugar, actividades, etc. | No hagas <i>recorta y pega</i> de listas y exponentes de forma irreflexiva y acrítica para hacer tu sílabo. |
| Proporciona a los aprendices exponentes lingüísticos contextualizados siempre que sea posible.    | No hagas que tus estudiantes se aprendan listas de exponentes. No es esa su función.                        |

## 5. CONCLUSIONES

En este trabajo hemos presentado un análisis de una muestra de datos empíricos procedentes de la puesta en práctica de un escenario "Organizar y realizar un viaje", más específicamente, de la tarea "compra de un billete de tren". Hemos identificado algunos rasgos de la especificidad de la situación no evidentes para los enfoques comunicativos tradicionales. Hemos señalado puntos novedosos de interés, de complejo descubrimiento y trabajo en el aula desde el punto de vista lingüístico (léxico, sintáctico), pragmático o discursivo (cohesión a través de determinantes, anáfora), y cuyo interés radica en que permiten graduar la complejidad creciente de la competencia gramatical y hallar pautas para trabajarla significativamente, de acuerdo con las recomendaciones de EAQUALS y el Consejo de Europa. Asimismo, hemos aportado un tipo de corpus original (respecto a los referenciados en el apartado 1), formado por nativos y no nativos enfrentados a la misma tarea, más un grupo de control donde la tarea no es simulada. Para su tratamiento y extracción de microgramática, hemos presentado un ejemplo de análisis de datos con herramientas libres, a fin de estimular en el profesorado las prácticas de investigación-acción en esta línea y contribuir a la transferencia de resultados y praxis.



Con el análisis de la micromuestra preparada para la investigación, se ha puesto de manifiesto que es posible localizar y trabajar razonablemente aspectos léxico-semánticos, gramaticales y pragmáticos de relevancia máxima para el escenario, y cuya contextualización en concordancias, etc. hace evidentes y elocuentes las reglas de uso que los rigen.

Observaciones y descripciones como las aquí recogidas permiten, a nuestro entender, investigar y elaborar *microgramáticas* emanadas de los escenarios que dirijan el trabajo del aula y cierren el círculo virtuoso que proponen los expertos (North, Ortega, Sheehan, pero también Thornbury, por ejemplo). Asimismo, concluimos, permitirán enlazar elegantemente con las pragmagramáticas de corpus multilingües, completando el trayecto de lo experimental a lo aplicado y de la teoría al aula.

FECHA DE PUBLICACIÓN: 20 DE DICIEMBRE DE 2011