

Versión aceptada

Pérez-Paredes, P., Sánchez, P., Aguado, P., & Criado, R. (2006). El ámbito de la oralidad en la investigación lingüística anglosajona reciente. *Revista Española de Lingüística Aplicada/Spanish Journal of Applied Linguistics* (RESLA/SJAL) 19, 163–177. ISSN 0213-2028

EL ÁMBITO DE LA ORALIDAD EN LA INVESTIGACIÓN LINGÜÍSTICA ANGLOSAJONA RECIENTE¹

Pascual Pérez-Paredes, Purificación Sánchez, Pilar Aguado, Raquel Criado

Universidad de Murcia

RESUMEN. *Los estudios sobre oralidad han adquirido una enorme relevancia en las últimas décadas coincidiendo con la eclosión de las investigaciones con corpus lingüísticos y la creación de herramientas de tratamiento textual cada vez más sofisticadas. En este trabajo llevamos a cabo una labor exhaustiva de análisis de artículos publicados en quince revistas internacionales de máximo prestigio y relevancia en el área de la lingüística aplicada durante los años 2004 y 2005 para determinar cuáles son las líneas de trabajo más notables en este campo y así valorar en qué punto se encuentran las investigaciones en este área. A efectos de estudio hemos clasificado los trabajos analizados en tres grandes bloques dentro del campo de la oralidad: estructuras mononucleares, estructuras complejas y unidades sintácticas y pragmáticas. Lo primero que cabe resaltar es la variedad en cuanto a los textos analizados, así como a los contextos de estudio y lenguas estudiadas. Sí existe, sin embargo, homogeneidad en cuanto al tipo de análisis llevado a cabo, que es fundamentalmente cualitativo para las estructuras mononucleares y cuantitativo para las complejas. Las unidades sintácticas y pragmáticas, sin embargo participan de ambos tipos de análisis. Una de las carencias detectadas en nuestra investigación es la ausencia de trabajos publicados que aborden un estudio longitudinal utilizando muestras de los mismos informantes en puntos temporales diferentes.*

PALABRAS CLAVE. Oralidad, investigación, análisis multidimensional, corpus.

ABSTRACT. *Studies on orality have acquired great relevance in the last decades along with the appearance of corpus linguistics research and the creation of increasingly more sophisticated textual analysis tools. In this work, we have carried out an exhaustive analysis of articles published in fifteen international journals of the highest prestige and relevance in the area of Applied Linguistics during the years 2004 and 2005. Our purpose is to determine the most outstanding lines of work in orality are and hence to uncover at which point the research in this area stands. For the purpose of our study we have classified the works analysed in three main blocks: mononuclear structures, complex structures and syntactic and pragmatic structures. The first thing that needs highlighting is the variety of texts under scrutiny, as well as the contexts of study and the languages examined. Nonetheless, there exists homogeneity regarding the type of analysis employed, which is fundamentally qualitative for the mononuclear structures and quantitative for the complex ones. Within the latter, both types of analysis are involved in the syntactic and pragmatic units. One of the shortcomings we found is the fact that none of the published works is based on a longitudinal study with samples of the same informants at different time points.*

KEYWORDS. Orality, research, multidimensional analysis, corpus.

1. EL CONCEPTO DE ORALIDAD

El análisis de la oralidad se puede realizar desde múltiples perspectivas, por lo que es preciso acotar el sentido que le daremos al término en el presente trabajo. Abascal (2004: 9) apunta que el concepto *oralidad* está siendo utilizado en distintos ámbitos disciplinares con significados no coincidentes y que la labor definitoria del término resulta hasta el momento insatisfactoria. La misma autora sostiene que una aproximación general a la oralidad supone referirse a uno de los dos modos (oral y escrito) en que se nos ofrece el lenguaje a los individuos de las sociedades letradas, y que este enfoque deja translucir tres contraposiciones, a saber, (a) la oposición de dos medios del lenguaje (Crystal (2001) habla de cuatro medios), (b) la oposición de dos modelos de lengua o dos normas y, finalmente, (c) la oposición de dos modos de interacción social. Estos contrastes nos remiten necesariamente a las diferentes peculiaridades que determinan los medios oral y escrito: contexto compartido vs. carencia de contexto compartido; edición en línea vs. edición cerrada; abundancia de contexto comunicativo vs. merma del mismo; mayor variabilidad lingüística vs. menor variabilidad lingüística, etc. El presente trabajo forma parte un proyecto de investigación titulado *Estudio longitudinal y comparativo de la producción oral (oralidad) de aprendices de inglés como lengua extranjera a partir del análisis de un corpus lingüístico oral recopilado y mediante procedimientos de análisis multidimensional*, cuyo fin es analizar precisamente la oralidad del discurso oral de aprendices avanzados de lengua inglesa. Se hace necesario profundizar, por lo tanto, en la naturaleza epistemológica del término.

Los estudios sobre la oralidad en el ámbito lingüístico han girado en torno a cuatro grandes ejes temáticos, a saber, la singularidad del medio oral frente al medio escrito, con aportaciones tanto teóricas como aplicadas de diferentes escuelas lingüísticas; la oralidad en la literatura y la correspondiente diferenciación estilística entre literatura escrita y oral, su diferente *textura*; la esencia fonética de la oralidad, la descripción de las cadenas y unidades sonoras, y, en último término, la variación lingüística. Este ámbito ha recibido una gran atención por parte de los investigadores del análisis del discurso (AD).

Sin ánimo de ser exhaustivos, y citando a Trappes-Lomax (2004), los analistas del discurso sitúan la esencia de la variación lingüística en cuatro

aspectos discursivos² bien diferenciados. En primer lugar, la *interacción*, considerada como el marco en el que se desarrollan el discurso oral y el escrito como proceso. Las investigaciones en este campo han abordado cuestiones tan diversas como la dinámica de las intervenciones de los participantes en el discurso o el papel de las intervenciones en el mismo. Este área ha vivido tal auge que el término *negociación del significado* ha abandonado la escena pragmática y se ha instalado en la pedagogía de lenguas extranjeras. En segundo lugar, el *contexto discursivo* ha protagonizado el debate sobre la variación lingüística, contexto entendido como un factor intrínseco a una situación comunicativa particular: conocimiento compartido, expectativas, situación comunicativa y contexto cultural son algunos de los conceptos que se han examinado con frecuencia bajo este epígrafe. Otro factor arraigado en el AD es *función*. No es posible entender este término sin el anteriormente referido contexto discursivo. Las denominadas funciones comunicativas se desarrollan en contextos discursivos concretos y, viceversa, determinados contextos discursivos favorecen la utilización de un limitadísimo conjunto de funciones comunicativas. Un último factor de interés en el AD es el de la *instrumentalidad* de los recursos lingüísticos del lenguaje, entendidos éstos como las características entonativas y léxico-gramaticales del mismo. Este factor no ha recibido tanta atención en el AD como la relación entre contexto y función por razones que expondremos más adelante. Sin embargo, estimamos que es un factor determinante para el analista ya que, como veremos a continuación, le permite evaluar de forma cuantitativa la manera en la que se determinan los llamados registros y los géneros. Trappes-Lomax (2004: 146) dice al respecto:

Discourse Analysis needs a functional model of language, one that can show how the resources of the language system are organized to meet the needs of “whos and whats” (context-function) in actual communication. Two distinct versions of functionalism can be identified here, which we may call function-external and function-internal. The function-external version is essentially an appropriateness model, derived from Hymes’ theory of communicative competence, which includes knowledge of what is appropriate use of language for a given context-function [...]. The functional-internal version is the systemic model, whose premise [...] is that the lexico-grammatical is organized, through the ideational, interpersonal and textual metafunctions, to meet the intrinsic needs of language-mediated communication in whatever situation. In this

model, the connection to the external is made through the categories of register and genre.

Para este autor, un registro sería una variedad del lenguaje, mientras que un género sería un tipo de acontecimiento comunicativo (*speech event*). Lee (2001), en un estupendo trabajo sobre la caracterización de los géneros, estilos, textos y dominios en el BNC³, coincide con este enfoque. Para él un género se configura a partir de criterios externos, mientras que las tipologías textuales se conforman según las características lingüísticas de orden interno. Por lo tanto, la categoría género nos permite situar los ejes de la investigación en la variabilidad lingüística, mientras que la tipología textual, los textos, el discurso, acapara toda la atención del investigador ya que se convierte en el auténtico protagonista de su análisis.

La corriente de investigación iniciada por Biber (1988) sitúa el análisis de la oralidad precisamente en el último escenario apuntado, esto es, en el análisis del discurso centrado en lo léxico-gramatical y por lo tanto en la categorización del registro o la tipología textual en tanto que variedad del lenguaje susceptible de caracterización intra-textual.

Ya que el trabajo de Biber (1988, 2003) parte de lo puramente léxico-gramatical e intra-textual, se podría criticar su falta de ambición analítica en el plano extra-textual y, por ende, en la relación contexto-función. Sin embargo, es difícil sustentar estos postulados por dos motivos. En primer lugar, hoy en día no existe un modelo unificado de análisis discursivo que integre los cuatro factores apuntados anteriormente. Por lo tanto, cualquier intento de exhaustividad ontológica y epistemológica en el campo que nos ocupa carece de veracidad. En segundo lugar, el trabajo de Biber es un ejemplo de investigación ascendente, *bottom-up*, la cual parte de datos lingüísticos que son metodológicamente pertinentes e instrumentales en una determinada inquisición lingüística. La adecuación de los datos a los objetivos es una de las ventajas de la investigación con corpus lingüísticos: "The choice of corpus to be used in a study depends upon the research questions being asked of the corpus, or the applications one wishes to base upon the corpus" (McEnery 2003: 451). Por otro lado, la introspección no parece una opción viable. Como diría el propio Teubert

(2004: 97): "If we want to find out what words, sentences and texts mean, we should opt for corpus linguistics". Esta metodología es, en consecuencia, de gran utilidad en la caracterización intra- e inter-textual del discurso dentro de un género determinado.

Tras esta brevísima exposición estamos en condiciones de entender las posibilidades de análisis lingüístico de la *oralidad* y las implicaciones que se derivan de la elección de cualquiera de estas opciones, en general, y, si acaso, de la metodología de análisis multidimensional, en particular.

2. METODOLOGÍA Y ÁMBITO DEL ANÁLISIS

En el presente estudio se han analizado quince revistas científicas cuyo ámbito de interés se centra expresamente en algunas de las áreas apuntadas en el apartado anterior y que están relacionadas directamente con el análisis del discurso, a saber, lenguaje, pragmática, comunicación, enseñanza y aprendizaje de idiomas y lingüística. Las revistas analizadas fueron las siguientes: *Applied Linguistics*, *CALL*, *ELT Journal*, *English for Specific Purposes*, *International Journal of Applied Linguistics*, *International Journal of Lexicography*, *Journal of English for Academic Purposes*, *Journal of Pragmatics*, *Journal of Second Language Writing*, *Language and communication*, *Language Learning*, *Linguistics and Education*, *System*, *The Modern Language Journal* y *World Englishes*. Se revisaron los volúmenes publicados durante los años 2004 y 2005 (Tabla 1) con la intención de examinar primero, y analizar y categorizar después, la investigación en el campo de la oralidad para así evaluar las líneas de investigación actuales y ofrecer una taxonomía de los estudios publicados en estas revistas, así como una discusión sobre las perspectivas de investigación futura.

Nombre de la revista	Editorial	Volúmenes y números revisados	Ámbito de la revista tal y como se recoge en su descriptor original
Applied Linguistics	OUP	<ul style="list-style-type: none"> • Año 2004, volumen 25, números 1, 2, 3 y 4 • Año 2005, volumen 25, números 1, 2, 3 y 4 	Research into language with relevance to real world problems, multidisciplinary approaches encompassed by applied linguistics.
ELT Journal	OUP	<ul style="list-style-type: none"> • Año 2004, volumen 58, números 1, 2, 3 y 4 • Año 2005, volumen 59, números 1, 2, 3 y 4 	Teaching English as a second or foreign language, education, linguistics and psychology.
International Journal of Lexicography	OUP	<ul style="list-style-type: none"> • Año 2004, volumen 17, números 1, 2, 3 y 4 • Año 2005, volumen 18, números 1, 2, 3 y 4 	Lexicography, including issues of design, compilation and use, and with dictionaries of all languages.
International Journal of Applied Linguistics	Blackwell Synergy	<ul style="list-style-type: none"> • Año 2004, volumen 14, números 1, 2 y 3 • Año 2005, volumen 15, números 1, 2 y 3 	Mediation between expertise about language and experience of language.
The Modern Language Journal	Blackwell Synergy	<ul style="list-style-type: none"> • Año 2004, volumen 88, números 1, 2, 3 y 4 • Año 2005, volumen 89, 	Research and discussion about the learning and teaching of foreign and second languages.

		números 1, 2, 3 y 4	
Language Learning	Blackwell Synergy	<ul style="list-style-type: none"> • Año 2004, volumen 54, números 1, 2, 3 y 4 • Año 2005, volumen 55, números 1, 2, 3 y 4 	Child, second and foreign language acquisition, language education, bilingualism, literacy, language representation in mind and brain, culture, cognition, pragmatics and intergroup relations.
World Englishes	Blackwell Synergy	<ul style="list-style-type: none"> • Año 2004, volumen 23, números 1, 2, 3 y 4 • Año 2005, volumen 24, números 1, 2, 3 y 4 	Theoretical and applied studies on language, literature and English teaching, with emphasis on cross-cultural perspectives and identities.
English for Specific Purposes	Elsevier	<ul style="list-style-type: none"> • Año 2004, volumen 23, números 1, 2, 3 y 4 • Año 2005, volumen 24, números 1, 2, 3 y 4 	Second language acquisition in specialized contexts, needs assessment, curriculum development and evaluation, materials preparation, discourse analysis, descriptions of specialized varieties of English, teaching and testing techniques, the effectiveness of various approaches to language learning and language teaching, and the training or retraining of teachers for the teaching of ESP.

Journal of English for Academic Purposes	Elsevier	<ul style="list-style-type: none"> • Año 2004, volumen 3, números 1, 2, 3 y 4 • Año 2005, volumen 4, números 1, 2, 3 y 4 	English for Academic Purposes, literacy, (contrastive) rhetorics.
Journal of Pragmatics	Elsevier	<ul style="list-style-type: none"> • Año 2004, volumen 36, números 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 y 12 • Año 2005, volumen 37, números 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 y 12 	Pragmatics, discourse.
Journal of Second Language Writing	Elsevier	<ul style="list-style-type: none"> • Volumen 13, números 1 al 4 • Volumen 14, números 1 al 4 	Second and Foreign Language writing, writing instruction.
Linguistics and Education	Elsevier	<ul style="list-style-type: none"> • Volumen 15, números 1 al 4 • Volumen 16, número 1 	Linguistics and Education, and the application of linguistics, sociolinguistics, psycholinguistics, discourse analysis, social semiotics, conversational analysis and ethnomethodology to educational issues.
Language and Communication	Elsevier	<ul style="list-style-type: none"> • Año 2004, volumen 24, números 1, 2, 3 y 4 • Año 2005, volumen 25, números 1, 2, 3 y 4 	Verbal and non-verbal communication, language and its communicational functions.

System	Elsevier	<ul style="list-style-type: none"> • Año 2005, volumen 32, números 1, 2, 3 y 4 • Año 2005, volumen 33, números 1, 2, 3 y 4 	Applications of educational technology and applied linguistics to problems of foreign language teaching and learning.
CALL	Routledge	<ul style="list-style-type: none"> • Año 2004, volumen 17, números 1, 2, 3, 4 y 5 • Año 2005, volumen 25, números 1, 2, 3, 4 y 5 	Matters associated with the use of computers in language learning (L1 and L2), teaching and testing.

Tabla 1. *Revistas analizadas en nuestra investigación*

Hemos dividido la exposición siguiente en tres grandes apartados que coinciden con la caracterización taxonómica que proponemos en este trabajo. En primer lugar abordaremos el análisis de las estructuras mono-nucleares. A continuación, ofreceremos el análisis de las estructuras léxicas complejas, para terminar con el análisis de las unidades sintácticas y pragmáticas.

3. LA INVESTIGACIÓN DE LAS ESTRUCTURAS LÉXICAS MONO-NUCLEARES

En los últimos dos años la producción científica relacionada con el campo de la oralidad ha sido, cuando menos, interesante, ya que en todos los trabajos revisados subyace la preocupación por la incorporación de las características y peculiaridades del lenguaje oral a la enseñanza de una segunda lengua/lengua extranjera, aspecto bastante desatendido en el aula de idiomas. El análisis de la oralidad en los trabajos que hemos seleccionado nos ha llevado a la clasificación

de los mismos atendiendo al tipo de unidades de análisis lingüístico en el que se centran, ya sean unidades léxicas mono-nucleares, unidades léxicas complejas, o unidades sintácticas y pragmáticas. En los estudios centrados en estructuras mono-nucleares destacan principalmente los usos y funciones de los marcadores del discurso oral, como es el caso de *well*, *actually* y *so*.

De Klerk (2005a) propone una revisión de las características semánticas y sintácticas del marcador del discurso *actually* en función del uso que hacen del mismo hablantes nativos de *Xhosa English*⁴ que aprenden inglés como segunda lengua. Para este trabajo, la autora ha recopilado un corpus de aproximadamente medio millón de palabras provenientes de la transcripción de diálogos espontáneos entre hablantes de la mencionada variedad de inglés. Mediante tablas de frecuencias, de Klerk ilustra la distribución de uso de este marcador, y realiza una clasificación en tres grandes áreas, con diferentes subdivisiones. El resultado de este análisis pone de manifiesto que existen dos contextos de uso de *actually* que divergen del empleo que le dan los hablantes nativos de inglés. La motivación que subyace en este estudio es el hecho de que los marcadores de discurso oral no se estudian en gramáticas, lo que resalta la ausencia de consideración de los aspectos orales en la enseñanza de una lengua como el inglés.

Esta misma autora (de Klerk 2005b) propone, tomando como fuente de datos el corpus de medio millón de palabras antes mencionado, un estudio sobre el marcador del discurso *well*. En él, realiza una clasificación de los usos contextualizados de este marcador, así como de sus significados procedimentales, estableciendo comparaciones con otros corpus de inglés. Con este trabajo, la investigadora somete a consideración los problemas que los aprendices de una L2 experimentan en el aprendizaje de los marcadores del discurso. En el caso de los hablantes implicados en este estudio, las dificultades en el uso de *well* se centran en la ausencia de un equivalente en su lengua materna, y en el contexto en el que aprenden inglés como L2. En este trabajo se señala la ausencia de materiales pedagógicos para el estudio de los marcadores del discurso, enfatizando la relación estrecha de uso con el lenguaje oral, desplazado de forma mayoritaria por materiales destinados a la práctica del lenguaje escrito.

Müller (2004) realiza un estudio sobre el marcador *well* analizando sus funciones en estudiantes alemanes que aprenden inglés, y compara sus resultados con el uso que hablantes americanos hacen de este marcador. Su análisis se basa en 70 conversaciones en las que parejas de alumnos comentan una película de cine mudo que han visto. Los diferentes usos de *well* se han categorizado en doce funciones, de las que sólo tres son coincidentes en frecuencia de uso con los hablantes americanos nativos. La autora compara también los datos obtenidos con respecto al marcador del discurso *so*, para concluir que, en el caso de *well*, la frecuencia de uso en los alumnos de inglés como lengua extranjera es mayor incluso que la frecuencia de uso de *so* por hablantes nativos. El análisis de estos resultados lleva a la autora a pensar que el uso de *well* se ve potenciado por el intento de evitar la utilización de *so* en alemán.

Por otra parte, O'Connell, Kowal y Dill (2004) realizan un estudio cuantitativo y cualitativo sobre marcadores de oralidad en ocho entrevistas realizadas en programas de noticias en televisión. En el análisis cuantitativo presentan el número y porcentaje de sílabas así como la frecuencia normalizada de los marcadores de oralidad encontrados, tanto en entrevistados como en entrevistadores. Los marcadores de oralidad considerados son *back-channeling*, *hesitations*, interrupciones, contracciones y elipsis, pronombres de primera persona, interjecciones y *question tags*. De los resultados se desprende que el recurso más empleado es *back-channeling*, seguido de las interjecciones. En el análisis cualitativo sí se aprecia el uso de estructuras mononucleares, tanto marcadores del discurso e interjecciones, como otras unidades con mayor carga de significado. Los autores concluyen que los entrevistados utilizan los marcadores de oralidad más frecuentemente, mientras que los entrevistadores adaptan el uso de los citados marcadores al contexto comunicativo concreto que se sustancia en las diferentes entrevistas.

Son más los artículos aparecidos en estos últimos dos años los que, de un modo central o marginal, tratan la oralidad y las unidades léxicas mononucleares. De forma general, los estudios encontrados trabajan con corpus orales que responden a diferentes diseños, y que en su mayoría ofrecen análisis

cualitativos de los datos manejados. En ningún caso hemos encontrado estudios que apliquen el *análisis multi-dimensional* como herramienta para estudiar la oralidad.

4. LA INVESTIGACIÓN DE LAS UNIDADES LÉXICAS COMPLEJAS

Tras el análisis de las estructuras mono-nucleares de las unidades léxicas, acometemos a continuación la exploración de la producción científica en otra área de gran importancia en el estudio global de la oralidad: las unidades léxicas complejas. En los trabajos de investigación examinados destaca fundamentalmente la investigación sobre *lexical bundles* y VBDUs (*Vocabulary-Based Discourse Units*).

La expresión *lexical bundles* podría definirse como las secuencias de tres o más palabras que aparecen con frecuencia en un registro (*Longman Grammar of Spoken and Written English (LGSWE)*, 1999). A pesar de los múltiples estudios llevados a cabo sobre estas unidades léxicas y de la relevancia que todos los autores otorgan a estas secuencias de palabras en los estudios sobre la oralidad, no hay mucho acuerdo en cuanto a la forma de definir las, ni en las metodologías para identificarlas. En lo que si están de acuerdo los investigadores es en que estas unidades léxicas no pueden considerarse ni como estructuras gramaticales completas ni como estructuras idiomáticas; sin embargo, funcionan como pilares básicos del discurso.

Biber, Conrad y Cortés (2004) investigan el uso de los *lexical bundles* en dos registros relacionados con la enseñanza universitaria: la instrucción en la clase y los libros de texto. Los autores comparan las *lexical bundles* usadas en la clase y en los libros de texto del Corpus T2K-SWAL con las encontradas en un trabajo de investigación anterior (Biber y Conrad 1999) sobre conversación y prosa académica. En primer lugar se describen los patrones estructurales y a continuación se presenta una taxonomía funcional incluyendo expresiones de opinión, organizadores del discurso y expresiones referenciales. El análisis indica

que los *lexical bundles* son un constructo lingüístico único. Los resultados del análisis sugieren que los *lexical bundles* deberían considerarse como un constructo lingüístico básico con importantes funciones para la construcción del discurso. Considerando que los *lexical bundles* se definen estrictamente sobre la base de la frecuencia sin consideraciones de criterios estructurales o funcionales, cabría esperar que pudieran etiquetarse como sucesiones de palabras arbitrarias sin estatus lingüístico. Los autores manifiestan que aunque no son el tipo de estructuras gramaticales reconocidas por la teoría lingüística tradicional, la mayoría de los *lexical bundles* tienen una estructura bien definida: generalmente están formadas por el comienzo de una frase más la primera palabra de una estructura subordinada (por ejemplo una frase preposicional).

El trabajo concluye que el uso de *lexical bundles* en el aula es especialmente sorprendente tanto en frecuencia como en función. Asimismo, se pone de manifiesto que se utilizan más grupos léxicos de opinión y de organización del discurso en el aula que en la prosa académica.

Cortés (2004) por su parte, trata de los *lexical bundles* en medio escrito. La autora investiga la función de estas unidades léxicas en prosa académica usando un corpus de publicaciones en historia y biología y lo compara con otro formado por la producción escrita de estudiantes universitarios de tres niveles de la Northern Arizona University de tres niveles diferentes de esas disciplinas. La investigadora identifica estructural y funcionalmente los *lexical bundles* de cuatro palabras en los dos corpus. El trabajo pone de manifiesto que existe una gran distancia entre el uso de estas construcciones por parte de los estudiantes universitarios de biología e historia y las usadas en los libros de estas disciplinas. Las conclusiones obtenidas tienen implicaciones pedagógicas en lo que se refiere a *academic writing*.

Cabría cuestionar la consideración de este trabajo dentro de un estudio sobre la oralidad. Siguiendo a Hudson-Ettle y Nilsson (2002) entendemos los textos orales no como contrapuestos a los textos escritos, sino como parte de un continuo en uno de cuyos extremos estarían las conversaciones informales y en el otro la prosa académica formal.

Por lo que respecta a las VBDUs (*Vocabulary-Based Discourse Units*) la producción no ha sido precisamente abundante en los dos últimos años, lo que

nos permite constatar el momento de cambio y de innovación investigadora que vive el campo del análisis del discurso, cada vez más influido por las técnicas computacionales para el tratamiento de los datos lingüísticos. De hecho hemos encontrado un solo trabajo en las revistas seleccionadas. Csomay (2004) lleva a cabo un estudio cuyo objetivo primordial es identificar la variación lingüística intra-textual en el discurso en un aula universitaria analizándola con técnicas basadas en el corpus. La metodología aplicada consiste en segmentar un gran número de sesiones de clase en unidades pequeñas de análisis. Posteriormente se identifican los patrones presentes de una serie de características lingüísticas aplicando técnicas analíticas multidimensionales. De acuerdo con esos patrones se identifican tres dimensiones de variación lingüística en el habla de la clase: (1) orientación directiva contextual frente a conceptual informativa; (2) *framing* personalizado frente a *framing* no personalizado; (3) dialogo interactivo frente a monólogo del profesor. Finalmente, y tomando en consideración estas dimensiones, se discuten las diferencias iniciales en los patrones del discurso en seis disciplinas y a tres niveles de instrucción. Este artículo ofrece nuevas evidencias desde múltiples perspectivas. En el estudio de registros basado en corpus lingüísticos supone una novedad ya que combina los estudios de corpus con estudios de patrones del discurso. El estudio llevado a cabo demuestra que las características lingüísticas del habla de la clase son parte de dimensiones más amplias de variación en este registro. Restaurando esas dimensiones al nivel de unidades del discurso en clases específicas el artículo describe la variación lingüística dentro de las sesiones de clase y como se relaciona esa variación con las disciplinas y niveles de instrucción.

5. LA INVESTIGACIÓN DE LAS UNIDADES SINTÁCTICAS Y PRAGMÁTICAS

Al contrario que los trabajos de Biber, Conrad y Cortés (2004) y de Csomay (2004), los artículos centrados en unidades sintácticas y pragmáticas no adoptan el entorno universitario formal como contexto de estudio. En este caso las

conversaciones son mayoritariamente espontáneas y libres. Destacan en esta sección las investigaciones de los siguientes fenómenos: elipsis, las partículas concesivas *of course* y *bien sûr* y *negative lexical have*.

Nariyama (2004) lleva a cabo un estudio en el que intenta descubrir los mecanismos que gobiernan la utilización de elipsis de sujetos sintácticos en inglés, una lengua que no pertenece a la tipología *pro-drop* (aquellas que presentan morfología de concordancia de sujeto y verbo). A pesar de ello, afirma el autor que las construcciones elípticas son bastante frecuentes en inglés, hecho que apenas ha recibido atención en el mundo investigador. Nariyama examina este fenómeno desde dos puntos de vista: sintáctico y pragmático; reconoce la importancia de los aspectos prosódicos aunque opta por no contemplarlos.

Para esta investigación se ha utilizado un corpus pequeño, de cuyo análisis resalta su carácter cualitativo. Los textos pertenecen al registro informal, tanto en la modalidad oral (conversaciones en familias pertenecientes al *Australian English Corpus* junto con guiones de series de televisión australianas) como escrita (postales, diarios y cartas informales). Si bien Nariyama reconoce en los guiones la falta de espontaneidad inherente a los diálogos libres, aduce que los ejemplos estudiados se consideraron naturales por hablantes nativos de inglés australiano.

Los resultados de Nariyama coinciden con otros estudios previos en relación al principio único de organización de la elipsis: la recuperación por parte del receptor de la identidad del sujeto sintáctico elidido gracias a información incluida en cualquier otra parte del contexto. Por tanto, se elimina la repetición del elemento sintáctico omitido, de tal manera que se respeta una de las “Máximas cooperativas de conversación” (Grice 1975), concretamente la de “Manera” (“sé breve”).

Nariyama proporciona una abundante y compleja exposición de los resultados, los más relevantes de los cuales se presentan ilustrados con extractos procedentes del mismo estudio.

Dada la importancia del contexto en la ejecución de la elipsis, Nariyama distingue dos tipos de aquél respecto a la elisión de sujetos: contexto situacional (información extra lingüística) y contexto meramente lingüístico. Dentro de esta distinción se encuadran los cuatro factores provocadores de la elipsis. El más

frecuente de ellos es la anáfora. El referente de ésta suele ser el sujeto de la oración inmediatamente precedente (*'Where's dad?' '(He's) birthday shopping, I bet'*). Como tal, depende del contexto lingüístico. El otro ingrediente de idéntica clase de contexto es *dummy subject* (*'I saw Rachel yesterday. I brought her flowers and we had a chat.'* *'(It) clearly did her a lot of good.'*) Los otros dos aspectos, pertenecientes a la esfera situacional, son expresiones convencionales o idiomáticas (*'Gotta go'*) y deixis.

Existen ciertas restricciones estructurales asociadas a estos cuatro elementos. El primero de ellos es la persona. En el corpus destaca el estatus preferente de la primera persona como sujeto elidido, hecho directamente relacionado con deixis. La primera persona se vincula además a oraciones declarativas, mientras que la segunda a interrogativas. Esta explicación se justifica desde un razonamiento epistémico: es más probable que el hablante esté familiarizado con los asuntos representados en primera persona, presumiblemente los suyos propios, que los relacionados y expresados mediante otra persona verbal. Igualmente, al poseer el emisor menos conocimiento del receptor, surge la elipsis de segunda persona en oraciones interrogativas, en número bastante inferior al de la primera.

Las expresiones convencionales constituyen el segundo tipo de restricción, en el que influyen persona, polaridad entre la primera y segunda persona en parejas de intercambios conversacionales, modalidad interrogativa y semántica verbal.

En el corpus, la mayor parte de estas expresiones incluyen la primera persona, y el resto, *it* (*'(It) doesn't matter'*). En algunas ocasiones, se puede omitir el sujeto en oraciones declarativas y no en las interrogativas (*'(I) shall /will do it.'* **(You) will do?'*). Además, la polaridad es frecuentemente fija en estos casos, de tal manera que mientras que es correcto decir *'(I) wouldn't mind a coffee'*, no lo es su respuesta de contenido idéntico **(You) wouldn't mind a coffee?'* Se descubre igualmente que ciertos verbos que requieren conocimiento epistémico son los que se asocian en mayor grado con las expresiones convencionales (*'(I) got in late'*; *'(I'd) like a coffee'*; *'(I'm) feeling good'*).

El tiempo y aspecto verbales con auxiliares representan la tercera restricción producida por los elementos causantes de elipsis. En el futuro es

aceptable omitir el sujeto pero no *will* (*'(I)'ll go to the football this weekend'*), contrariamente al caso de *going to*, en el que es correcto suprimir el sujeto y el auxiliar (*'(I'm) going to the football this weekend'*). Por otro lado, el aspecto imperfectivo debe eliminarse cuando se elide el sujeto (*'(You are/It is) looking good'*).

En cuanto al nivel sintáctico, en general los sujetos elididos se encuentran en posición inicial, estructura que coincide con la del imperativo. La escasa ambigüedad resultante se descifra con facilidad gracias a los significados de los verbos (**'Like a coffee!' vs 'Have a coffee'*); el tiempo verbal (**'Got a house!' vs 'Get a house'*) y el tipo de objeto (*'(I'll) tell you' vs 'Tell me'*). Asimismo, únicamente se pueden eliminar sujetos en oraciones principales (ej. *'(I'll) tell you if I've gotta go'*), siempre y cuando no aparezca inversión del tipo auxiliar-sujeto (**'Will (you) tell me if you've got to go?'*). Por otra parte, la elipsis sujeto-auxiliar parece ocurrir cuando la oración es idiomática, por ejemplo: *'(I'm) sorry'*; *'(I) dunno'*; *'(I) gotta go'*.

Desde un punto de vista pragmático, la elisión de sujetos implica connotaciones distintas a las oraciones sintácticamente completas. Ello es debido a la implicatura resultante de las primeras (significado derivado de la desviación tácita de las máximas conversacionales). Dichas implicaturas conllevan un matiz pragmático de desgana y evasión en el contenido. Por ejemplo, *'I don't know'* expresa genuinamente la idea de *'I thought about it, but I have no idea'*. Por el contrario, en *'(I) dunno'* se percibe *'I don't want to think about it'*, o incluso *'I don't care'*.

Respecto a los resultados concernientes a la elipsis en cartas informales, el corpus indica una abrumadora mayoría de sujetos de primera persona elididos en oraciones coordinadas (concretamente la segunda cláusula). Este hecho se corresponde con el contexto epistolar en relación a la descripción del escritor de eventos experimentados por él mismo.

La denominada *computer mediated communication* (CMC) contrastada en dos idiomas también es objeto de estudio en relación a marcadores sintácticos y pragmáticos. Lewis (2005) sostiene que, debido a su vertiginosa naturaleza cambiante, este tipo de interacción no puede identificarse simplemente cotejando sus similitudes y diferencias con el modo oral y escrito, sino que también se

deben tener en cuenta otros factores sociales y materiales.

Esta autora estudia la argumentación en diálogos electrónicos u *on-line* ingleses y franceses con participantes anónimos en situaciones de mínima información contextual. Para ello se basa en dos corpora comparables en inglés (360.000 palabras) y en francés (400.000 vocablos) procedentes de foros de análoga temática política servidos por cinco periódicos del Reino Unido y Francia.

Lewis localiza una estructura recurrente de los mensajes integrada por [reacción]+ postura + apoyo. Dentro de dicha estructura se hallan dos marcadores concesivos (*of course* y *bien sûr*) cuyo uso se asemeja más a aquel de la conversación que al del discurso político monológico. En una tabla de frecuencias de porcentajes, Lewis ofrece la distribución de los seis usos pragmáticos de *of course* y *bien sûr* identificados en ambos corpus: afirmación o negación enfática; introductor de un sub-tema; marcador de concesión; distanciamiento del hablante (ironía o desaprobación); *hedging* de relevancia; acompañante de un comentario de la idea previa. Sin embargo, los dos marcadores difieren en su distribución y hábitos de estructuración de la información. *Of course* aparece con bastante mayor frecuencia en el corpus, especialmente como un *hedge* de relevancia cuya omisión apenas altera el significado global de la oración. En cambio *bien sûr* se muestra en construcciones bastante más regulares y específicas. Este hecho realza su valor de estructuración de la información en contra del de oposición concesiva, que se representa por *certes* (“en verdad”, “claro es que”) y *il est vrai (que)/c’est vrai (que)* (“es cierto que”). Todas estas vicisitudes entrañan la imposibilidad traductológica de considerar *of course* y *bien sûr* equivalentes pragmáticos.

Corpus monolingües e internacionales constituyen también el material de análisis de unidades sintácticas a nivel oral. En esta línea Nelson (2004) investiga los patrones estructurales de *negative lexical have*: (1) *I don’t have (any) money*; (2) *I haven’t any money*; (3) *I haven’t got any money*; (4) *I have no money*; (5) *I’ve got no money*. Este autor identifica la recurrencia de cada modelo en el inglés de las variedades externas o *Outer Circles* (Nueva Zelanda, India, Hong Kong y Singapur) en comparación con las internas o *Inner Circles* (integradas por Estados Unidos y Gran Bretaña). El autor aclara que se limita a

inglés conversacional pues estima que la informalidad y espontaneidad de este registro proporcionarán preferencias más auténticas de los hablantes.

Mediante una comparación de frecuencias porcentuales, y basándose en los datos de Estados Unidos y Gran Bretaña suministrados en *LGSWE* (1999), Nelson corrobora para el inglés británico tanto su variabilidad como su tendencia del patrón (3). Este resultado contrasta con los dos usos preferidos en inglés americano (mayormente (1) seguido de (5)), hallados igualmente en *LGSWE* por Biber et al. Asimismo, el análisis de correlación de Pearson utilizado desvela que las variedades externas conforman un grupo independiente del inglés americano y británico.

En ninguno de los estudios de las unidades sintácticas y pragmáticas se ha utilizado un análisis de tipo multidimensional. Por otra parte, este apartado, al contrario que los anteriores, ha revelado la ausencia de objetivos o recomendaciones explícitas de índole didáctica. El carácter de los estudios es predominantemente descriptivo, con el propósito principal de realizar contrastes bilingües o de investigar ciertos aspectos lingüísticos en determinadas variedades del inglés.

6. CONCLUSIONES Y PROPUESTAS

La presente revisión de la literatura ha desvelado, como cabía esperar, una amplia disparidad en lo que se refiere a tipos de análisis, contextos de estudio, registros, número de lenguas estudiadas y variedades lingüísticas. La peculiaridad más importante respecto al primer aspecto reside en el escaso empleo de técnicas multidimensionales con excepción de los trabajos de estructuras complejas. En este sentido parece evidente que las investigaciones de Biber y su equipo siguen siendo pioneras en el campo, por lo que es factible que en el futuro mediato no se cuestione el modelo multidimensional original. Por otro lado, sorprende que la evolución en el tratamiento computacional de los datos lingüísticos no haya traído consigo una relectura de las relaciones entre características lingüísticas discretas y el concepto de dimensión lingüística.

Las estructuras complejas se han analizado casi exclusivamente en el

contexto universitario y formal, mientras que la minoría de los estudios presenta un registro que cubre desde la espontaneidad total hasta la formalidad de entrevistas fuera de un contexto universitario. Además, mientras que en la sección de estructuras mononucleares el tipo de análisis es primordialmente cualitativo, los artículos de las estructuras complejas son exclusivamente cuantitativos, quedando los trabajos de los elementos sintácticos y pragmáticos como casos intermedios. Es evidente que carecemos de un paradigma de análisis estandarizado que pueda dar respuestas a los intereses investigadores de la comunidad lingüística en materia de *oralidad*. Esto último no ha de verse como una crítica, sino como la constatación de la diversidad de enfoques que sustentan la praxis investigadora en el campo. Así, la metodología tradicional de análisis del discurso oral, centrada en gran medida en el ámbito conversacional, se ha ocupado del acto social, mientras que desde el ámbito de la lingüística descriptiva se ha hecho más énfasis en el lenguaje como producto. No es de extrañar, pues, que existan diferentes paradigmas de investigación que confluyen en el mismo hecho lingüístico. La metodología multidimensional se articula en este contexto como una interfaz extremadamente útil para la conjunción del concepto de función y el de caracterización morfo-sintáctica de los textos.

Asimismo, los trabajos de unidades sintácticas y pragmáticas han sido los únicos que han mostrado corpus bilingües y monolingües de distinta etiología. Por otra parte, los informantes de todos los artículos eran hablantes nativos, bien de francés (en uno de los corpus de Lewis (2005)) o de inglés, aparte del estatus de segunda lengua de este último idioma en los trabajos de de Klerk (2005a y 2005b). Hemos constatado un único estudio centrado en lengua inglesa cuyos informantes eran tanto americanos hablantes nativos de inglés como hablantes no nativos de inglés, en concreto alemanes (Müller2004).

Por último, cabe resaltar el hecho de que ningún trabajo sea de índole longitudinal, es decir, en el cual se comparen varios corpus de L2 de los mismos sujetos recopilados en puntos temporales distintos. Precisamente una comparación entre esta producción y la equivalente de nativos de la lengua en cuestión permitiría delinear semejanzas y divergencias y diseñar las opciones pedagógicas para la enseñanza de lenguas más apropiada a cada contexto educativo.

NOTAS

1. Trabajo realizado dentro del marco del proyecto de investigación 00580/PI/04, *Estudio longitudinal y comparativo de la producción oral (oralidad) de aprendices de inglés como lengua extranjera a partir del análisis de un corpus lingüístico oral recopilado y mediante procedimientos de análisis multidimensional*, financiado por la Fundación Séneca de la Comunidad Autónoma de la Región de Murcia.
2. Deliberadamente excluimos el debate sobre el término *texto* por ser aquél más propio de la meta-ciencia lingüística que de la pragmática.
3. *British National Corpus*.
4. *Xhosa English* es una sub-variedad del *Black South African English*. Son hablantes nativos de *Xhosa* que han aprendido Inglés como lengua adicional. Los hablantes de *Xhosa English* no han aprendido inglés con profesores genuinamente nativos, sino con profesores que pertenecen a segundas o terceras generaciones de hablantes en Sudáfrica.

BIBLIOGRAFÍA

- Abascal Vicente, M. D. 2004. *La teoría de la oralidad*. Málaga: Universidad de Málaga.
- Biber, D. 1988. *Variation across speech and writing*. Cambridge: Cambridge University Press.
- Biber, D. 2003. "Variation among university spoken and written registers: A new multidimensional analysis". *Corpus Analysis. Language Structure and Language Use*. Eds. P. Leistyna y C. Meyer. Amsterdam: Rodopi. 47-70.
- Biber, D., y S. Conrad. 1999. "Lexical bundles in conversation and academic prose". *Out of Corpora: Studies in honor of Stig Johansson*. Eds. H. Hasselgard y S. Oksefjell. Amsterdam: Rodopi. 181-189.
- Biber, D., S. Conrad, y V. Cortés. 2004. "If you look at ...: Lexical Bundles in University Teaching and Textbooks". *Applied Linguistics* 25, 3: 371-405.
- Biber, D., S. Johansson, G. Leech, S. Conrad, y E. Finegan. 1999. *Longman Grammar of Spoken and Written English*. Londres: Longman.
- Cortés, V. 2004. "Lexical bundles in published and student disciplinary writing: Examples from history and biology". *English for Specific Purposes* 23, 4: 397-423.
- Crystal, D. 2001. *Language and the Internet*. Cambridge: Cambridge University

Press.

- Csomay, E. 2004. "Linguistic variation within university classroom talk: a corpus-based perspective". *Linguistics and Education* 15, 3: 243-274.
- de Klerk, V. 2005a. "The use of actually in Spoken Xhosa English: a corpus study". *World Englishes* 34, 3: 275-288.
- de Klerk, V. 2005b. "Procedural meanings of *well* in a corpus of Xhosa English". *Journal of Pragmatics* 37: 1183-1205.
- Grice, H. P. 1975. "Logic and conversation". *Syntax and Semantics, 3: Speech Acts*. Eds. P. Cole y J. L. Morgan. New York: Academic Press. 43-58.
- Hudson-Ettle, D. y T. Nilsson. 2002. "Orality and noun phrase structure in registers of British and Kenyan English". *ICAME Journal* 26: 33-62.
- Lee, D. 2001. "Genres, Registers, Text Types, Domains, and Styles: Clarifying the Concepts and Navigating a Path through the BNC Jungle". *Language Learning & Technology* 5, 3: 37-72.
- Lewis, D. M. 2005. "Arguing in English and French asynchronous online discussion". *Journal of Pragmatics* 37: 1801-1818.
- McEnery, T. 2003. "Corpus Linguistics". *The Oxford Handbook of Computational Linguistics*. Ed. R. Mitkov. Oxford: Oxford University Press. 448-463.
- Müller, S. 2004. "'Well you know that type of person': functions of *well* in the speech of American and German students". *Journal of Pragmatics* 36: 1157-1182.
- Nariyama, S. 2004. "Subject ellipsis in English". *Journal of Pragmatics* 36: 237-264.
- Nelson, G. 2004. "Negation of lexical *have* in conversational English". *World Englishes* 23, 2: 299-308.
- O'Connell, D. C., S. Kowal y E. J. Dill III. 2004. "Dialogicality in TV news interviews". *Journal of Pragmatics* 36: 185-205.
- Teubert, W. 2004. "Language and Corpus Linguistics". *Lexicology and Corpus Linguistics*. Eds. M. A. K. Halliday, W. Teubert, C. Yallop y A. Cermáková. Londres: Continuum.
- Trappes-Lomax, H. 2004. "Discourse Analysis". *The Handbook of Applied Linguistics*. Eds. A. Davies y C. Elder. London: Blackwell. 133-164.