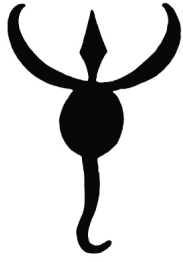
Revista Iberoamericana de Lingüística

RIL



nº 8 2013

R.I.L. 8

Fundador y Editor:
Ricardo de la Fuente Ballesteros (Universidad de Valladolid)

Co-editor:
Department of Modern Languages and Literatures (University of Texas at San Antonio)

Director:
Francisco Marcos Marín (University of Texas at San Antonio)

Comité asesor:
Humberto López Morales (Secretario de la Asociación de Academias de la Lengua Española)
José Antonio Pascual (Real Academia Española)
Liliana Sánchez (Rutgers University)
José Camacho (Rutgers University)
Alejandra Balestra (George Mason University)
Mariana Achugar (Carnegie Mellon University)
Brunello Natale di Cussatis (Università di Peruggia)
Luis Santos Río (Universidad de Salamanca)
Alfredo Torrejón (Auburn University)
Miguel Casas Torres (Universidad de Cádiz)
José Antonio Samper Padilla (Universidad de las Palmas de Gran Canaria)
Francisco Ocampo (University of Minnesota)
Francisco Javier Satorre Grau (Universidad de Valencia)
Antonio Salvador Plans (Universidad de Extremadura)

Comité de redacción:
María Jesús Leal (Hamline University)
Nelsy Echávez-Solano (College of Saint Benedict / S. John's University)
Antonio Gragera (Texas State University, San Marcos)
Barbara Gori (Università di Peruggia)
Antonio Carrasco (Universitas Castellae)

Edición, fotomecánica e impresión: Universitas Castellae, edificio 2
Plaza del Viejo Coso, 5
47003 Valladolid
España
Telf. 34 983 377 508 / 629 388 777
E-mail: cuc@universitascastellae.es
www.universitascastellae.es
www.reviblin.com
ISSN: 1887-407X
Fotomecánica e impresión: Universitas Castellae

ARTÍCULOS

ANÁLISIS DE LA PRODUCCIÓN Y DE ERRORES EN UN CORPUS ORAL DE ESPAÑOL COMO LENGUA EXTRANJERA*

Leonardo Campillos Llanos
Universidad Autónoma de Madrid

Resumen: este trabajo resume un estudio realizado sobre la adquisición de la expresión oral de cuarenta universitarios (N = 40) que estudiaban español como lengua extranjera en los niveles A2 y B1 (*MCER*). El método de obtención de datos es una entrevista semiespontánea, tomando cuatro participantes por cada lengua materna (principalmente, portugués, italiano, francés, inglés, neerlandés, alemán, polaco, chino y japonés). Igualmente, se entrevistó a cuatro hablantes nativos de edad similar (grupo de control) para establecer comparaciones acerca de la producción oral y los errores del habla. El estudio parte de la meto-dología de la lingüística de corpus, la didáctica del español, el análisis de errores asistido por ordenador y la investigación en adquisición de segundas lenguas. Uno de los resultados es la recogida de un corpus de interlengua oral de aprendices de español. Dicho banco de datos ha permitido analizar los aspectos formales del habla no nativa (recuento de palabras o categorías más frecuentes), para lo cual se anotó morfológicamente el corpus mediante el analizador informático GRAMPAL. Asimismo, se analizaron los errores de la producción oral en el nivel léxico, gramatical, de pronunciación y pragmático-discursivo. Para este fin se diseñó una tipología de errores y se anotó manualmente el corpus con etiquetas XML.

Palabras clave: adquisición, computacional, corpus, didáctica, entrevista, error, español, lengua materna, lengua extranjera, oral.

Abstract: This paper summarizes the results of a study on the acquisition of the oral expression by forty university learners (N = 40) who studied Spanish as a Foreign Language at A2 (N=20) and B1 (N=20) levels (*CEFR*). The data collection method was a semi-spontaneous interview, and learners were divided into 10 groups (with four participants in each group) according to their mother tongue (L1). Learners' L1 was principally Portuguese, Italian, French, English,

Dutch, German, Polish, Chinese, and Japanese. Likewise, four native speakers of similar age were interviewed with the same procedures (control group) in order to compare spoken language phenomena and errors between native and non-native speakers. The methodology used departs from the following approaches: Corpus linguistics, Didactics of Spanish as a foreign language, Computer-assisted error analysis, and Second language acquisition research. A result from the project is the collection of a Spanish learner language oral corpus for search and analysis purposes. This data bank has been used to examine formal aspects of non-native speech (e.g. word count or most frequent categories). To accomplish this task, the corpus was part-of-speech tagged by means of an automatic tagger for Spanish language (GRAMPAL). Furthermore, the errors were analysed at different linguistic levels: lexis-semantics, grammar, pronunciation, and pragmatics-discourse. For that goal, an error typology was designed and the errors were manually tagged with XML tags.

Keywords: acquisition, computational, corpus, didactics, error, foreign language, interview, mother tongue, oral, Spanish.

1. Introducción

El presente trabajo pretende contribuir al conocimiento del habla de diferentes grupos de aprendices extranjeros. Para ello se ha analizado su lengua en desarrollo (*interlengua*), atendiendo a las categorías empleadas y los errores cometidos durante una entrevista. Las grabaciones fueron transcritas y conforman un banco de datos (*corpus*) de la producción oral de 40 alumnos, cuyas lenguas maternas pertenecen a la familia románica (portugués, italiano y francés), germánica (inglés, neerlandés y alemán), eslava (polaco), sino-tibetana (chino) y a las lenguas de Japón (japonés).

Una de las motivaciones que fundamentan esta investigación es el parco volumen de estudios rigurosos sobre la adquisición del español (especialmente en niveles bajos), lo cual contrasta con el peso demográfico de esta lengua. Este trabajo tiene como objetivo concretar los aspectos

más espinosos de la adquisición del español en los grupos de nivel intermedio-bajo, a fin de mejorar la didáctica o los materiales de enseñanza. Pero también existe un interés tecnológico que se relaciona con la carencia de recursos lingüísticos especializados (como corpus de datos de estudiantes), y que este proyecto pretende paliar aportando un bancos de datos obtenidos de aprendices no nativos.

Las preguntas de nuestra investigación se pueden formular del siguiente modo:

- ¿En qué medida ciertos contenidos del aprendizaje del español causan dificultades a todos los alumnos en la expresión oral?
- ¿Por qué en el habla ciertos grupos de alumnos presentan más problemas que otros en determinados puntos? ¿Cuáles son esos contenidos y en qué grado?

En nuestro análisis abordamos todos los niveles lingüísticos (gramática, léxico, pronunciación y pragmática-discurso), y no ignoramos el hecho de que quizá hemos querido abarcar demasiado empleando únicamente una entrevista oral para obtener los datos –lo cual ha limitado el análisis completo o riguroso de ciertos aspectos–. Hemos optado por un enfoque global, en lugar de limitarnos a estudiar en profundidad un solo punto (p. ej., la adquisición de los clíticos). Es cierto que, actuando así y teniendo en cuenta el volumen de nuestro corpus, la generalización de los resultados puede ser limitada, pero preferimos tratar todas las áreas a fin de aproximarnos a la adquisición del español desde diferentes perspectivas, lo cual que creemos que enriquece el análisis.

Nuestro trabajo se enmarca en la encrucijada de los estudios de adquisición en segundas lenguas, la didáctica del español para extranjeros, el análisis de errores asistido por ordenador (Dagneaux *et al.*, 1998; Barlow, 2005) y la lingüística de corpus. Asimismo, conviene realizar algunas precisiones terminológicas: aunque se distinguen ambos conceptos, hemos usado los términos *español lengua extranjera* y *español como segunda lengua* (en adelante, L2 o LE, respectivamente) para referirnos indistintamente al proceso de adquisición/aprendizaje en el país donde se habla la lengua que se

adquiere. Finalmente, cabe subrayar el carácter aplicado de esta investigación, que ha intentado estrechar la distancia entre la investigación básica y la práctica docente. Uno de los resultados del proyecto ha sido el desarrollo de una interfaz de consulta para el corpus oral, que se piensa destinar a la formación de profesionales de la enseñanza[1].

Este artículo resume los puntos destacados del estudio que se explica con detalle en Campillos Llanos (2012), y se organiza del siguiente modo. La sección 2 aborda los planteamientos teóricos, y seguidamente se presentan la investigación (sección 3) y la metodología (sección 4). La sección 5 explica el análisis formal de la producción (recuento de palabras y unidades léxicas, y uso de categorías) y la sección 6 detalla el análisis de errores (léxicos, gramaticales, de pronunciación y pragmático-discursivos). Por último, se exponen las conclusiones y las posibilidades de trabajo futuro.

2. Planteamientos teóricos

En esta sección, repasamos los paradigmas desde los que se ha estudiado la interlengua, centrándonos después en el análisis de errores asistido por ordenador. Seguidamente apuntamos nociones de la metodología de la lingüística de corpus; y en último lugar, exponemos el marco de investigación con corpus para el análisis de errores.

2.1. Paradigmas de investigación y análisis de la interlengua de los estudiantes

El concepto de la *interlengua* (según el término forjado por L. Selinker, 1972) hace referencia al sistema lingüístico de los aprendices en su proceso de adquisición, y viene siendo referido por otras denominaciones como *dialecto idiosincrásico* (Corder, 1971) o *sistema aproximado* (Nemser, 1971). Las investigaciones sobre la adquisición de L2s o LEs han abordado la interlengua desde diferentes paradigmas, de entre los cuales pueden destacarse los siguientes (véanse más referencias en Larsen-Freeman y Long, 1991[1994]; Mackey y Gass, 2005; o Ellis y Barkhuizen, 2005):

1. Análisis contrastivo.
2. Análisis de la actuación y análisis de la interlengua o de la frecuencia.
3. Análisis del discurso.
4. Análisis de errores.

El Análisis Contrastivo fue desarrollado entre los años 1940 y 1960 por Charles Fries y Robert Lado, y parte de las teorías conductistas de la adquisición de lenguas, que consideran que el lenguaje se adquiere mediante formación de hábitos. La piedra angular de este paradigma es la Hipótesis del Análisis Contrastivo, que defiende que los elementos similares entre la lengua materna (L1, en adelante) y la lengua que se adquiere serán más sencillos de aprender (por transferencia positiva de la L1 a la L2); y a la inversa, los elementos que divergen serán más difíciles de adquirir. Aunque este supuesto ha sido refutado por los datos de las producciones de los alumnos –no siempre los elementos de las lenguas más cercanas tipológicamente son más fáciles de adquirir–, la investigación resultante ha sido fecunda en estudios entre pares de lenguas.

El análisis de la actuación (*Performance Analysis*) estudia las producciones lingüísticas de los aprendices, tanto las correctas como las incorrectas (Ellis y Barkhuizen, 2005: 73). En los años 70 se empezó a aplicar a la lingüística adquisicional, con estudios longitudinales y transversales, y dio lugar, por ejemplo, a estudios sobre el orden de adquisición de los clíticos. Análogamente, el análisis de la interlengua (*Interlanguage Analysis*) o de la frecuencia (*Frequency Analysis*) estudia un contenido gramatical o comunicativo para dar cuenta de su variabilidad en la lengua del aprendiz, desde perspectivas generativas o cognitivas. Una de las críticas que han recibido es la dificultad para describir la variabilidad de la interlengua en diferentes contextos cuando se centran solo en una modalidad de lengua (oral o escrita) (Dussias, 2003: 237).

Por su parte, el *análisis del discurso* considera el lenguaje como *uso*, poniendo en relación los enunciados con el contexto o situación lingüística; en este marco destacan el *análisis*

de la conversación y el *análisis interaccional.* Las críticas a estos enfoques provienen de estudiosos con un perfil más cuantitativo, que los acusan de falta de rigor por su carácter cualitativo. En realidad, la combinación de ambos tipos de métodos enriquece el análisis, como indica Dörnyei (2007).

Finalmente, el Análisis de Errores surge en la década de 1970 como evolución del Análisis contrastivo tras comprobar que sus premisas no tenían el carácter predictivo otorgado. Corder (1971[1991]) establece diferentes pasos en esta metodología:

1. Identificación del error.
2. Descripción y clasificación del error: se tiende a señalar el nivel lingüístico en que se produce y el mecanismo de cambio que lo causa (p. ej. el orden de palabras).
3. Explicación del error: se diagnostica si se debe a interferencia (interlingüístico), si es de desarrollo (intralingüístico) o se explica por otros factores.
4. Valoración de la importancia del error para la comunicación.

Uno de los reproches hacia este enfoque es que no permite describir completamente la competencia lingüística de los aprendices, puesto que se centra solo en lo que se equivocan. Los hablantes pueden *evitar* la producción de un término o estructura y puede resultar opaco esclarecer las dificultades que existen. En cambio, a diferencia de otros paradigmas, el análisis de errores intenta explicar la causa del error, aunque el diagnóstico no resulte fácil de establecer. Por ello, el análisis de errores no ha sido relegado, sino incorporado al *análisis de la producción* (Larsen-Freeman y Long, 1991).

2.2. La lingüística de corpus

Un *corpus* se puede definir como un conjunto de textos auténticos, procesados informáticamente y escogidos según un criterio para un estudio. Rara vez se trata de documentos al azar, sino que reúnen ciertas características con un fin determinado. Así, la etapa de diseño es cru-

cial, pues define cuestiones como las siguientes (Marcos Marín, 1994; McEnery, Xiao y Tono, 2006: 4-ss.; Granger, 2008: 261-263):

- Representatividad: un tamaño de la muestra equilibrado y proporcional.
- Procesamiento informático: metadatos y anotación o etiquetado.
- Disponibilidad: el acceso a los datos (ej. condicionado por tipo de contenidos).

La anotación lingüística del corpus se puede mantener ajena en otro documento independiente (*stand-alone annotation*), o disponerla en cada archivo correspondiente (*embedded annotation* o 'incrustada'), como procedemos en nuestro trabajo.

Otra discusión acerca de la lingüística de corpus es de si se trata de una metodología o de una rama de la Lingüística. Nosotros seguimos el criterio de McEnery *et al.* (2006: 7), y la consideramos una metodología de estudio, no una disciplina como la gramática.

2.3. Tendiendo puentes entre la lingüística de corpus, el análisis de errores y la adquisición del español como L2/LE

Los estudios con corpus de estudiantes (*Learner Corpus Research*) siguen dos métodos:

> Análisis contrastivo de la interlengua (*Contrastive Interlanguage Analysis*): realiza bien un análisis contrastivo entre producciones lingüísticas de nativos y de no nativos –para observar, por ejemplo, el uso excesivo o limitado de una estructura por los aprendices–, o bien entre las producciones de no nativos de diferentes L1 –para estudiar el grado de interferencia lingüística– (Gilquin, 2008). Este método no precisa de etiquetado de errores, aunque en algunos proyectos se realiza la anotación morfológica para comparar el uso de categorías entre nativos y no nativos.
>
> Análisis de errores asistido por ordenador (*Computer-aided error analysis*): se fundamenta en la meto-

dología del análisis de errores, y a las fases establecidas por Corder (1971) se le añade una etapa más, que sería el etiquetado de errores a partir de su descripción y explicación. Mediante las etiquetas para cada incorrección y técnicas de la lingüística de corpus se pueden recuperar las incorrecciones según criterios de búsqueda y extraer frecuencia de errores (Dagneaux *et al.*, 1998).

Cada método tiene sus partidarios (Myles, 2005: 378): el análisis de errores asistido por ordenador interesa a los profesores de idiomas, pero no a los investigadores en adquisición de lenguas, quienes prefieren no estudiar los errores. En nuestra opinión, resulta muy positiva la combinación de metodologías, como intentamos aplicar en este trabajo.

3. Presentación de la investigación

3.1 Diseño del corpus

El objetivo fue construir un corpus de análisis multilingüe con datos de aprendices de nivel plataforma (A2) y umbral (B1) (*MCER*). A fin de lograr resultados comparables, buscamos la misma cifra de informantes por nivel (N = 20). Igual procedimos respecto al número de sujetos de cada L1 (4 alumnos por grupo), logrando 9 grupos homogéneos (más uno heterogéneo). Con todo, no se alcanzó en todos un equilibrio en cuanto al nivel o al sexo del hablante debido a la disponibilidad de los alumnos. En total, se entrevistó a cuarenta (N = 40) estudiantes extranjeros de español, algunos aprendices internacionales del programa ERASMUS, y otros, alumnos de intercambio por otros convenios. La mayoría estudiaban español en el Servicio de Idiomas de la Universidad Autónoma de Madrid, y otros, en los cursos de español de la Universidad Complutense de Madrid. La tabla 1 resume los datos de las grabaciones (incluyendo los nativos): sexo (H, hombre, o M, mujer), lengua materna, nivel (*MCER*), duración de la grabación (respecto a cada informante y al grupo de L1), y turnos de habla por entrevista.

	Archivo	Sexo	L1	Nivel (MCER)	Duración (mm : ss)	Duración grupo L1	Nº de turnos
Lenguas romances	PORMA2	H	Portugués	A2	25:10	1:26:52	524
	PORWA2_1	M		A2	20:09		328
	PORWA2_2	M	Portugués (brasileño)	A2	19:51		462
	PORWB1	M		B1	21:42		496
	ITAMA2	H	Italiano	A2	20:45	1:13:25	540
	ITAWA2	M		A2	13:09		304
	ITAMB1	H		B1	23:16		436
	ITAWB1	M		B1	16:15		280
	FREMA2	H	Francés	A2	24:08	1:23:17	584
	FREWA2	M		A2	20:31		250
	FREMB1	H		B1	21:56		566
	FREWB1	M		B1	16:46		522
Lenguas germánicas	ENGWA2	M	Inglés	A2	15:04	1:20:39	348
	ENGMB1	H		B1	18:44		436
	ENGWB1_1	M		B1	18:02		347
	ENGWB1_2	M		B1	28:49		733
	DUTMA2	H	Neerlandés	A2	18:19	1:16:46	454
	DUTWA2_1	M		A2	17:33		180
	DUTWA2_2	M		A2	23:05		582
	DUTWB1	M		B1	17:49		370
	GERMA2	H	Alemán	A2	18:23	1:13:24	306
	GERWA2	M		A2	19:45		526
	GERWB1_1	M		B1	15:35		284
	GERWB1_2	M		B1	19:41		336
Lenguas eslavas	POLMA2_1	H	Polaco	A2	22:20	1:32:25	510
	POLMA2_2	H		A2	30:28		656
	POLMB1	H		B1	26:46		443
	POLWB1	M		B1	12:51		268
Lenguas sino-tibetanas	CHIWA2_1	M	Chino	A2	18:48	1:17:27	478
	CHIWA2_2	M		A2	18:45		450
	CHIMB1	H		B1	18:56		425
	CHIWB1	M		B1	20:58		449
Lengua japonesa	JAPWA2	M	Japonés	A2	28:52	1:32:41	552
	JAPWB1_1	M		B1	16:28		466
	JAPWB1_2	M		B1	20:59		498
	JAPWB1_3	M		B1	26:22		679
Otras lenguas	FINWA2	M	Finés	A2	20:27	1:19:05	544
	HUNWA2	M	Húngaro	A2	21:28		164
	KORWB1	M	Coreano	B1	21:14		462
	TURWB1	M	Turco	B1	15:56		288
Grupo control	SPAM_1	H	Español	-	18:57	1:22:29	401
	SPAM_2	H		-	26:47		626
	SPAW_1	M		-	16:49		307
	SPAW_2	M		-	19:56		333

Tabla 1. Datos de los participantes en el corpus

3.2. Identificación del error

En la consideración de error, hemos tenido en cuenta las características de la lengua oral. Las estructuras reformuladas las consideramos *fallos* (Corder, 1971) y no las incluimos en el recuento. Por ejemplo, en este enunciado no marcamos el uso de *ser*:

*JHA: la pizzería *es* cerrada [/] *está* &ce [/] cerrada /// (FREWA2)

Una cuestión más compleja es la propia consideración de errores que también son cometidos por hablantes nativos de español en la conversación informal. Por ejemplo, también es frecuente reconocer las siguientes desviaciones en el habla espontánea (debido a la planificación deficiente o improvisada del discurso oral):

- La omisión de ciertas preposiciones ante subordinadas: ej. *me acuerdo* (*de*) *que*
- El uso innecesario de ciertos conectores, especialmente *y, entonces*, y *por eso*.

Al analizar las transcripciones del grupo de control, se comprueba que estos fenómenos también aparecen en el habla nativa, a la que se están expuestos los no nativos. Sucede lo mismo con los siguientes fenómenos (se indica el archivo del grupo de nativos):

- Concordancia de persona errónea (especialmente con el verbo impersonal *haber* o el sustantivo colectivo singular *gente*): *la gente *hablasen* (SPAM_2).
- Concordancia de género o número erróneo (especialmente cuando el modificador no se encuentra cerca del sustantivo): *la plaza que es un poco *nuevo* (SPAM_1).

No obstante, decidimos corregir estas discordancias, estimando que son desviaciones gramaticales. Pero en la consideración de error no solo se ha de tener en cuenta la variación diafásica (registro informal y formal), sino también la variación dialectal, a fin de evitar la hipercorrección en el proceso de etiquetado. Por ejemplo, no son errores ni la pronunciación de *c + e/i* con sonido [s] (*seseo*) ni la de /x/ con sonido [h], dado que estos fenómenos son comunes en las variedades de Sudamérica o el sur de España. Como consecuencia, estos fenómenos dialectales no se marcan como error.

Otras dificultades surgen ante la interpretación de nuestros datos orales, para los cuales no hay más evidencia que el habla grabada –a diferencia del análisis de errores de la producción escrita, donde la forma de las palabras está presente–. De hecho, no solo se registran ciertas

palabras incomprensibles en la grabación –lo cual vuelve borroso el significado del mensaje–, sino que también ciertos errores son ambiguos. Por ejemplo, existen ambigüedades respecto al nivel lingüístico del error, como en el este contexto:

> *AMA: es una → *lengua {%pho: ['lɪŋgwa]}* muy [/] muy <linda> (PORWB1)

A partir del sonido ['lɪŋgwa], podemos tener dos interpretaciones: 1. que es un error formal de nivel léxico (**lingua* en vez de *lengua*); o 2., de pronunciación (*e,* /e/ → [ɪ]). En este caso, la L1 del aprendiz (portugués) no ayuda a despejar la ambigüedad, ya que en dicho idioma la vocal *e* se pronuncia con un sonido más cerrado ([i]), y la palabra que corresponde a *lengua* también se escribe *língua.* En tanto en cuanto no poseemos mayor evidencia de que el alumno conoce la forma de la palabra (a partir de datos escritos), este tipo de ambigüedad queda sin ser resuelta. Con todo, no cabe duda de que la L1 interfiere en este caso específico.

Otros problemas se relacionan con las deficiencias de la metodología del análisis de errores. A veces es imposible diagnosticar la causa de algunas desviaciones porque no podemos conocer el mensaje que pretendía expresar el aprendiz. Veamos el siguiente enunciado en el que una estudiante explica cómo prepara el sushi:

> *MSU: primero / *tiene que comprar* / arroz japonés /// (JAPWB1_2)

En una interpretación, el error estribaría en la concordancia verbal –la alumna usa la tercera persona (*tiene que comprar*) en lugar de la segunda singular, que puede expresar impersonalidad (*tienes que comprar*). En otra interpretación, el error se relacionaría con la omisión de pronombre *se*, si se pretendía usar la pasiva refleja (*se tiene que comprar arroz*). Ambas correcciones son posibles, y el análisis de errores resulta insatisfactorio. No obstante, este tipo de errores idiosincrásicos no parece revelar ninguna estrategia de aprendizaje o de desarrollo, ni ningún fenómeno de

interferencia, así que deberíamos otorgarle una importancia relativa en el conjunto del análisis.

3.3. Tipología de errores

Se ha diseñado una tipología de errores a partir de estudios previos para el inglés (James, 1998; Nicholls, 2003), el francés (Granger, 2003) o el español (Fernández López, 1990; Vázquez, 1999). Una taxonomía ideal de errores debería ser, como explica Granger (2003), consistente, informativa, flexible y reutilizable. Con todo, la mayoría de las clasificaciones encuentran los siguientes problemas (Heift y Schulze, 2007: 102):

- Cobertura: algunos errores pueden quedar sin ser clasificados cuando la taxonomía se aplica a otro corpus, dada la imposibilidad de anticipar nuevos casos.
- Homogeneidad: es difícil una descripción completa y significativa de los errores.
- Características de la clasificación: muchas categorías son dependientes de cada lengua, o no pueden ser distinguidas fácilmente (por ejemplo, en español el verbo *hacer* + expresión de tiempo tiene la misma función que un adverbio de tiempo).

Los siguientes son los criterios para la clasificación de errores en nuestra taxonomía:

- Categoría, elemento fonológico o categoría sintáctica.
- Mecanismo de cambio: el proceso que cause el error:
  - Omisión: ej., artículo: **camarero está*
  - Uso innecesario: ej., **un mi amigo*
  - Mezcla: ej., *novelas de históricos* ⇒*novelas históricas* o *novelas de historia*
  - Orden: ej., colocación de las palabras en el enunciado.
  - Deformación: ej., **higenias* ⇒ *higiene*
  - Falsa selección: ej., **soy satisfecho* ⇒ *estoy satisfecho*
- Nivel lingüístico:
  - Pronunciación: ej., acento léxico: **abanico [a'ßaniko]* ⇒ abaNIco.

    - Léxico-semántico:
        - Errores formales: ej., acuñaciones (**omeleta* ⇒ *tortilla,* de 'omelet').
        - Errores semánticos: ej., confusión entre *ir/venir*.
    - Gramática: ej., conjugación (**tiengo* ⇒ *tengo*).
    - Pragmática-Discurso: coherencia, cohesión y registro.
- Tipo: grupos específicos de error frecuente: ej., *ser/estar* (**soy satisfecho* ⇒ *estoy*).
- Etiología (causa del error):
    - Intralingüística (errores de desarrollo): **escribido* ⇒ *escrito*
    - Interlingüística (interferencia)[2]: ej., calcos (*frutas de mar* ⇒ *marisco*).
    - Desconocida: ej., errores inducidos por la enseñanza o los materiales.

Estas categorías descriptivas pueden ser problemáticas. El grado de precisión al distinguir errores similares (la *granularidad*) puede resultar excesivo y difícil de comprender por los no especialistas. Igualmente, esta propuesta está determinada por nuestro banco de datos y puede necesitar una adaptación para otros análisis. Este problema de *robustez* se debe en parte a la carencia de estándares entre las taxonomías de errores. Por último, respecto a los errores que no se pueden asignar a una categoría, dudamos de si crear una nueva clase para ellos; nuestra decisión evitó el riesgo de generar en exceso subdivisiones para pocos casos.

Errores ambiguos y criterios de clasificación de error

Ante los errores ambiguos, seguimos el principio de mínima corrección (*Minimum Correction Principle*; Köhlmyr, 2003; tomado de Díaz-Negrillo y Valera, 2009). Con un criterio semejante, se procedió en el Cambridge Learner Corpus, en el cual, ante varias interpretaciones, se decidió marcar el error que menos cambio suponga en la estructura original o el que menos severo resulte (Nicholls, 2003). En la clasificación de errores es preciso

observar el contexto discursivo anterior, sobre todo en los errores *concomitantes* (James, 1998: 117), causados por otra desviación en el entorno próximo.

Los criterios más importantes para el marcado y recuento de errores son los siguientes:

- Si concurren varios errores en una palabra o sintagma, contamos todos, incluso si el uso de un elemento no es necesario en la forma correcta; por ejemplo:

  *SOR: una → [/] *un otro* comida (ENGWB1_2)

  Contamos dos errores: la selección de género del sustantivo *comida* (se concuerda en masculino con *un* y *otro*), y el uso innecesario del artículo con el indefinido.
- La tendencia ha sido minimizar el recuento de errores. Por ejemplo, en este caso:

  *DAK: y en otro → / ¿ *olla {%err: ollo}* ? (POLMA2_2)

  contamos un error por asignación de género (en vez de dos, deformación y género).
- En el caso de las palabras erróneas que se repiten en el mismo enunciado, solamente etiquetamos y contabilizamos la incorrección en la última forma; p. ej.:

  *REM: que *era* → [/] que *era* muy bien ('estuvo') (FREMB1)

  En este contexto, corregimos solamente el error en la última forma de *era.*

- Marcamos solo un error en incorrecciones que afectan a varias palabras pero que se deben a la lógica de una misma regla errónea; ej., la asignación de plural a *gente*:

  *AIS: *las* gente siempre *tienen* / tiempo ('la gente … tiene') (ENGWA2)

4. Metodología

El proceso ha consistido en las siguientes fases (figura 1):

1. Compilación del corpus (recogida de entrevistas).
2. Transcripción de las grabaciones.
3. Sincronización del sonido y el texto.
4. Conversión al formato XML.
5. Procesamiento y segmentación de las grabaciones.
6. Etiquetado de errores.
7. Análisis de errores.
8. Anotación morfológica.

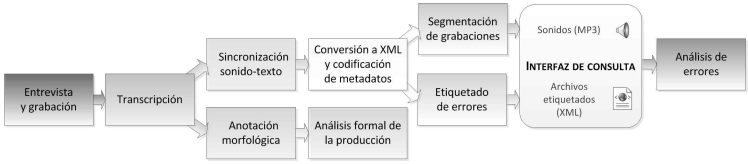


Figura 1. Fases y metodología del proyecto

En los puntos siguientes abordamos dichas fases, excepto el Análisis de errores (§4).

4.1. Compilación del corpus

Los participantes (tanto nativos como no nativos) fueron grabados mientras realizaban una entrevista semiestructurada en la que todos tenían que realizar las mismas tareas. Las técnicas de obtención de datos son similares a las de los exámenes de idiomas:

- Descripción de dos fotografías (en este trabajo, dos tipos de comida).
- Una tarea narrativa a partir de dos historias representadas en viñetas, a fin de que los alumnos usen tiempos de pasado y mecanismos de cohesión discursiva, y para plantearles una pregunta sobre dos actos de habla (una sugerencia y una petición).

En la parte final de la entrevista, los informantes tenían que expresar su opinión acerca de diferentes tópicos (p. ej.,

sobre los cambios en los hábitos de alimentación actual), con el propósito de obtener un habla más espontánea.

4.2. Transcripción de las grabaciones. Marcas de transcripción

Las entrevistas se transcribieron manualmente utilizando el programa Transana© (desarrollado en la universidad de Wisconsin) y fueron revisadas por otro lingüista, a fin de asegurar la consistencia de los contenidos. Las marcas de transcripción son una adaptación del formato CHAT (como se utilizó en el proyecto C-Oral-Rom; Cresti y Moneglia, 2005) y de las convenciones usadas en SPLLOC (Mitchell *et al.*, 2008). En la cabecera de cada archivo se codifican los metadatos de la entrevista, a saber:

- Título y nombre del archivo.
- Datos de los participantes: nombre, sexo, edad, nivel educativo, profesión y origen.
- Fecha, lugar de grabación, situación y tema del discurso.
- Datos sobre la adquisición del español: L1, conocimiento de otras lenguas, nivel, tiempo y contexto de estudio, o tiempo de estancia en país hispanohablante.
- Otras informaciones sobre la entrevista: duración, transcriptor y revisor, etc.

En la transcripción se marcan los siguientes fenómenos (ejemplos de ENGWA2):

- Fenómenos del habla:
  - Pausas: ej., y solo / gramática
  - Actos paralingüísticos/extralingüísticos (risa, tos...): ej. ok hhh {%act: laugh}
  - Solapamientos: ej., *ENT: y <francés> ///
    *AIS: [<] <sí> ///
  - Disfluencias:
    - Repeticiones y reformulaciones: ej. es [/] es muy diferente
    - Apoyos vocálicos o titubeos: ej., &ah &mm / me gusta / &ah leer
    - Palabras fragmentadas: &ma [/] mucho // sí
- Fenómenos del proceso de adquisición de lenguas extranjeras:

- Errores de pronunciación: ej. son más / como / tranquilo {%pho: [tran'kwilo]}
- Errores de deformación: ej., &meji [/] mejillones {%err: melijones} ///
- Acuñaciones o neologismos: ej., el cabarero@n está / enfadada

Los datos personales del aprendiz (como el nombre o el apellido) fueron anonimizados para proteger la identidad de los participantes.

4.3. Sincronización del sonido y el texto. Procesamiento y conversión a XML

La sincronización del texto transcrito con cada segmento de tiempo es una tarea manual realizada con el programa Transana©. El proceso consiste en ir escuchando la grabación e insertar cada marca temporal de final de fragmento mediante una combinación de teclas. Una vez añadidos los códigos de tiempo en las transcripciones, los archivos se procesan y convierten al formato XML. Los archivos procesados incluyen los códigos de tiempo correspondientes a cada fragmento, y dichas marcas se usan para segmentar las grabaciones mediante un programa que automatiza el proceso.

4.4. Etiquetado de errores

El etiquetado se realiza mediante el lenguaje de marcado extensible (eXtensible Markup Language, XML), un formato flexible para el diseño personalizado de etiquetas que ya se ha empleado en proyectos de lingüística, digitalización de textos o corpus de errores (respecto a este tema, véase Izumi *et al.*, 2005; Lüdeling *et al.*, 2005; o Granger *et al.*, 2007). Los elementos de cada fichero XML son los siguientes:

- Metadatos de la entrevista: corresponden a la cabecera de cada transcripción.
- Transcripción: incluye los códigos de tiempo de cada enunciado y turno.
- Etiquetas de errores.

El objetivo de este diseño de datos es permitir la recuperación automática de los fragmentos de texto y sonido donde aparecen errores específicos (según se han definido en las etiquetas de error). En las secciones siguientes explicaremos el conjunto de etiquetas usado para codificar las transcripciones y para marcar y describir los errores.

4.4.1. Etiquetas XML para los contenidos y los metadatos de la transcripción

El diseño de las etiquetas se basa en las usadas en C-Oral-Rom (Cresti y Moneglia, 2005), a las que se añadieron otras específicas que recogen datos sobre la adquisición del español. Los metadatos de la entrevista (que se registran en la cabecera) codifican las variables consideradas en el análisis de errores, entre las que se suele distinguir las *de la tarea* y las *del estudiante* (Barlow, 2005: 337; Granger, 2008: 263-ss.):

- Variables de la tarea (o del contexto): p. ej. condiciones de realización de la prueba.
- Variables del estudiante: p. ej., origen, nivel de competencia (*MCER*), L1, etc.

Algunos de los descriptores (como el nivel educativo o de competencia en español) deben ser tomados con cautela a causa de variabilidad que puede registrar cada categoría (p. ej., debido a la dificultades en alcanzar clasificaciones similares de nivel entre diferentes instituciones o países; Barlow, 2005: 337).

Por lo que respecta a la transcripción, el formato de etiquetas correspondiente a cada unidad de sincronización entre texto y sonido se ha diseñado para codificar interrupciones del otro hablante, o solapamientos de habla, que son muy frecuentes en la conversación espontánea.

4.4.2. Etiquetas XML para los errores

Las etiquetas para el marcado de errores se han diseñado teniendo en consideración los etiquetarios de proyectos anteriores (Granger, 2003; Izumi *et al.*, 2004; Nicholls, 2003; Lüde-

ling *et al.*, 2005; véase un resumen en Díaz-Negrillo y Fernández, 2006). En los archivos XML, se identifica el error en el turno de habla del aprendiz, se describe y se explica mediante los criterios expuestos en la tipología de errores. Un conjunto de etiquetas permite marcar también los errores ambiguos, incluyendo dos interpretaciones diferentes. El diseño de las etiquetas permite marcar errores dentro de otros, en una estructura recursiva para identificar errores que coaparecen en una misma palabra o sintagma; por ejemplo, un error de pronunciación dentro de otra estructura incorrecta.

4.5. Anotación morfológica

Para el estudio cuantitativo de la producción y uso de categorías por los aprendices es necesario contar con corpus anotados morfológicamente. Creemos que la anotación morfosintáctica aporta más ventajas que inconvenientes para el análisis (véase discusión en McEnery, Xiao y Tono, 2006: 29-ss.). El proceso de anotación es semiautomático: en primer lugar, se usa un analizador morfológico informático (GRAMPAL; Moreno Sandoval y Guirao, 2006) que asigna categorías a cada palabra. En segundo lugar, se revisa la anotación mediante una interfaz para la revisión asistida por ordenador (figura 2)[3].

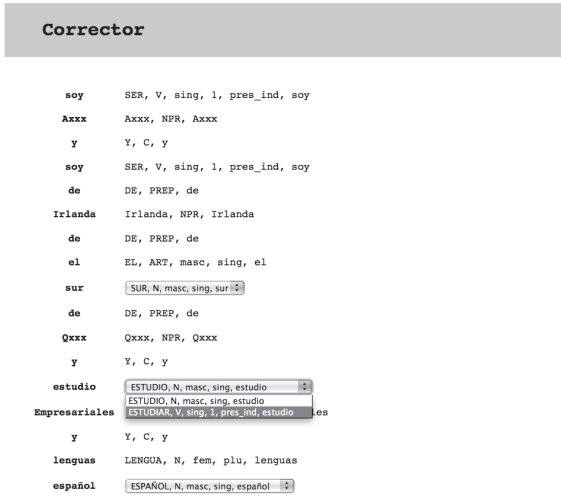


Figura 2. Interfaz de GRAMPAL para la revisión de la anotación morfológica

Una muestra del resultado final de la anotación en XML se presenta en la figura 3.

```
<w cat="V" lem="SER" num="sing" per="1" tie="pres_ind" for="soy" id="1-1-3">soy</w>
<w cat="NPR" lem="Axxx" for="Axxx" id="1-1-4">Axxx</w>
<w cat="C" lem="Y" for="y" id="1-1-5">y</w>
<w cat="V" lem="SER" num="sing" per="1" tie="pres_ind" for="soy" id="1-1-6">soy</w>
<w cat="PREP" lem="DE" for="de" id="1-1-7">de</w>
<w cat="NPR" lem="Irlanda" for="Irlanda" id="1-1-8">Irlanda</w>
<a for="del"> del
<w cat="PREP" lem="DE" for="de" id="1-1-9"/>
<w cat="ART" lem="EL" gen="masc" num="sing" for="el" id="1-1-10"/>
</a>
<w cat="N" lem="SUR" gen="masc" num="sing" for="sur" id="1-1-11">sur</w>
```

Figura 3. Muestra de un fragmento del archivo ENGWA2 anotado con GRAMPAL

5. Análisis cualitativo y formal

5.1. Recuento de palabras y unidades léxicas por grupos de L1

Hemos realizado un recuento en bruto del número de palabras, y de una manera menos mecánica, de unidades léxicas, según los criterios que define Fernández López (1990):

- Palabras simples: *patata, ella, compramos, una, verdaderamente…* En el caso de las contracciones (*al* y *del*), contamos separadamente cada unidad.
- Palabras compuestas que forman una unidad de significado: cuantificadores (*un poco*), sustantivos (*colegio mayor*) o verbos (formas compuestas: *he estudiado*; perífrasis: *va a comer*; pasiva: *son obligados*; y verbos pronominales: *levantarse*).
- Locuciones verbales (*tomar en cuenta*), adjetivales (*de toda la vida*), preposicionales (*alrededor de*), adverbiales (*al final*) o conjuntivas (*para que*).
- Marcadores discursivos (*es decir, por ejemplo*).
- Fórmulas oracionales o rutinarias: *de nada* (para responder a un agradecimiento).
- Nombres propios (*Sierra Morena*), materias académicas (*Filología Clásica*), eventos históricos (*Segunda Guerra Mundial*) o títulos de películas (*El perro*).

Cuando una parte de la unidad léxica se repite (debido a las disfluencias de la oralidad: ej. *a [/] a veces*), consideramos la repetición dentro de una misma unidad. Exceptuamos los

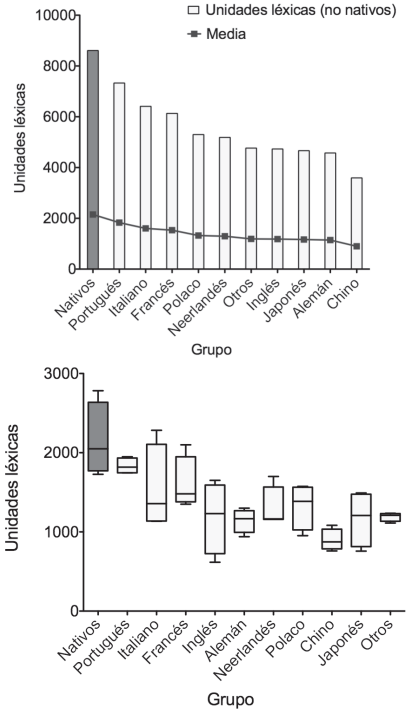
casos en que una parte aparece en otro turno porque el entrevistador se la proporciona. En el caso de los calcos formales (p. ej., **frutas de mar,* 'marisco'), preferimos contar conjuntamente todas las palabras, dado que forman una unidad semántica para expresar un concepto en la L1 de los aprendices (aunque no en español).

La tabla 2 y los gráficos 4 y 5 muestran el recuento de palabras en bruto y por unidades léxicas, así como la media y la desviación estándar (D) de cada grupo.

L1	Archivo	Recuentos en bruto				Unidades léxicas			
		Archivo	Grupo	Media	D	Archivo	Grupo	Media	D
Portugués	PORMA2	1860	7712	1928,00	103,76	1748	7329	1832,25	102,21
	PORWA2_1	1845				1745			
	PORWA2_2	1935				1887			
	PORWB1	2072				1949			
Italiano	ITAMA2	1710	6603	1650,75	569,69	1572	6133	1533,25	539,92
	ITAWA2	1200				1135			
	ITAMB1	2434				2283			
	ITAWB1	1259				1143			
Francés	FREMA2	1538	6760	1690,00	347,21	1467	6413	1603,25	337,13
	FREWA2	1563				1496			
	FREMB1	2206				2100			
	FREWB1	1453				1350			
Inglés	ENGWA2	1115	5009	1252,25	483,16	1049	4731	1182,75	451,18
	ENGMB1	1767				1651			
	ENGWB1_1	648				617			
	ENGWB1_2	1479				1414			
Neerlandés	DUTMA2	1195	5473	1368,25	279,22	1160	5186	1296,50	269,03
	DUTWA2_1	1230				1158			
	DUTWA2_2	1785				1700			
	DUTWB1	1263				1168			
Alemán	GERMA2	1241	4800	1200,00	148,38	1159	4573	1143,25	149,53
	GERWA2	1194				1174			
	GERWB1_1	1004				940			
	GERWB1_2	1361				1300			
Polaco	POLMA2_1	1280	5546	1386,50	310,29	1241	5300	1325,00	288,89
	POLMA2_2	1679				1575			
	POLMB1	1589				1531			
	POLWB1	998				953			
Chino	CHIWA2_1	895	3731	932,75	139,25	862	3592	898,00	135,50
	CHIWA2_2	1118				1084			
	CHIMB1	783				760			
	CHIWB1	935				886			
Japonés	JAPWA2	783	4953	1238,25	397,44	757	4664	1166,00	354,03
	JAPWB1_1	1028				985			
	JAPWB1_2	1549				1430			
	JAPWB1_3	1593				1492			
Otros	FINWA2	1289	4990	1247,00	66,66	1236	4767	1191,75	54,80
	HUNWA2	1148				1112			
	KORWB1	1273				1215			
	TURWB1	1280				1204			
Total		55577		1389,425	399,48	52688		1317,20	374,16
Español (grupo de control)	SPAM_1	2363	9389	2347,25	514,14	2195	8610	2152,50	462,47
	SPAM_2	3054				2783			
	SPAW_1	1859				1727			
	SPAW_2	2113				1905			

Tabla 2. Número de palabras (recuento en bruto) y de unidades léxicas por grupos de L1

En nuestro corpus, la producción más baja se recoge los grupos chino (con diferencia), alemán y japonés (por este orden). En cambio, los lusófonos, franceses e italianos presentaron la producción más alta. Estos resultados podrían asociarse a causas interlingüísticas (la cercanía o lejanía entre la L1 y el español explicaría sus dificultades de expresión), pero también a factores socioculturales (su participación activa o pasiva).



Figuras 4 y 5. Unidades léxicas; frecuencia absoluta y media de grupo, y representación de desviación

5.2. Categorías más utilizadas

Las listas de frecuencia de categorías se obtuvieron procesando informáticamente los archivos anotados con GRAMPAL y haciendo uso de la herramienta SketchEngine (Kilgarriff *et al.*, 2004). La tabla 3 desglosa el número de unidades léxicas por categoría en cada grupo (el total

de unidades se ordena de mayor a menor frecuencia absoluta).

Cat / L1	POR	ITA	FRE	ING	NER	ALE	POL	CHI	JAP	OTH	TOTAL	NATIVOS
Adverbios	1043	832	1093	884	792	898	884	668	843	889	8826	935
Verbos	1336	1070	1041	770	875	709	796	566	761	861	8785	1631
Sustantivos	1112	894	907	685	733	661	839	655	831	703	8020	1183
Conjunciones	771	649	718	516	646	468	550	329	539	601	5787	983
Preposiciones	691	623	595	378	470	422	604	284	347	381	4795	764
Artículos	572	554	550	359	441	364	318	270	221	252	3901	685
Adjetivos	301	219	232	134	194	170	253	167	219	212	2101	298
Pronombres	319	323	174	154	193	153	185	121	149	185	1956	592
Marc. discurs.	201	197	352	191	155	133	109	72	163	127	1700	582
Indefinidos[4]	268	212	150	151	199	104	217	115	126	160	1702	289
Nom. propios	152	92	122	106	85	86	115	96	128	103	1085	65
Numerales	71	76	97	41	72	60	52	55	68	51	643	91
Interjecciones	71	15	61	99	54	81	35	55	84	45	600	32
Relativos	94	129	98	33	78	23	44	16	7	28	550	208
Desconocidas	79	40	19	58	53	75	117	40	27	25	533	49
Demostrativos	97	70	36	23	46	43	68	22	44	37	486	98
Interrogativos	57	25	54	42	46	35	58	25	38	35	415	40
Posesivos	47	59	54	44	25	26	33	23	42	49	402	49
Fórmulas	35	39	53	57	23	56	17	1	24	14	319	26
Auxiliares[5]	12	15	7	6	6	6	6	12	3	9	82	10
Total	7329	6133	6413	4731	5186	4573	5300	3592	4664	4767	52688	8610

Tabla 3. Distribución de unidades léxicas por categorías en cada grupo de L1

Se puede observar que el uso de ciertas categorías es muy pobre en ciertos grupos de alumnos, lo cual revelaría problemas interlingüísticos por ausencia en su L1, a saber:

- Artículos entre polacos, chinos, japoneses y el grupo heterogéneo.
- Relativos y preposiciones entre japoneses y chinos.
- Conjunciones, marcadores discursivos y fórmulas orales en el grupo chino.

6. Análisis de errores

En este trabajo se ha distinguido entre errores no ambiguos y ambiguos. Por lo que se refiere a estos últimos, hemos registrado un total de 304 ambigüedades, solo el 4,45% de todos los errores (6838). Esto quiere decir que el 95,55% del análisis de errores realizado no representa una grave ambigüedad, y que la proporción entre ambiguos y no ambiguos que hemos encontrado es de un error ambiguo cada más de 22 aproximadamente. En la mayoría de las ocasiones la ambigüedad estriba en el nivel lingüístico del error, siendo tan solo unos pocos (16) ambiguos respecto a otros criterios, aunque resulta clara su adscripción a un determinado nivel. Los tipos de ambigüedades registradas han sido las siguientes (figura 6):

- Gramática-léxico (4,61%): p. ej., *me gusta // ¿ película ?* (gramática, uso del plural: 'las películas'; léxico: 'el cine').
- Gramática-pronunciación (14,47%): p. ej., *llegue* ['ʝeɣe] (acento o error verbal).
- Léxico-pronunciación (75,66%): ej., *lengua* ['liŋgwa] (préstamo o fonética).
- Pronunciación-pragmática (0,33%): p. ej. *la camarera* [kama'ɾerə] *responda que…* (pronunciación de vocal /a/, o problema de comprensión de la historia).
- Gramática-pragmática (0,33%): p. ej., [Refiriéndose en segunda persona en un contexto informal] *¿Quiere ir a otro restaurante?* (gramática, uso de la persona verbal correcta; pragmática, uso de un registro no formal).

- Ambigüedades dentro del nivel gramatical (3,62%): *cambio la hora* (omisión de pronombre: 'se cambia'; concordancia de persona: 'cambias').
- Ambigüedades en el nivel léxico (0,99%): *tomas un *plazo* (selección léxica de verbo: 'necesitas un plazo [para pedirlo]' o relación semántica: 'pides una plaza').

El mayor número de ambigüedades se registró, con diferencia, entre nivel léxico y de pronunciación; y en segundo lugar, entre gramática y pronunciación (figura 6).

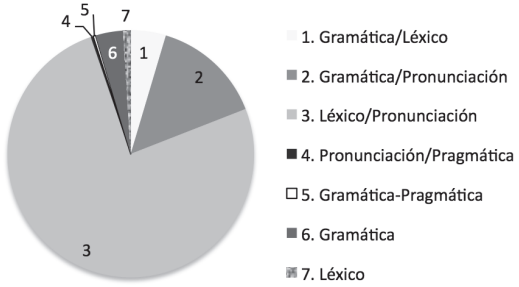


Figura 6. Frecuencia de cada tipo de error ambiguo

La tabla 5 y la figura 7 resumen las cifras de error de cada tipo y nivel lingüístico.

Errores			*Total*
	Nivel lingüistico	*Total*	*(%)*
No ambiguos respecto al nivel	Pronunciación	970	14,19%
	Gramática	3324	48,61%
	Léxico-semántico	2008	29,37%
	Pragmática-Discurso	245	3,58%
	No asignado	3	0,04%
Ambiguos		288	4,21%
Total		6838	

Tabla 4. Frecuencia relativa (%) de errores no ambiguos y ambiguos por nivel lingüístico

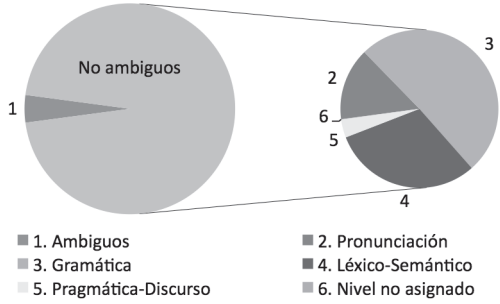


Figura 7. Errores ambiguos y no ambiguos (estos, desglosados por niveles lingüísticos).

Los errores de gramática son los más frecuentes en el habla de los alumnos, seguidos de los errores léxico-semánticos. Estos resultados coinciden con el análisis de errores escritos realizado por Fernández López (1990). A pesar de estas cifras, no todos los errores de gramática impiden la comunicación; p. ej., muchos errores de concordancia (p. ej., *las cosas *españoles* -> españolas), mientras que ciertos errores léxicos son realmente ambiguos respecto a la información expresada. Cabe señalar que no clasificamos en ningún nivel 3 errores que parecen más bien desviaciones idiosincrásicas o de adecuación (un hablante nativo no usaría tales construcciones).

5.1. Recuento global

El total de errores marcados es 6838, incluyendo errores no ambiguos (6534, el 95,55% del total de errores) y ambiguos (304, el 4,45%), con una media de 171 errores por estudiante prácticamente (tabla 6). Como se esperaba de los resultados obtenidos en otros análisis a partir de textos escritos (Fernández López, 1997 [1990]: 65), la tasa de error en el nivel A2 cae notablemente en el nivel B1. Las tablas 6 y 7 y el gráfico 8 muestran la frecuencia absoluta y relativa (expresada en %) de los errores (no ambiguos y ambiguos) respecto al nivel de competencia (*MCER*).

TOTAL DE ERRORES (VALORES ABSOLUTOS)						
Nivel	*No ambiguos*		*Ambiguos*		*Total*	
	Err.	*Media*	*Err.*	*Media*	*Err.*	*Media*
A2	3826	191,30	218	10,90	4044	202,50
B1	2708	135,40	86	4,30	2794	139,25
Total	6534	163,35	304	7,60	6838	170,63

Tabla 5. Frecuencia de errores no ambiguos y ambiguos de acuerdo al nivel del estudiante (MCER).

TOTAL DE ERRORES (PORCENTAJES)				
Nivel MCER	No ambiguos		Ambiguos	
	Fre. relativa	% sobre un. léxica	Frec. relativa	% sobre un. léxica
A2	56,95% (sobre 6838)	14,54% (26317)	3,19% (sobre 6838)	0,83% (26317)
B1	39,60% "	10,27% (26371)	1,26% "	0,33% (26371)
Total	95,35% "	12,40% (52688)	4,45% "	0,58% (52688)

Tabla 6. Frecuencia relativa de errores y proporción por unidad léxica de cada nivel (expresadas en %).

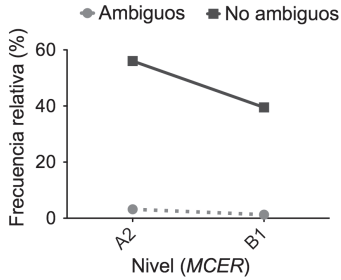


Figura 8. Frecuencia relativa de errores por nivel (% sobre el total de errores cometidos)

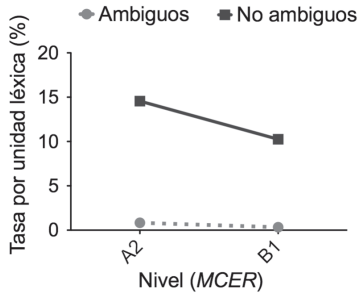


Figura 9. Tasa de error por unidad léxica por nivel (% sobre el total de unidades léxicas de cada nivel)

Los siguientes son los 20 errores más frecuentes según el corpus y nuestra corrección:

Tipo de error	*Nivel lingüístico*	*No ambiguos*	*% del total de errores (6838)*
1. Extranjerismos	Léxico	875	12,80%
2. Artículo	Gramática	634	9,27%
3. Preposiciones	Gramática	541	7,91%
4. Formación no atestiguada	Léxico	359	5,25%
5. Pronombres personales	Gramática	302	4,42%
6. Género	Léxico	272	3,98%
7. Estructura de la oración	Gramática	248	3,63%
8. Concordancia de persona	Gramática	202	2,95%
9. Relación semántica	Léxico	196	2,87%
10. Subordinación	Gramática	187	2,73%
11. Concordancia de número	Gramática	184	2,69%
12. Concordancia de género	Gramática	183	2,68%
13. Tiempos de pasado	Gramática	152	2,22%
14. Conjugación	Gramática	136	1,99%
15. Marcadores discursivos	Pragmático-discursivo	133	1,95%
16. *Ser/estar*	Gramática	122	1,78%
17. Oclusivas sordas/sonoras	Pronunciación	122	1,78%
18. Acento	Pronunciación	118	1,73%
19. Cambios de punto/modo art.	Pronunciación	72	1,05%
20. Pérdida de consonantes	Pronunciación	71	1,04%

En ciertos puntos las cifras realmente reflejan una frecuencia de error acusada solo en ciertos alumnos: los extranjerismos se concentran en el grupo lusófono; los artículos, en los japoneses, chinos, polacos y el grupo heterogéneo; los cambios de acento, sobre todo en el grupo francés; y la no distinción entre oclusivas sordas y sonoras, en los chinos. Los errores más frecuentes y generalizados se relacionan con las preposiciones, las formaciones no atestiguadas, los pronombres personales, la asignación de género (excepto en el grupo italiano), la estructura de la oración, la subordinación, las concordancias y los tiempos de pasado. Por niveles, se pueden apuntar estas tendencias:

- En léxico, sobresalen los errores formales: en A2, deformaciones y extran-jerismos (existe gran transferencia entre nativos de lenguas románicas, sobre todo lusófonos); y aunque en B1 disminuyen su frecuencia, persisten otros errores por relación semántica y asignación de género.
- En gramática, además de los puntos comentados, destaca la alta cifra de incorrecciones por omisión y orden de palabras, las discordancias y el abuso del presente sobre otros tiempos, que se explicaría por las características de la oralidad.
- En pronunciación, el problema más generalizado es el sonido [r], aunque cada grupo tiende a transferir rasgos: p. ej. chinos y japoneses no distinguen /ɾ/ y /l/, o los nativos de lenguas germánicas cometen más errores por reducción vocálica.
- En pragmática-discurso destacan los errores de cohesión (p. ej. marcadores discursivos), aunque se registra gran variabilidad individual entre los alumnos.

Respecto a la etiología de los errores, cabe decir que los resultados han de tomarse con prudencia, ya que en muchos contextos presuponen conocer el mensaje pretendido por el hablante, dando rienda suelta a nuestras (mejores o peores) capacidades interpretativas. De hecho, cuando consultamos la literatura de investigación sobre el tema, constatamos las diferentes conclusiones a las que llegan distintos investigadores. Por ejemplo, según Dulay

y Burt (1976), los errores de interferencia no llegan al 5%, mientras que Nickel (1981) los cifra en una media del 60% (otros lingüistas calculan incluso hasta un 80%); en realidad, la definición de lo que se considera *errores de interferencia* condiciona los resultados (véase un resumen en Lott, 1983). Con todo, estimamos que al menos el 49,21% de las incorrecciones se debe a causas interlingüísticas (tabla 8 y figura 10), lo cual parece estar en la línea de los resultados de Lott (1983) para el inglés escrito, o de Valverde Mateos (2012) respecto al francés oral.

Etiología	*Errores*			
	No ambiguos	*Ambiguos*	*Total*	*F. r. (%)*
1. Interlingüísticos	3274	91	3365	49,21%
2. Intralingüísticos	1534	27	1561	22,83%
3. Desconocida	1726	46	1772	25,91%
4. Varias causas	-	140	140	2,05%
Total	6534	304	6838	

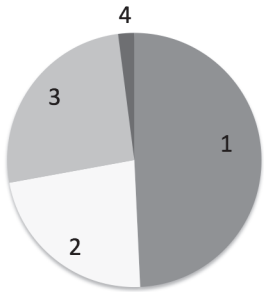


Tabla 7 y Figura 10. Etiología estimada de los errores (frecuencia absoluta y relativa en %)

La distribución por niveles revela un descenso de los interlingüísticos de A2 a B1, y al contrario, un leve aumento de intralingüísticos, como apuntan otros trabajos (Vázquez, 1999) (tabla 9 y gráfico 11; la frecuencia relativa se mide respectivamente sobre el total de errores de cada nivel: por una parte, los no ambiguos; por otra, los ambiguos; y el total, sobre ambos; la cifra de cada uno se indica entre paréntesis). La alta tasa de errores de etiología desconocida se debe a nuestro criterio prudente de marcado.

Errores		*Etiología*									
		Interlingüísticos		*Intralingüísticos*		*Desconocida*		*Varias causas*[4]			
		Frec	*Frec. rel. (%)*	*Frec*	*Frec. rel. (%)*	*Frec*	*Frec. rel. (%)*	*Frec*	*Frec. rel. (%)*
A2	No ambiguos	2024	52,92% (3825)	809	21,15% (3825)	988	25,93% (3825)	-	-
	Ambiguos	85	38,99% (218)	9	4,13% (218)	45	20,64% (218)	79	36,24% (218)
	Total	2109	52,16% (4043)	818	20,35% (4043)	1033	25,65% (4043)	79	1,95% (4043)
B1	No ambiguos	1249	46,11% (2709)	726	26,80% (2709)	725	27,09% (2709)	-	-
	Ambiguos	6	6,98% (86)	18	20,93% (86)	1	1,16% (86)	61	70,93% (86)
	Total	1255	44,90% (2795)	744	26,62% (2795)	726	26,30% (2795)	61	2,18% (2795)

Tabla 8. Distribución de errores por nivel según la etiología (frecuencia absoluta y relativa en %)

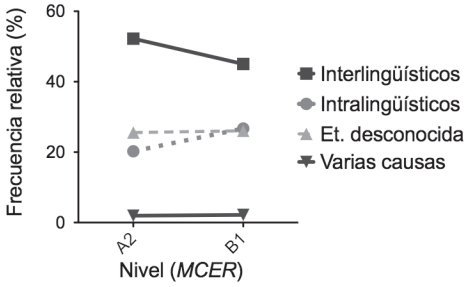


Figura 11. Etiología de errores (no ambiguos y ambiguos) por nivel (frecuencia relativa en %)

6. Conclusiones

Esta investigación ha analizado la interlengua oral de cuarenta estudiantes de español. Los resultados del análisis formal de la producción en nuestro corpus son los siguientes:

- En el recuento de unidades, apenas hay diferencia entre A2 y B1, pero es más baja en chinos, alemanes y japoneses, y más alta en lusófonos, franceses e italianos.
- En grupos específicos de L1 hay una producción más pobre de ciertas categorías, quizá por causas interlingüísticas: artículos (polacos, chinos y japoneses), relativos y preposiciones (japoneses) y conjunciones y marcadores discursivos (chinos).

El análisis de errores de nuestro corpus reveló las siguientes dificultades:

- El progreso de A2 a B1 muestra un descenso de errores (de una media de 191,30 a 135,40) aunque esta disminución puede deberse a la evitación de estructuras.
- Los errores más abundantes afectan a la gramática (48,61%) y el léxico (29,37%) –frente al 14,19% de la pronunciación, y el 3,58% de la pragmática o el discurso)–; aproximadamente, un 4,45% son ambiguos, y el 49,21% se deberían a interferencia.

- Las incorrecciones del nivel léxico se concentran en los aspectos formales más que en los semánticos (que son más resistentes a la adquisición).
- Las errores gramaticales más frecuentes y generalizados afectan al artículo, la estructura de la oración, las concordancias y los tiempos de pasado; además, en B1 persisten los que se registran en pronombres, preposiciones y la subordinación.
- En la pronunciación, la interferencia suele persistir en B1 y quizá es donde más influye la L1, aunque algunos errores son generalizados, como la articulación de /r/.
- Los errores pragmático-discursivos revelan gran variación individual no siempre relacionable con el nivel, y quizá dependen de las propias habilidades del alumno; con todo, los chinos muestran relaciones de cohesión más pobres.

Entre las limitaciones de nuestro estudio, se pueden destacar las siguientes:

- Al basarse en datos orales, no es claro si ciertas desviaciones se deben a la competencia o la actuación, o respecto al tipo de dificultad o el nivel lingüístico.
- Solo reunimos datos de alumnos de nivel intermedio-bajo, y no es posible extraer conclusiones acerca de su capacidad para lograr una competencia casi bilingüe.
- No es posible generalizar los resultados dado el reducido número de alumnos participantes en cada grupo, aunque sí se vislumbran tendencias particulares.

Se pueden proyectar las siguientes tareas futuras de investigación:

- Ampliar el corpus con más participantes, más lenguas maternas y más niveles.
- Confirmar con pruebas complementarias (ej., escritas) el tipo o el nivel lingüístico de algunos fenómenos, o si se deben a la competencia o la actuación.
- Validar la taxonomía de errores entre dos anotadores independientes y con un subconjunto del cor-

pus, estimando el grado de coincidencia entre analistas y la consistencia del etiquetado (mediante el índice Kappa, κ; Herrera *et al.*, 2011: 194).

Si bien el alcance de este proyecto es reducido y los resultados no son todos generalizables, la principal aportación del estudio es un análisis de la producción y de los errores del español oral no nativo con una metodología rigurosa y extendible a otro corpus. Más aún, el banco de datos recogido ha sido procesado informáticamente y es accesible para la consulta por medio de una interfaz en línea (figuras 12 y 13). Creemos ciertamente que será aprovechable tanto por la comunidad investigadora en la adquisición de lenguas como por los docentes de español/LE.



Figura 12. Muestra de la pantalla inicial de la interfaz de consulta al corpus

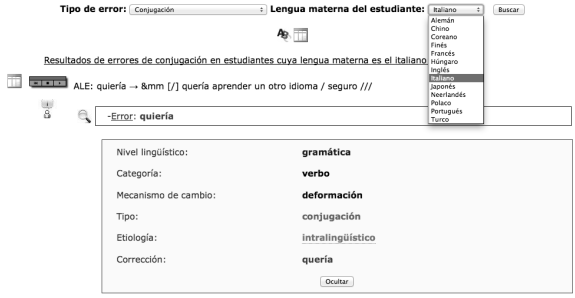


Figura 13. Muestra de consulta de errores mediante la interfaz informática en línea

BIBLIOGRAFÍA

BARLOW, M. (2005) «Computer-based analyses of learner language», en Ellis y Barkhuizen, *op. cit.*, chapter 14, pp. 335-357.

CAMPILLOS LLANOS, L. (2012) *La expresión oral en español lengua extranjera: interlengua y análisis de errores basado en corpus.* Tesis doctoral inédita. Universidad Autónoma de Madrid. Dpto. de Lingüística y Lenguas modernas.

CORDER, P. (1971/1991) «Idiosyncratic dialects and error analysis». *International Review of Applied Linguistics*, nº 9, 2, pp. 147-60. Traducción al español: «Dialectos idiosincrásicos y análisis de errores», en J. Muñoz Liceras (comp.) (1991) *La adquisición de las lenguas extranjeras*, pp. 63-77. Madrid: Visor.

CRESTI, E., y M. MONEGLIA (eds.) (2005). *C-ORAL-ROM. Integrated Reference Corpora for Spoken Romance Languages*. Amsterdam: John Benjamins.

DAGNEAUX, E., S. Denness, y S. Granger (1998) «Computer-aided error analysis». *System*, 26(2), pp. 163-174.

DÍAZ-NEGRILLO, A., y J. FERNÁNDEZ DOMÍNGUEZ (2006) «Error tagging systems for learner corpora», *RESLA*, pp. 83-102. Disponible en: <http://dialnet.unirioja.es/ servlet/dcfichero_articulo?codigo=2198610&orden=72810> [11/01/2013]

DÍAZ-NEGRILLO, A., y S. VALERA, (2009) «What results give which strategies in error annotation?». *CALICO* Pre-Conference Workshop.
<www.sfs.uni-tuebingen.de/~dm/events/aall09/diaz-negrillo-valera.pdf> [11/01/2013]

DÖRNYEI, Z. (2007) *Research Methods in Applied Linguistics.* Oxford: Oxford University Press.

DULAY, H. C., y M. K. BURT (1976) «Creative construction in second language learning and teaching». *Language Learning,* número especial 4.

DUSSIAS, P. E. (2003) «Cognitive Perspectives on the Acquisition of Spanish as a Second Language». En B. Lafford y R. Salaberry (eds.) *Spanish Second Language Acquition,* capítulo 8, pp. 233-261.

ELLIS, R., y G. BARKHUIZEN (2005) *Analyzing Learner Language*. Oxford: Oxford University Press.

FERNÁNDEZ LÓPEZ, S. (1990) *Análisis de errores e interlengua en el aprendizaje del español como lengua extranjera.* Tesis doctoral. Universidad Complutense de Madrid. Facultad de Filología, Dpto. de Filología Románica.

GILQUIN, G. (2008) «Combining contrastive and interlanguage analysis to apprehend transfer: detection, explanation, evaluation». En Gaëtanelle Gilquin, Szilvia Papp y María Belén Díez-Bedmar (eds.) (2008) *Linking up contrastive and learner corpus research*. Amsterdam/New York, NY: Rodopi. Language and Computers - Studies in Practical Linguistics Series, 66.

GRANGER, S. (2003) «Error-tagged Learner Corpora and CALL: a promising synergy», *CALICO Journal*, 20 (3), pp. 465-480.

GRANGER, S. (2008) «Learner corpora». En Lüdeling & Kytö (eds.), *op. cit.*, pp. 259–275.

GRANGER, S., O. Kraifa, C. Pontona, G. Antoniadisa, y V. Zampa (2007) «Integrating learner corpora and natural language processing: A crucial step towards reconciling technological sophistication and pedagogical effectiveness». *ReCALL*, 19, pp. 252-268.

HEIFT, T. y M. SCHULZE (2007) *Errors and Intelligence in Computer-Assisted Language Learning: Parsers and Pedagogues.* London: Routledge.

HERRERA SOLER, H., R. Martínez Arias, y M. Amengual Pizarro (2011) *Estadística aplicada a la investigación lingüística.* Madrid: EOS Universitaria.

IZUMI, E., K. UCHIMOTO, y H. ISAHARA (2004) «SST Speech corpus of Japanese learners' English and automatic detection of learners' errors». *ICAME Journal,* nº. 28, pp. 31-48. Documento accesible en:
<http://icame.uib.no/ij28/Izumi.pdf> [11/01/2013]

IZUMI, E., K. Uchimoto y H. Isahara (2005) «Error Annotation for Corpus of Japanese Learner English». En *Proceedings of the Sixth International Workshop on Linguistically Interpreted Corpora (LINC 2005),* pp. 71-80. <http://aclweb.org/anthology-new/I/I05/I05-6009.pdf> [11/01/2013]

JAMES, C. (1998) *Errors in Language Learning and Use. Exploring Error Analysis.* London/New York: Longman. Applied Linguistics and Language Study Series.

KILGARRIFF, A., P. RYCHLY, P. SMRZ, y D. TUGWELL (2004) «The Sketch Engine». En *Proc. EURALEX 2004, Lorient, France,* pp 105-116. <http://www.sketchengine.co.uk>

KÖHLMYR, P. (2003) *«To Err is Human…» An investigation of Grammatical Errors in Swedish 16-year-old Learners' Written Production in English.* Göteborg: Acta Universitatis Gothoburgensis.

LARSEN-FREEMAN, D., y M. H. LONG (1991)[1994] *An Introduction to Second Language Acquisition Research.* London/New York: Longman. Applied Linguistics and Language Study.

LAFFORD, B., y M. R. SALABERRY (2003) *Spanish second language acquisition: state of the science.* Washington: Georgetown University Press.

LOTT, D. (1983) «Analysing and counteracting interference errors». *ELT Journal*, vol. 37/3, pp. 256-261.

LÜDELING, A., M. WALTER, E. KROYMANN, y P. ADOLPHS (2005) «Multi-level error annotation in learner corpora». *Proc. of the Corpus Linguistics Conference 2005.* <http://www2.huberlin.de/korpling/projekte/falko/FALKOCL2005.pdf?PHPSESSID=5bf51ebe3e6c9d694c16ffc9b27efc15> [11/01/2013]

LÜDELING, A., y M. KYTÖ (eds.) (2008) *Corpus Linguistics. An International Handbook*. 2 Vols. Berlin: Mouton de Gruyter.

MACKEY, A., y S. Gass (2005) *Second Language Research. Methodology and Design*. London/N.J.: Lawrence Erlbaum Associates.

MARCOS MARÍN, F. (1994) *Informática y humanidades.* Madrid: Gredos.

MCENERY, A., R. XIAO, y Y. TONO (2006) *Corpus-Based Language Studies. An Advanced Resource Book*. London: Routledge (Routledge Applied Linguistics).

MITCHELL, R., L. DOMÍGUEZ, Mª. ARCHE, F. MYLES y E. Marsden (2008) «Research Report: Linguistic Development in L2 Spanish: Creation and analysis of a learner corpus». En Roberts, L., F. Myles, y A. David. *EuroSLA Yearbook, 8.* Amsterdam; Philadelphia: John Benjamins, pp. 287-304. Documento disponible en: <http://www.splloc.soton.ac.uk/doc/SPLLOC_1_Report_Final.pdf> [11/01/2013]

MORENO SANDOVAL, A. (1998) *Lingüística computacional.* Madrid: Síntesis.

MORENO SANDOVAL, A., y GUIRAO, J. M. (2006) «Morpho-syntactic Tagging of the Spanish C-ORAL-ROM Corpus: Methodology, Tools and Evaluation». En Y. Kawaguchi, S. Zaima y T. Takagaki (eds.) *Spoken Language Corpus and Linguistic Informatics,* pp. 199-218. Amsterdam: John Benjamins. Una demostración de GRAMPAL está en: <http://cartago.lllf.uam.es/grampal/grampal.cgi> [11/01/2013]

MYLES, F. (2005) «Interlanguage corpora and second language research». *Second Language Research,* 21(4), pp. 373-391.

NEMSER, W. (1971) «Approximative systems of foreign language learners». *International Review of Applied Linguistics*, IX (2), pp. 115-123.

NICHOLLS, D. (2003) «The *Cambridge Learner Corpus* – error coding and analysis for Lexicography and ELT». En Archer *et al.* (eds.) *Proceedings of the Corpus Linguistics 2003 Confe-*

crence (CL2003). Lancaster University: University Centre for Computer Corpus Research on language, pp. 572-581. Disponible en: <http://ucrel.lancs.ac.uk/publications/CL2003/papers/nicholls.pdf> [11/01/2013]

NICKEL, G. (1981) «Aspects of error analysis: "errare humanum est"». *AILA Bulletin,* 29, pp. 1-27.

SELINKER, L. (1972) «Interlanguage», *International Review of Applied Linguistics*, vol. 10, nº 3, pp. 209□301. Traducción española: «La interlengua», en J. Muñoz Liceras (1992) *La Adquisición de las lenguas extranjeras: hacia un modelo de análisis de la interlengua.* Madrid: Visor. Pp. 79-101.

VALVERDE MATEOS, A. (2012) *Análisis de errores de aprendientes de Francés Lengua Extranjera (FLE) basado en corpus orales.* Tesis doctoral sin publicar. Madrid: Universidad Autónoma de Madrid. Dpto. de Lingüística y Lenguas modernas.

VÁZQUEZ, G. (1999) *¿Errores? ¡Sin falta!* Madrid: Edelsa.

PROGRAMAS

Transana 2.05. Universidad de Wisconsin. <www.transana.org> [11/01/2013]

NOTAS

[*] Trabajo financiado por la Comunidad de Madrid y el Fondo Social Europeo (contrato predoctoral).

[1] El acceso a la interfaz se encuentra disponible en: http://cartago.lllf.uam.es/corele/index.html

[2] En los errores interlingüísticos, marcamos información sobre de la(s) lengua(s) que quizá interfiere(n).

[3] Agradecemos al Dr. José Mª. Guirao de la Universidad de Granada el desarrollo y la puesta a disposición de dicha herramienta informática, que nos ha agilizado este proceso.

[4] En los cuantificadores indefinidos reunimos formas como *nadie, algo* o *un poco*.

[5] El recuento de auxiliares aquí presentado corresponde solamente a los que no pertenecen a las formas verbales complejas (tiempos compuestos, perífrasis o pasiva), las cuales se computan dentro de la categoría del verbo. Si contáramos estos casos, la producción total de auxiliares sería de 992.

[6] Los errores ambiguos tienen dos interpretaciones, y a veces dos diagnósticos distintos; p. ej., **creo [ˈkrei]* puede ser de conjugación (intralingüístico) o de pronunciación (por factores desconocidos).

LA ANOTACIÓN DE LA NEGACIÓN EN UN CORPUS ESCRITO ETIQUETADO SINTÁCTICAMENTE
ANNOTATION OF NEGATION IN A WRITTEN TREEBANK

Antonio Moreno Sandoval y Marta Garrote Salazar
Universidad Autónoma de Madrid

Resumen: En este artículo se describe la anotación de la negación y su ámbito de afectación en un corpus formado por 1501 oraciones extraídas de textos periodísticos y anotadas sintácticamente. En ellas, se han detectado las marcas de negación, los elementos de concordancia negativa y su alcance. Se muestran además los resultados estadísticos preliminares respecto a la frecuencia y funcionamiento de los elementos negativos anotados en el corpus. Este trabajo supone una primera aproximación al tratamiento computacional de la negación. El corpus anotado está disponible de forma gratuita para investigación.
Palabras clave: Anotación de la negación, corpus anotado sintácticamente, marcas de negación, ámbito de la negación.

Abstract: In this paper, it is described the annotation of negation and its scope in a corpus made up of 1501 sentences extracted from newspapers and syntactically annotated. For each sentence, negation cues, negative concordance and their scope have been marked. Also, we show some preliminary statistical results regarding frequency and function of negative elements annotated in the corpus. This work is a first approach to the computational treatment of negation. The annotated corpus is freely available.
Keywords: Negation annotation, treebank, negation cues, negation scope.

1. Introducción

En los últimos años se ha visto un auge de algunas tareas lingüístico-computacionales como la anotación de corpus para el aprendizaje automático conocido como Inferencia Gramatical (Nivre 2006; Seginer 2007). Básicamente, con-

siste en aprender o inferir las reglas gramaticales de una lengua a partir del análisis sintáctico realizado por lingüistas. Tomando la muestra de datos anotados, el programa propone la estructura sintáctica más probable para una oración dada. Esta inducción de reglas gramaticales se realiza mediante métodos estadísticos.

En este trabajo, pretendemos aportar algo de luz sobre uno de los desafíos en la formalización del lenguaje natural: la negación. Mucho se ha escrito sobre ella en el plano teórico de la Lingüística; sin embargo, las teorías no siempre son suficientes en la práctica de formalizar el funcionamiento de la negación para su procesamiento computacional. En los últimos años se ha producido un interés creciente por el tratamiento automático de la negación. El mejor ejemplo de ello es el número especial de la revista *Computational Linguistics* dedicado al tema (Morante y Sporleder 2012), donde se publican cinco artículos que tratan de temas generales aplicados al inglés.

Como explican Morante et al. (2011a:4), "Detecting negated events is important because negated events are not facts". Si un sistema de extracción automática de información se encuentra con la siguiente oración:

(1) La presidenta de Unió Mallorquina no descarta pactar con los socialistas.

debería ser capaz de detectar que el evento principal, *descartar*, está modificado por una marca de negación. Si el sistema extrae que dicha presidenta descarta un pacto con los socialistas, la información que obtenemos es falsa. Para evitar que esto ocurra, el tratamiento de la negación debe ser incorporado al sistema.

Otro ejemplo, del campo de las aplicaciones computacionales en análisis de la opinión, es distinguir entre "Juan lo hace mal", "Juan no lo hace mal" y "Juan no lo hace nada mal". La primera es una valoración negativa mientras que las dos últimas implican una apreciación positiva sobre el sujeto. Las tres contienen elementos negativos.

De entre los trabajos que se están llevando a cabo en este campo, destacaremos algunos por la similitud que tienen con nuestro proyecto: la anotación de la negación y su ámbito en un corpus compuesto por dos historias de Conan Doyle (Morante *et al.* 2011a) y el Bioscope Corpus (Szarvas *et al.*

2008) conjunto de textos médicos en los que se han anotado la negación y su alcance. Ambos corpus están disponibles de forma gratuita para fines de investigación.

Otra muestra de la actualidad del tema son las dos evaluaciones recientes: las tareas de resolución del ámbito y el foco de la negación en *SEM 2012 (Morante y Blanco 2012), donde se anotaron dos conjuntos de datos del PropBank. Otro corpus similar es el *gold standard* elaborado para la tarea *Processing modality and negation* en QA4MRE en el CLEF2012. Este último recurso contiene 8 documentos y 1244 eventos, es decir, en torno a unas 1000 oraciones. Todos ellos son conjuntos de datos en inglés[1].

Este documento, en la línea de los mencionados proyectos, describe una primera etapa del proceso de anotación de información negativa en un corpus de árboles sintácticos, el UAM Spanish Treebank. La principal razón para trabajar con este tipo de corpus y no con otros menos estructurados, mencionados anteriormente es que la anotación sintáctica existente facilita nuestra tarea, especialmente para determinar el ámbito de la negación, y nos permitirá extraer conclusiones más complejas en el futuro. Todas las oraciones y sintagmas negativos han sido anotados, incluyendo tanto las marcas de negación como su ámbito o alcance. El resultado, en formato XML, se puede conseguir de forma gratuita en la dirección http://www.lllf.uam.es/ESP/Treebank.html.

El artículo se estructura de la siguiente manera: en primer lugar, se hace un breve repaso de la situación lingüística teórica de la negación y se expone nuestro modelo de anotación; en segundo lugar, se describe de forma escueta el corpus utilizado para nuestro trabajo; a continuación, se explica el procedimiento de anotación; en cuarto lugar, se muestran algunos de los resultados en términos cuantitativos; por último, se exponen las conclusiones y el trabajo previsto en el futuro. Los ejemplos utilizados pertenecen todos ellos al UAM Spanish Treebank.

2. Marco teórico

La negación se ha tratado en el plano teórico desde perspectivas gramaticales, semántico-lógicas y pragmáticas. Es

precisamente su carácter interdisciplinar lo que la convierte en un elemento lingüístico difícil de formalizar en términos computacionales. Igualmente, ese carácter de la negación la hace computacionalmente inabarcable en un solo paso. Por ello es necesario desglosar su análisis en diferentes fases, desde la puramente formal hasta la más pragmática o dependiente del contexto. Nuestro análisis comienza pues en el nivel gramatical, partiendo de tres problemas fundamentales: las diferentes formas o *manifestaciones lingüísticas básicas* de la negación, la *concordancia negativa* y el *ámbito* de la negación.

Bosque y Gutiérrez-Rexach (2009) destacan la asimetría entre la naturaleza semántica uniforme de la negación y su polivalencia categorial: adverbios (*no*, *tampoco*), pronombres (*nadie*, *nada*), conjunciones (*ni*), preposiciones (*sin*) etc. Todos ellos pueden funcionar como Inductores Negativos (IN), elementos lingüísticos que legitiman a los Términos de Polaridad Negativa (TPN) o palabras y sintagmas que sólo pueden aparecer en entornos negativos (Sánchez 1999, RAE-AALE 2009). En (2), *no* es el IN y *nada* el TPN.

Él *no* hace *nada* por eliminar esa imagen.

Los TPNs son palabras negativas, o palabras-n (*nadie*, *nada*, *ninguno*, *nunca*, *jamás* y *tampoco*) que expresan el significado negativo por sí mismas cuando preceden al verbo. En esta posición, pueden legitimar a otro TPN.

Ayer *nadie* comentó *nada* sobre su estado de salud ni sobre las causas de la enfermedad.

Además, presentan concordancia negativa con el IN, una propiedad morfosintáctica propia de las lenguas románicas (Bosque y Gutiérrez-Rexach, 2009). En (3) no se produce una doble negación, sino la expresión de la negación en dos constituyentes. Esta propiedad diferencia a las palabras-n de aquellas otras palabras con significado negativo intrínseco que combinadas con un IN provocarían una doble negación ("*No* es *im*posible" equivale a la afirmación "Es posible").

Nuestras pautas de anotación contemplan todos los elementos que puedan funcionar como IN y como TPN coordinado con un IN.

Pero la identificación de marcas negativas en un texto no es suficiente: es necesario detectar el *ámbito de la negación*, es decir, el dominio sintáctico sobre el que esta tiene efecto (Sánchez 1999). Sin embargo, el ámbito de la negación es ambiguo

ya que "[…] las interpretaciones de la negación son sumamente flexibles […]" (Bosque y Gutiérrez-Rexach, 2009:638). La decisión más básica es si la marca de negación afecta a la oración entera o solamente a un constituyente.

La teoría gramatical (Sánchez, 1999, Moreno y De Molina 2002, RAE-AALE 2009) habla de *negación externa* o *metalingüística*, que es aquella que abarca toda la oración y *negación interna* o *descriptiva*, que alcanza parcialmente a la proposición o a algunos de sus componentes. Moreno y De Molina, (2002:10) afirman que "tanto la negación como los cuantificadores son operadores que tienen bajo su dominio a los operadores situados a su derecha". Así, se distingue entre negación oracional y sintagmática, la que afecta al sintagma al que precede, o morfológica, la que recae en unidades léxicas mediante prefijos negativos (esta última no se va a tratar en el presente trabajo). Sánchez (1999) también mantiene que el elemento negativo precede a su ámbito y lo domina sintácticamente.

Siguiendo estos supuestos sobre el ámbito de la negación en esta fase del proyecto se van a anotar los siguientes niveles:

1. La negación gramatical (con el inductor "no" o cualquier palabra-n que funcione como inductor en ausencia de este)
    1.1. Oracional: afecta a toda la oración.
        No + (P) + V
    1.2. Sintagmática: afecta sólo al sintagma.
        No + Q/N/ADJ/ADV
2. La negación léxica de las palabras-n
    2.1. Pronombres: *nadie, ninguno, nada.*
    2.2. Adverbios: *nunca, jamás, tampoco.*

Este ha sido el modelo formal empleado para anotar nuestro corpus. Las reglas de 1.1 y 1.2 siguen la teoría sobre la distribución del ámbito de la negación a la derecha del IN. Dependiendo de si a continuación del IN encontramos un verbo o no, la negación será oracional o sintagmática. La negación morfológica mediante prefijos (como *imposible)*, o interna al significado de la palabra (como *rechazar*) y la negación interoracional, representada en su mayoría por conectores oracionales, se tratarán en futuras fases del proyecto. Finalmente, tampoco se tratarán de momento los enunciados negativos sin IN ni TPN: *En lugar alguno, encontrará refugio.*

En resumen, el objeto de esta investigación es centrarse en los aspectos más generales y frecuentes del fenómeno, siempre desde una perspectiva sintáctica oracional. Para ello, se aprovecha el trabajo previo en análisis sintáctico y las herramientas computacionales que permiten localizar y anotar rápidamente las palabras en nuestro corpus. Esto permite realizar un recuento y dar una primera aproximación cuantitativa a estos fenómenos.

3. Descripción del UAM Spanish Treebank

El UAM Spanish Treebank (Moreno *et al.* 2003) es un corpus compuesto por 1501 oraciones anotadas sintácticamente (22695 palabras), todas ellas extraídas de periódicos (*El País Digital* y *Compra Maestra*). Su manual de anotación, disponible junto con el corpus, incluye un inventario de categorías y rasgos, el esquema de anotación e indicaciones específicas sobre una gran variedad de fenómenos del español. Los árboles están codificados en estructura anidada, con los elementos de cada nivel, incluyendo la categoría sintáctica, los rasgos sintácticos y semánticos y los nodos constituyentes, siguiendo el modelo del Penn Treebank. La estructura refleja la sintaxis superficial. La primera fase de este proyecto se desarrolló entre los años 1997 y 2000, y fue el primer corpus anotado sintácticamente del español.

El UAM Spanish Treebank se puede conseguir de forma gratuita, en la dirección http://www.lllf.uam.es/ESP/Treebank.html. Desde 2002, fecha en que se puso a disposición de los investigadores interesados, se han dado más de 40 licencias de uso. De entre las investigaciones más relevantes realizadas con este recurso podemos destacar la llevada a cabo por el grupo de investigación Pattern Recognition and Human Language Technology (PRHLT) en cuya página web (http://prhlt.iti.upv.es/) podemos consultar una herramienta de anotación sintáctica de oraciones. Recientemente, Santamaría (2013) ha demostrado en su tesis doctoral que el UAM Treebank es suficiente para conseguir resultados parecidos a los obtenidos con el PennTreebank con sistemas de inferencia gramatical no supervisada. La Tabla 1 muestra los resultados de la medida armónica F con los sistemas de inducción gramatical CCL (Seginer 2007) y MCM (Santamaría 2013), tomados de la tesis de este último. La evaluación se realizó sobre un sub-

conjunto de los corpus PennTreebank WSJ (inglés) y del UAM Treebank (español) con oraciones de menos de 10 palabras y con el con el corpus completo. El tamaño de los cuatro conjuntos es de 7422 oraciones (WSJ10), 49208 (WSJ-entero), 396 (UAM-10) y 1501 (UAM-entero).

Corpus	10		Completo	
	CCL	MCM	CCL	MCM
WSJ	75	88	56	70
UAM	67	78	59	65

Tabla 1. Evaluación de inferencia gramatical con dos sistemas no supervisados (Santamaría 2013).

Santamaría(2013) muestra que con un conjunto reducido de árboles se pueden obtener resultados bastante similares a los de otro mucho mayor.

Igualmente, llamamos la atención de que el tratamiento de la negación no era ni siquiera un tópico en los corpus sintácticos de primera generación, como se puede comprobar en la monografía de Abeillé (2003).

4. Metodología

La nueva extensión del UAM Spanish Treebank ha sido anotada por dos investigadores expertos, ambos especialistas en Lingüística de Corpus. Para ello, se ha seguido el modelo formal presentado en la sección 2. En los casos de ambigüedad, se ha llegado a un consenso tras una discusión sobre el fenómeno concreto y siempre manteniendo las pautas del modelo establecido[2].

El UAM Spanish Treebank se presenta en formato XML en estructura anidada. Cada una de las oraciones está representada por la etiqueta <Sentence> como elemento raíz dentro del cual se van anidando los diferentes elementos que constituyen la oración. Esto nos permite acotar el ámbito de la negación introduciendo el rasgo Neg="YES" en el constituyente sintáctico sobre el cual recae su significado. Si la negación es total, afectando a la oración completa, el rasgo se incluye en la etiqueta <Sentence>; si por el contrario el ámbito de la negación sólo abarcase, por ejemplo, un sintagma adjetival, representado como <ADJP>, el rasgo Neg="YES" se incluiría en esta etiqueta. El IN o marca de negación se ha marcado con

el rasgo Type=”NEG”, y cuando se produce el fenómeno de concordancia negativa, la palabra en función de TPN se marca con el rasgo Neg=”YES”. Los ejemplos de las Figuras 1, 2 y 3 muestran el resultado de la anotación.

En la Figura 1 vemos un ejemplo de negación oracional (*No comprendía la situación*) en el que se marca el ámbito de la negación (Neg=”YES”) y el IN (Type=”NEG”). La Figura 2 (*No pretende dar ninguna impresión concreta*) se muestra cómo se ha marcado la concordancia negativa: la palabra *ninguna* funciona como TPN concordando con el IN *No* y a su vez actúa como IN del sintagma nominal al que precede. Finalmente, en la Figura 3 (*Carlos Saura y José Luis Garci competirán por la estatuilla a la mejor película de habla no inglesa*) es una muestra de negación cuyo ámbito abarca solamente un sintagma (<ADJP Gender=”FEM” Number=”SG” *Neg=”YES”*>). Los árboles simplificados de las tres oraciones se muestran en la Figura 4. En ellos se puede observar el ámbito de alcance de la negación. Los elementos hermanos del IN (en todos estos casos, el adverbio negativo *no*, excepto en el último NP, cuyo IN es el cuantificador *ninguna*) son núcleos de un constituyente mayor: en el caso del verbo, este es núcleo de la oración; en el caso de otras categorías sintácticas (N, ADJ, etc.), estas son núcleos de un sintagma. Ese constituyente mayor es el alcance de la negación.

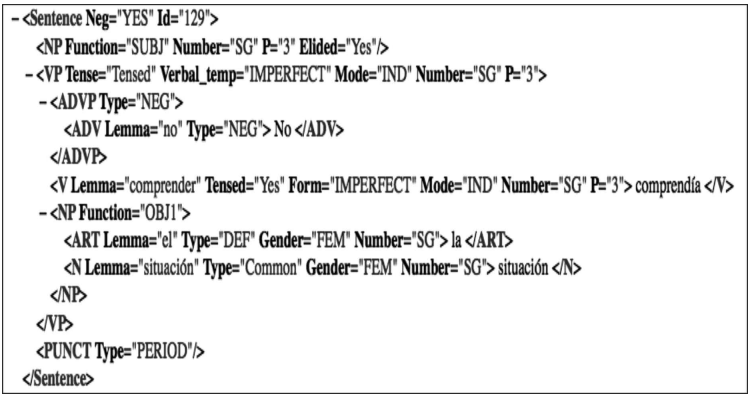


Figura 1: Ejemplo de negación oracional, *No comprendía la situación*

```
- <Sentence Neg="YES" Id="140">
    <NP Function="SUBJ" Id="1" Gender="SG" P="3" Elided="Yes"/>
  - <VP Tense="Tensed" Verbal_temp="PRES" Mode="IND" Number="SG" P="3">
    - <ADVP Type="NEG">
        <ADV Lemma="no" Type="NEG"> No </ADV>
      </ADVP>
      <V Lemma="pretender" Tensed="Yes" Form="PRES" Mode="IND" Number="SG" P="3"> pretende </V>
    - <CL Type="INFINITIVE" Function="OBJ1">
        <NP Function="SUBJ" Ref="1" Elided="Yes"/>
      - <VP Tense="Untensed" Verbal_temp="INFINITE">
          <V Verbal_temp="dar" Lemma="dar" Tensed="No" Form="INFINITE"> dar </V>
        - <NP Function="OBJ1" Neg="YES">
          - <QP>
              <Q Lemma="ninguno" Gender="FEM" Number="SG" Type="NEG"> ninguna </Q>
            </QP>
            <N Lemma="impresión" Type="Common" Gender="FEM" Number="SG"> impresión </N>
          - <ADJP Gender="FEM" Number="SG">
              <ADJ Lemma="concreto" Gender="FEM" Number="SG"> concreta </ADJ>
            </ADJP>
          </NP>
        </VP>
      </CL>
    </VP>
    <PUNCT Type="PERIOD"/>
  </Sentence>
```

Figura 2: Concordancia negativa en *No pretende dar ninguna impresión concreta*

A diferencia del trabajo de Morante et al. (2011a) y al igual que en el proyecto BioScope (Szarvas et al., 2008), el evento negado (el verbo en el caso de las oraciones o el núcleo en el caso de los sintagmas) no se marca en la actual versión de anotación, aunque se podría incorporar en un futuro. En nuestro corpus, las marcas de negación están incluidas dentro del ámbito de la negación. Todos los argumentos del evento negado se incluyen dentro del ámbito de la negación, incluyendo el sujeto en el caso de las oraciones. La negación afijal y la interoracional no se han tratado en esta fase de anotación, pero ambas están contempladas como proyecto para el futuro.

```
- <Sentence Id="1321">
  - <NP Function="SUBJ" Number="PL" P="3" Coordinated="Yes">
    - <NP Function="SUBJ" Number="SG" P="3">
        <N Lemma="Carlos_Saura" Type="PROPER"> Carlos Saura </N>
      </NP>
      <C Lemma="y" Type="COORDINATING"> y </C>
    - <NP Function="SUBJ" Number="SG" P="3">
        <N Lemma="José_Luis_Garci" Type="PROPER"> José Luis Garci </N>
      </NP>
    </NP>
  - <VP Tense="Tensed" Verbal_temp="FUT" Mode="IND" Number="PL" P="3">
      <V Lemma="competir" Tensed="Yes" Form="FUT" Mode="IND" Number="PL" P="3"> competirán </V>
    - <PP Type="POR" Class="OBL">
        <PREP Lemma="por"> por </PREP>
      - <NP>
          <ART Lemma="el" Type="DEF" Gender="FEM" Number="SG"> la </ART>
          <N Lemma="estatuilla" Type="Common" Gender="FEM" Number="SG"> estatuilla </N>
        - <PP Type="A">
            <PREP Lemma="a"> a </PREP>
          - <NP>
              <ART Lemma="el" Type="DEF" Gender="FEM" Number="SG"> la </ART>
            - <ADJP Type="COMPARATIVE" Gender="FEM" Number="SG">
                <ADJ Lemma="bueno" Type="COMPARATIVE" Gender="FEM" Number="SG"> mejor </ADJ>
              </ADJP>
              <N Lemma="película" Type="Common" Gender="FEM" Number="SG"> película </N>
            - <PP Type="DE">
                <PREP Lemma="de"> de </PREP>
              - <NP>
                  <N Lemma="habla" Type="Common" Gender="FEM" Number="SG"> habla </N>
                - <ADJP Gender="FEM" Number="SG" Neg="YES">
                  - <ADVP Type="NEG">
                      <ADV Lemma="no" Type="NEG"> no </ADV>
                    </ADVP>
                    <ADJ Lemma="inglesa" Gender="FEM" Number="SG"> inglesa </ADJ>
                  </ADJP>
                </NP>
              </PP>
            </NP>
          </PP>
        </NP>
      </PP>
    </VP>
    <PUNCT Type="PERIOD"/>
  </Sentence>
```

Imagen 3

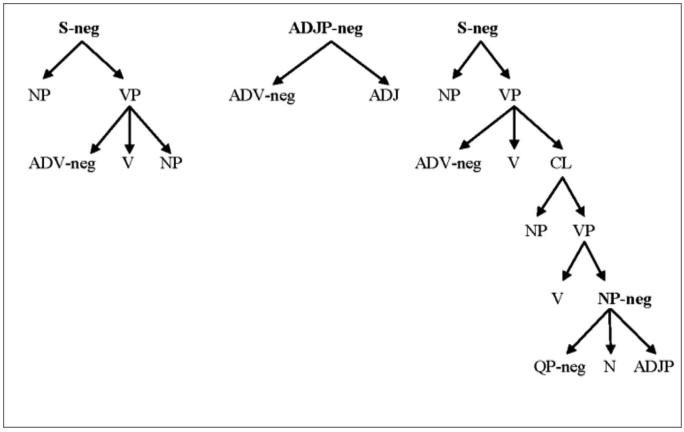


Imagen 4

5. Resultados

Los niveles de la negación anotados en el corpus nos han permitido extraer unos resultados preliminares. El Gráfico 1 muestra el porcentaje que representan las oraciones negativas en el corpus.

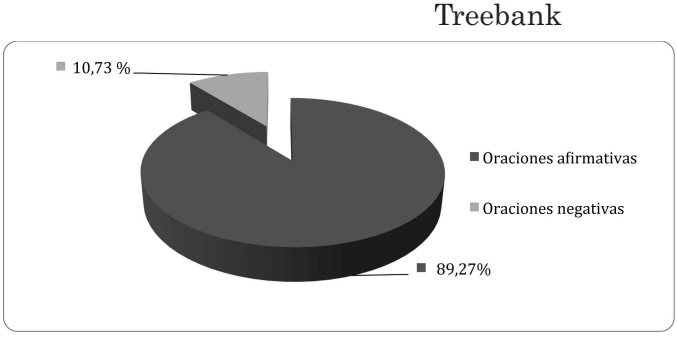


Gráfico 1. Porcentaje de oraciones negativas en el corpus

El hecho de que las oraciones negativas sólo supongan alrededor de un 10%[3] todo el corpus se explica de forma

sencilla porque la negación es un elemento marcado. "En todas las lenguas conocidas, las oraciones negativas se marcan mediante una palabra o morfema especial y las oraciones afirmativas no conocen en general un marcador obligatorio de afirmación." (Moreno Cabrera, 1991: 587). Los elementos lingüísticos marcados son menos frecuentes en las lenguas del mundo por economía lingüística. Por ello, las oraciones afirmativas, y por lo tanto no marcadas, son mucho más productivas en cualquier lengua.

Entre las marcas de negación, el adverbio negativo *no* es la marca más frecuente, un 84,63% del total de las marcas de negación (Tabla 2).

Marcas de negación

Tipo	*Frec. relativa sobre el total del UAM-Treebank*	*Frec. relativa con respecto al total de marcas de negación*
no	9,93	84,63
ni	0,27	2,27
ni_siquiera	0,13	1,14
nunca	0,27	2,27

Tabla 2. Marcas de negación encontradas en el corpus.

Los constituyentes lingüísticos afectados por el ámbito de la negación en nuestro corpus han sido principalmente oraciones, seguidas por sintagmas nominales, adjetivales y preposicionales, todos ellos con una baja frecuencia de afectación (Tabla 3).

Elementos negados	*Frec. total UAM-Treebank*	*Frec. respecto a elementos negados*
Oración	5,13	45,28
Subordinada	5,60	49,39
NP	0,27	2,35
ADJP	0,20	1,76
PP	0,13	1,18

Tabla 3. Elementos lingüísticos afectados por la negación

Las oraciones (ya sean simples o complejas) suponen el 94,70% de total de los elementos negados.

Por último, nos parece significativo el dato sobre la función de las palabras-n (Tabla 4). En nuestro corpus, *nunca* y *tampoco* funcionan en todos los casos de aparición como IN, mientras que *nada* aparece siempre como TPN en concordancia con otro IN. Finalmente, las funciones de *nadie* y *ninguno* están repartidas entre IN y TPN.

Palabras-n	*IN*	*TPN/CN*
ni	100%	
ni_siquiera	100%	
nunca	100%	
nada		100%
nadie	50%	50%
tampoco	100%	
ninguno	25%	75%

Tabla 4. Funcionamiento de las palabras-n

6. Conclusión y trabajo futuro

En este artículo hemos descrito la primera fase del trabajo realizado para anotar la negación y su ámbito. Supone una primera aproximación en español y usando un *treebank*, pues los esfuerzos anteriores se han concentrado en el inglés. Como venimos comentando a lo largo del artículo, la negación se manifiesta en el lenguaje de muy diversas formas y no todas ellas se han tenido en cuenta para este trabajo. que es una contribución desde el campo de la Lingüística de Corpus a la Lingüística descriptiva, por una parte, y a la Lingüística Computacional, por otra. Es incompleto, porque siempre estará limitado por los datos presentes. No podemos incluir toda la tipología de casos que mencionan las gramáticas, simplemente porque para la mayoría de estos no hay ejemplos en las 1501 oraciones del corpus. Nos hemos centrado en lo más básico y

frecuente, la negación sintáctica oracional y sintagmática. La morfoléxica y la discursiva se quedan muy fuera de lo que podemos abordar. Tampoco tratamos la interpretación del foco de la negación y las estructuras negativas que no tienen elementos negativos obvios. La aportación de este artículo es la cuantificación de la negación sintáctica en un corpus equilibrado de oraciones, lo que permite tener una idea aproximada del porcentaje que suponen estas estructuras en el uso lingüístico en textos periodísticos.No obstante, se han diseñado nuevas etapas del proyecto en las que se incluyen la anotación de la negación afijal, la de unidades léxicas intrínsecamente negativas y la negación interoracional. El resultado final será un corpus en el que todos los elementos lingüísticos negativos hayan sido anotados, a partir del cual se puedan extraer patrones de funcionamiento de la negación, con la finalidad de desarrollar sistemas de inferencia gramatical.

Por el momento, podemos concluir que nuestro trabajo demuestra que un corpus sintáctico facilita la identificación de la negación y, sobre todo, de su alcance, gracias a la anotación no sólo de partes del discurso, sino también de elementos superiores como los sintagmas o las oraciones subordinadas. Además, los datos estadísticos nos muestran que solamente tratando el IN "no" y las oraciones como ámbito de alcance se puede abarcar más del 80% de la negación en un texto. A pesar de que las oraciones negativas apenas supongan un 10% del total de oraciones de un texto, su identificación es fundamental para una correcta comprensión del mismo.

Agradecimientos

Esta investigación ha sido financiada por la Comunidad Autónoma de Madrid en el marco del proyecto MA2VICMR (S2009/TIC-1542) y por el MINECO en el marco del proyecto MULTIMEDICA (TIN2010-20644-C03-03).

BIBLIOGRAFÍA

ABEILLÉ, A. (2003). *Treebanks. Building and Using Parsed Corpora*. Netherlands, Kluwer Academic Publishers.

BOSQUE MUÑOZ, I. y GUTIÉRREZ-REXACH, J. (2009). *Fundamentos de sintaxis formal*. Madrid, Akal.

MORANTE, R., SCHRAUWEN, S. y DAELEMANS W. (2011a) "Annotation of negation cues and their scope. Guidelines v1.0". *Technical Report Series CTR-003, CLiPS*. University of Antwerp, Antwerp.

MORANTE, R., SCHRAUWEN, S. y DAELEMANS W. (2011b) "Corpus-based approaches to processing the scope of negation cues: an evaluation of the state of the art" in *Proc. Ninth International Conference on Computational Semantics*, Oxford.

MORANTE, R. y SPORDELER, C. (eds.) (2012) *A Special Issue of the Computational Linguistics Journal on Modality and Negation. Computational Linguistics*, 38:2.

MORENO, A. y DE MOLINA, JA. (2002). *La negación en español*. Granada, Port-Royal.

MORENO CABRERA, JC. (1999) *Curso universitario de lingüística general. Tomo I: Teoría de la gramática y sintaxis general*. Madrid, Síntesis.

MORENO SANDOVAL, A., LÓPEZ RUESGA, S., SÁNCHEZ, F. y GRISHMAN, R. (2003) "Developing a Spanish Treebank". A. Abeillé (Ed.) *Treebanks. Building and Using Parsed Corpora*. Netherlands, Kluwer Academic Publishers, pp. 149-163.

NIVRE, J. (2006) *Inductive Dependency Parsing*. Dordrecht, Kluwer.

RAE-AALE (2009). *Nueva gramática de la lengua española*. Madrid, Espasa.

SÁNCHEZ LÓPEZ, C. (1999). "La negación". En BOSQUE y DEMONTE (Eds.) *Gramática descriptiva de la lengua española*. Madrid, Espasa Calpe, pp. 2561-2634.

SANTAMARÍA, J. (2013) *Inducción gramatical no supervisada basada en patrones léxicos.* Tesis doctoral inédita. ETSI Informática, UNED.

SEGINER, Y. (2007). *Learning sintactic structure*. PhD Thesis. University of Amstersdam.

SZARVAS, G., VINCZE, V., FARKAS, R. y CSIRIK, J. (2008) "The BioScope corpus: anotation for negation, uncertainty and their scope in biomedical texts". *BioNLP 2008: Current Trends in Biomedical Natural Language Processing*. Columbus, Ohio, pp.38–45.

NOTAS

[1] Morante *et al.* 2011b presentan un breve estado de la cuestión en corpus anotados con la negación, todos ellos recientes y en inglés.

[2] Es habitual en los trabajos de anotación de corpus en los que participan varias personas proporcionar la tasa de acuerdo (*inter-annotator agreement*) para conocer el nivel de certeza o confianza en los datos anotados. En este caso, no procede hablar de ello pues ambos expertos están de acuerdo en el análisis realizado.

[3] En el *dataset* de CLEF12 el porcentaje de eventos negativos es del 5,14 para el inglés. Esto significa que nuestro corpus contiene más información sobre la negación.

ESTUDIO ETIMOLÓGICO DE CUATRO PARES DE COGNADOS EN ESPAÑOL Y URDU

María Isabel Maldonado García
University of the Punjab, Lahore

Resumen: El estudio que nos ocupa debe enmarcarse dentro de la disciplina de la lingüística aplicada y dentro de ella en el ámbito del análisis contrastivo entre el español y el urdu. En este caso se pretende determinar aquellas similitudes existentes entre cuatro pares de cognados en estas dos lenguas. Las similitudes entre lenguas relacionadas a través de su tipología como el inglés y el español han sido analizadas desde tiempo inmemorable. Lo mismo no puede decirse del par español-urdu. El hecho de que ambas lenguas pertenezcan a la familia indoeuropea y hayan recibido influencia del árabe y el persa nos conduce a la hipótesis de que en alguno de los pares de cognados debe existir una relación etimológica común.

Palabras clave: español, urdu, cognados, etimología, análisis contrastivo, similitudes lingüísticas, bilingüismo.

Abstract: The current study must be framed within the discipline of applied linguistics and within the scope of contrastive analysis between Spanish and Urdu. In this case, the purpose is to try to determine any similarities between four pairs of cognates in the two languages. The similarities between related languages trough their typology such as English and Spanish have been analyzed from time immemorial. The same cannot be said of the pair Spanish-Urdu. The fact that both languages belong to the Indo-European family and received influence from Arabic and Persian leads to the hypothesis that in any of the cognate pairs there must be a common etymological relationship.

Keywords: Spanish, Urdu, cognates, etymology, contrastive analysis, linguistic similarities, bilingualism.

1. Introducción

El propósito de este artículo es analizar etimológicamente cuatro pares de cognados en español y urdu con el propósito de averiguar si poseen el mismo origen o simplemente son cognados a través de otros aspectos lingüísticos. Los cognados

del español *aceituna, armario, bazar y toalla* además de en urdu se observan en otras lenguas.

El vocablo c*ognado* deriva del latín *cognātus* y de acuerdo con la Real Academia Española posee dos definiciones siendo una de ellas *Emparentado morfológicamente* y la otra *Pariente por cognación* donde cognación significa parentesco.

Diferentes autores han intentado definir el vocablo a través del tiempo. Johnston define el cognado como "*born together, kindred, same stock"* (1939, 4). Por supuesto esta definición lingüísticamente hablando es imprecisa, dejando demasiadas posibilidades abiertas a error. Holmes & Ramos lo definen como "*ítems of vocabulary in two languages which have the same roots and can be recognized as such*" (1993, 88). Para ellos una común etimología es crucial para el reconocimiento de los cognados. Si nos ajustamos estrictamente a esta definición, surge el interrogante de cuál sería la situación de vocablos sinónimos en dos lenguas que además fueran ortográfica y fonológicamente similares aunque etimológicamente no lo fueran.

Para Whitley, por otro lado, "a given word W from Language X and a word W from Language Y are termed "cognates" if and only if they have been inherited from the same ancestor language of X and Y". Y si su parecido es una coincidencia no son considerados verdaderos cognados. (2002, 305)

Lotto & de Groot definen los cognados como vocablos similares desde un punto de vista fonológico y/o morfológico a sus equivalentes en la segunda lengua. (1998,31-69). En este sentido, los autores no tienen en cuenta la etimología de los cognados. En la misma línea, Diane August el Al. se refieren a la percepción para identificar cognados como "*The extent to which cognate relationships are perceived is related to the degree of semantic, orthographic, and phonological overlap they share*" (2005, 52). Sunderman & Schwartz realizan una puntualización sobre estos aspectos. *"In the psycholinguistic literature, cognates are often defined as words that share aspects of spelling, sound and meaning across languages*" (2008, 527).

Lo anterior clarifica que ambos grupos de definiciones son igualmente aceptables; aquellos que requieren un origen común entre ambos términos y aquellos que basan el nivel de similitud, en la semántica, morfología, ortografía y fonología.

Las solidaridades léxicas están constituidas por relaciones semánticas entre vocablos que se ponen de manifiesto en restricciones combinables. Coseriu en su obra *Las Solidaridades Léxicas*, en un esfuerzo por aclarar el estudio sobre las solidaridades léxicas de Porzig, las define como "*determinación semántica de una palabra por medio de una clase, un archilexema o un lexema, precisamente, en el sentido de que una clase determinada, un determinado archilexema o un determinado lexema funciona como un rasgo distintivo de la palabra considerada. Dicho de otro modo, se trata del hecho de que una clase, un archilexema o un lexema pertenece a la definición semántica de esa palabra, en el plano de las diferencias semánticas mínimas"* (1977,148). Gutiérrez Ordoñez a su vez señala "*más que estructuras sintagmáticas, las solidaridades, tal como se presentan son relaciones semánticas que se establecen entre significados de signos […]. Por consiguiente, su estudio debería estar ubicado en el capítulo de las relaciones semánticas, junto con la neutralización, la antonimia, sinonimia, homonimia, etc.*" (1989, 114-115). Para nuestro estudio sobre los cognados, uno de los aspectos relevantes de las solidaridades léxicas o relaciones semánticas, (aparte de la etimología) es sin duda alguna la sinonimia.

Baldinger consideraba la sinonimia como "*uno de los platos fuertes de la semántica*" (1970, 205) y la Real Academia Española define la sinonimia como: "Del lat. *synonymĭa,* y este del gr. συνωνυμία). *1.* f. Circunstancia de ser sinónimos dos o más vocablos. *2.* f. *Ret.* Figura que consiste en usar intencionadamente voces sinónimas. Donde sinónimo posee el siguiente significado: "sinónimo, ma. (Del lat. *synonymus,* y este del gr. συνώνυμος; de σύν, con, y ὄνομα, nombre). *1.* adj. Dicho de un vocablo o de una expresión: Que tiene una misma o muy parecida significación que otro. Según López García "*la sinonimia se basa en la onomasiología, en que un significado puede ser evocado por varios significantes*" (1990, 42). La relación semántica de la sinonimia entre dos vocablos es una de las bases principales de los cognados.

Francisco Marcos Marín en su "Curso de Gramática Española" explica que "*un significado no se define solo en términos positivos, sino también en términos negativos: por lo que no es (dentro del mismo código)*" (1980, 451). El autor define la sinonimia como "*distintos significantes tienen el*

mismo significado" (1980, 452). Esta es la definición que utilizaremos para nuestro estudio de los cuatro pares de cognados del español y el urdu teniendo en cuenta que la mayoría de los autores apoyan el precepto de que la sinonimia perfecta no existe. En el ámbito de este trabajo, los cognados son en realidad, sinónimos en lenguas diferentes; distintos significantes que representan el mismo significado.

2. Metodología

La muestra sobre la que se investiga está constituida por cuatro pares de cognados del español y el urdu. Se evaluará el grado de similitud con referencia a diferentes aspectos lingüísticos. Dichos aspectos son los siguientes:

1. Aspectos etimológicos. Se extraerá la etimología de cada vocablo español y se comparará con su par en urdu con el propósito de contrastar los orígenes de ambos vocablos.

2. Aspectos relacionados con la sinonímia interlingüística.

2.1 Análisis semántico. Se compararán las definiciones con el propósito de averiguar si los cognados son sinonímicos o no.

2.2 Se compararán los pares de vocablos en su forma a través de la fonética. Esta comparación estará fundamentada en las siguientes categorías:

a. No existe diferencia en la fonética.

b. La diferencia se encuentra en uno o dos sonidos, normalmente al final.

c. La diferencia se encuentra en dos o más sonidos diferentes, algunas veces al principio.

d. La diferencia se encuentra en más de la mitad de los sonidos.

e. La diferencia se basa en que la mayoría de los sonidos son diferentes o poseen una disposición desigual.

Una vez se haya extraído toda la información, se realizara un análisis contrastivo de estos resultados.

3. Resultados

Los pares de cognados se encuentran representados en la siguiente tabla.

	COGNADOS	
PARES	*ESPAÑOL*	*URDU* *Escritura latina y persa*
Par 1	ACEITUNA	ZAYTUN (زَیتُون)
Par 2	ARMARIO	ALMARI (اَلماری)
Par 3	BAZAR	BAZAR (بازار)
Par 4	TOALLA	TOLIA (تَولِیا)

Tabla 1. Pares de cognados.

3.1 Aspectos Etimológicos

Etimología: español	*Etimología : urdu*
Significante: *Aceituna* Del ár. hisp. *azzaytúna*, este del ár. clás. *zaytūnah,* y este del arameo *zaytūnā*, dim. de *zaytā*.	Significante: زَیتُون Fonética: *['zaj.tuun]* Originalmente tiene su origen en el árabe y se utiliza como nombre. Fue utilizada por primera vez por Qutb Shah en 1611.

Tabla 2. Par 1.

Etimología: español	*Etimología : urdu*
Significante: *Armario* Del lat. *armarĭum*	Significante: اَلماری Fonética: *[al.'ma.ri]* Del prácrito y existe la posibilidad de que esta palabra pudiera tener su origen en el inglés. Fue usada por primera vez en 1847 en el libro titulado "Las maravillas de la sífilis" (Autor Yusuf Khan Kambalposh).

Tabla 3. Par 2.

Etimología: español	*Etimología: urdu*
Significante*: Bazar* Del persa *bāzār.*	Significante: بازار Fonética: *[ba.'zar]* Originalmente una palabra persa utilizada como nombre común por primera vez por Khwaja Banda Nawaz en 1421 en la obra "Skarnamh"

Tabla 4. Par 3.

Etimología: español	*Etimología : urdu*
Significante*: Toalla* Del ant. *tobaja,* y este del germ. **thwahljô.*	Significante: تَولیا Fonética: *[ˈto.l̪ia]* El autor se encuentra inseguro. Probablemente del inglés.

Tabla 5. Par 4.

3.2 Aspectos relacionados con la sinonimia interlingüística

3.2.1/2 Semántica y forma

Signo 1: español	*Significado común*	*Signo 2: urdu*
Significante: *Aceituna* Fonética*: [a.θ·e̯i̯.ˈt̪u.na]*	*1.* f. Fruto del olivo.	Significante: زَیتُون Fonética: *[ˈzai̯.t̪ʊ̯n]*

Tabla 6. Par 1.

Signo 1: español	*Significado común*	*Signo 2: urdu*
Significante: *Armario* Fonética*: [ar.ˈma.ɾ̪io]*	*1.* m. Mueble con puertas y anaqueles o perchas para guardar ropa y otros objetos.	Significante: اَلْماری Fonética: *[al.ˈma.ri]*

Tabla 7. Par 2.

Signo 1: español	*Significado común*	*Signo 2: urdu*
Significante: *Bazar* Fonética*: [ba.ˈθ·ar]*	*1.* m. En Oriente, mercado público o lugar destinado al comercio. *2.* m. Tienda en que se venden productos de varias industrias, comúnmente a precio fijo.	Significante: بازار Fonética: *[ba.ˈzar]*

Tabla 8. Par 3.

Signo 1: español	*Significado común*	*Signo 2: urdu*
Significante: Toalla Fonética*: [to.ˈa.ʎa]*	*1.* f. Pieza de felpa, algodón u otro material, por lo general rectangular, para secarse el cuerpo.	Significante: تَولیا Fonética: *[ˈto.l̪ia]*

Tabla 9. Par 4.

4. Análisis de los resultados

4. 1. Aspectos etimológicos

Tabla 2. Par 1: *Aceituna* – زَيتُون *[‘zai̯.tu̯n]*

El par de cognados comparte el mismo origen. Tanto en español, como en urdu, ambos vocablos proceden del árabe *az-zaytūna*, (الزيتونة). El DRAE anteriormente al árabe *zaytūn* “olivo” (زيتون), menciona al arameo *zaytā* (ܙܰܝܬܳܐ) como origen de la voz española, que al igual que el hebreo *zayit* (זַיִת) , significa “olivo”. Aparentemente, debido al hecho de que el aceite de oliva fue llamado *zayt* (زيت) tuvo que recurrirse a otro vocablo derivado que utilizaba el morfema final *–ōn,* diminutivo arameo que más tarde derivaría el árabe *zaytūn* (زيتون).

Corominas cita la primera documentación del árabe *al-zayt* (aceite) en el año 1251 (1987, Vol.I, 31-32). Según él, *aceituna* es el fruto del aceituno y primaba junto con oliva para designar al primero (1987, Vol.I, 32). Por otro lado Teresa Garulo no coincide con él y como primera documentación cita el *Lapidario de Alfonso X* en el año 1250 (1983, 131). En el urdu la primera referencia existente se produce, de acuerdo con el autor en 1611.

Aceituna en otras lenguas: Alemán: olive. Árabe: زيتون. Catalán: oliva. Danés: oliven. Esperanto: olivo. Francés: olive. Inglés: olive. Italiano: oliva. Finlandés: oliivi. Persa: زیتون Polaco: oliwka.

Es muy obvio que la mayoría de las lenguas que han utilizado cognados del término *oliva* derivan del latín. En árabe se sigue utilizando el mismo término que en urdu y persa constituye un préstamo léxico y que es objeto de nuestro estudio.

Tabla 3. Par 2: *Armario* – اَلْماری *[al.’ma.ri]*

Aparentemente el par de cognados difiere en su origen. Para Corominas el término *armario* es un latinismo que se origina en la primera mitad del siglo XIII cuya definición es “*lugar donde se guardan las armas*” (2011, 42). En An Etymologycal Dictionary of Latin Language Valpy define

armarĭum como *"A place where any (arma) article of dress and c., are deposited*" (1828, 33). En dicha definición *arma*, colectivo neutro plural, posee la significación de artefacto o artículo. El DRAE menciona el latín como origen del vocablo. De acuerdo con el Urdu Dictionary, *almari* deriva del prácrito y su primer uso data de 1847. Por otro lado, la conexión en urdu con el inglés podría conllevar un anterior origen en el latín ya que esta lengua se encuentra influenciada por él.

Armario en otras lenguas: Alemán: Schrank. Árabe:خزانة. Catalán: armari. Danés: skab. Esperanto: ŝranko. Finlandés: Kaappi siivouskomero. Francés: Armoire. Inglés: Cabinet, closet. Italiano: Armadio. Persa: گنجه Polaco: szafa.

Como puede verse, en aquellas lenguas derivadas del latín tales como el catalán, el francés y el italiano, la similitud es obvia, mientras que tal similitud no existe en las palabras usadas en las otras lenguas. Por esta razón debería investigarse más profundamente el origen de la palabra en urdu. Posiblemente el resultado sea el descubrimiento de un origen común.

Tabla 4. Par 3: **B***azar* – بازار *[ba.'zar]*

La palabra bazar se refiere a *waazaar*, que es una palabra del antiguo persa (1995, 391). El DRAE menciona que bazar deriva de la palabra persa *bāzār.* El autor del Urdu Dictionary, menciona la primera documentación de esta palabra en 1421. Corominas no incluye la palabra bazar en su *Diccionario crítico etimológico de la lengua castellana*, *Volumen I*, ya que el último término incluido es *barzón* e inmediatamente después *bencina*. Algunos autores mencionan que la palabra *baza*r apareció en el español en *Viaje de Turquía* que fue atribuido a Cristóbal de Villalón y con fecha aproximada de 1555. "*Dos panes, que llaman de «*bazar*», como quien dice de mercado*". Incluso entonces ya se usaba con el mismo significado. De acuerdo con el diccionario de Random House, el término apareció en el inglés aproximadamente del 1590 al 1600 aunque en el italiano antes de esa fecha en la forma de *bazarro*, procediendo este del persa *bāzār*.

Bazar en otras lenguas: Alemán: basar. Árabe: بازار . Catalán: basar. Danés: basar. Esperanto: bazaro. Finlandés: basaari. Francés: bazar. Inglés: bazaar. Italiano: bazar. Persa: بازار . Polaco: Szafa.

Como puede verse la similitud es muy evidente todas las lenguas mencionadas excepto el polaco. El par de cognados comparte el mismo origen. Ambos derivan del persa.

Tabla 5. Par 4: *Toalla* – تَولِیا *[ˈto.lia]*

El par de cognados aparentemente presenta orígenes diferentes. De acuerdo con el DRAE en español, el término proviene del anterior *tobaja* y este del germánico **thwahljô*. El autor de la etimología del urdu se encuentra inseguro sobre el origen de este vocablo y cita al inglés como probabilidad. El inglés es una lengua germánica por lo que podría existir una conexión en el origen.

Según Corominas : *"TOALLA, antiguamente tovaja, del germ. thwahljo íd. (a. alem. ant. dwahila, alem. dial. zwehle, b. alem. ant. thwahila, ags. thwehlae); la forma del castellano moderno hubo de tomarse de otro romance (italiano o catalán). 1ª doc.; toaja, Alex.; tovaja, APal.; toalla, 1570. C. de las Casas".*

De hecho, nos referimos a una antigua palabra germánica, *thwahljo* que significa 'baño', el gótico *twahl* 'baño' y en el antiguo escandinavo *thvâl* 'jabón'. Llegada al español probablemente del catalán (*tovalla),* del portugués, *toalha* o del italiano *tovaglia.* Existen textos del castellano antiguo en que los usos observados son *toaja, tobaja* y *toballa.*

Toalla en otras lenguas: Alemán: handtuch. Árabe: منشفة Catalán: tovalla. Danés: håndklæde Esperanto: mantukon. Finlandés: pyyhe. Francés: serviette. Inglés: towel. Italiano: tovaglia. Persa: حوله Polaco: ręcznik.

Como puede verse las similitudes se encuentran en el italiano, el portugués y el catalán y no en las demás lenguas. El autor del Urdu Dictionary se encuentra inseguro y menciona el inglés como probabilidad de su origen. El

inglés es una lengua germánica por lo que tal suposición podría ser correcta.

4. 2. Aspectos relacionados con la sinonímia interlingüística:

4.2.1 Análisis semántico:

Tabla 6. Par 1: *Aceituna* – زَيتُون *[‘zai̯.t̪u̯u̯n]*
Significado común: Fruto del olivo. Los dos significados restantes en el español no son comunes en el urdu.

Tabla 7. Par 2: *Armario* – الماری *[al.’ma.ɾi]*
Significado común: Mueble con puertas y anaqueles o perchas para guardar ropa y otros objetos.

Tabla 8. Par 3: *Bazar* – بازار *[ba.’zaɾ]*
Significado común: 1. En Oriente, mercado público o lugar destinado al comercio. *2.* Tienda en que se venden productos de varias industrias, comúnmente a precio fijo.

Tabla 9. Par 4: *Toalla* – تَولِيا *[‘t̪o.l̪i̯a]*
Significado común: Pieza de felpa, algodón u otro material, por lo general rectangular, para secarse el cuerpo.
En el ámbito de la significación los pares presentan significados comunes en ciertas definiciones mientras que no comparten otros de los significados.

4.2.2 Análisis de la forma:

La forma escrita en el español y el urdu, es sin lugar a dudas completamente diferente. El español utiliza la escritura latina mientras que el urdu utiliza la escritura arábigo-persa, estilo Nastaliq. Es por esta razón que ortográficamente los pares de cognados no comparten ningún rasgo. Analizaremos la forma a través de la fonética del español y del urdu.

Par 1: *Aceituna [a.θ+e̯i̯.’t̪u.na]* – زَيتُون *[‘zei̯.t̪u̯n]*
Nivel de similitud fonética: 5/8 *-eitun-*
a-θ-e-i-t-u-n-*a* ≠ *z*-e-i-t-u-n

La diferencia en la forma es más de dos sonidos diferentes, dos al principio y uno al final. Las letras *-eitun* constituyen el nivel de similitud que es de un *62.5%.*

Par 2: *Armario [ar.'ma.ɾio]* – الْمارى *[al.'ma.ɾi]*
Nivel de similitud fonética: 5/7 *a- mari-*
a-*r*-m-a-r-i-*o* ≠ a-*l*-m-a-r-i
La diferencia en la forma es obvia en el segundo sonido con grafía *r/l* y en la ausencia de la *o* final en urdu. Se presenta en dos sonidos diferentes al principio y al final.
Los sonidos comunes *[a][m][a][r][i]* constituyen el nivel de similitud que es de un *71.4%.*

Par 3: *Bazar [ba.'θar]* – بازار *[ba.'zaɾ]*
Nivel de similitud fonética: 3/4 *ba-ar*
b-a-*θ*-a-r ≠ b-a-*z*-a-ɾ
La única diferencia se encuentra en *[θ]* versus *[z].* La diferencia en la forma se presenta en un sonido en el centro de la palabra.
Los sonidos comunes *[b][a][a][r]* constituyen el nivel de similitud que es de un *75%.*

Par 4: *Toalla [t̪o.'a.ʎa]* – تَولِيا *['to.l̪ia]*
Nivel de similitud fonética: 3/5 *to-a*
t̪-o-*a*-*ʎ*-a ≠ t̪-o-*l*-*i*-a
La diferencia se presenta en dos o más sonidos diferentes en el centro de la palabra. Los sonidos *[t] [o] [a]* constituyen el nivel de similitud que es de un *60%.*
El nivel de similitud más alto lo presenta el par de cognados bazar *[ba.'θar]* – بازار *[ba.'zar]* con sólo un sonido de diferencia *[θ]* versus *[z]* y un 75% de similitud.

5. Conclusión

Con respecto a los aspectos etimológicos referentes a los cuatro pares de cognados el par *aceituna* – زَيتُون *['zai̯.t̪un]* comparte el mismo origen y la definición de fruto del olivo, mientras que no comparte las otras dos definiciones presentes en el DRAE. La fonética presenta un 62.5% de similitud con cinco sonidos comunes de ocho. El par *armario* –

آلْمارى *[al.'ma.ri]* presenta un origen latino de la palabra española mientras que en el urdu el origen es incierto, citando al prácrito como lengua vehículo de llegada al urdu y suponiendo al inglés como origen probable de la palabra. El nivel de similitud fonética es de cinco sonidos de siete y del 71.4%. Por otro lado, el par *bazar* – بازار *[ba.'zar]* comparte el mismo origen persa y es el par de cognados con mayor nivel de similitud fonética del 75% con un sonido que varía solamente. Por último, el par *toalla* – تَولِيا *['to.lia]* aparentemente presenta orígenes diferentes. El vocablo español deriva del germánico a través del anterior mientras que en urdu la etimología es incierta citando el autor al inglés como posible étimo. El par presenta un nivel de similitud del 60% con tres de cinco sonidos comunes, constituyendo este par el de menor similitud fonética.

Una investigación más exhaustiva del étimo de los vocablos آلْمارى *[al.'ma.ri]* y تَولِيا *['to.lia]* se hace necesaria factiblemente revelando un origen común de los pares en ambas lenguas. En este caso el mayor problema relacionado con este tema lo constituye la escasez de fuentes relacionadas con la lingüística en el urdu.

BIBLIOGRAFÍA

AUGUST et Al. (2005) *Development of Literacy In Spanish-Speaking English-Language Learners: Findings From Longitudinal Study of Elementary School Children*. Perspectives, 31, (2).

BALDINGER, K. (1970) *Teoría semántica. Hacia una semántica moderna*. Madrid. Alcalá.

COROMINAS, J. y PASCUAL, J. A. (1987) *Diccionario Crítico Etimológico Castellano E Hispánico*, Madrid, 1987, 2ª Reimpresión,6 Vols.

COROMINAS, Joan. (2011) *Breve Diccionario Etimológico de la Lengua Castellana*. Ed. Gredos. ISBN 9788424920401.

COSERIU, Eugenio (1977). *Principios de semántica estructural*. Madrid: Gredos

Diccionario de la Lengua Española. (2001) Real Academia Española. 22 Edición. Madrid http://www.rae.es

Dictionary.com Unabridge (2013) Random House Inc. http://dictionary.reference.com/browse/bazaar

GARULO MUÑOZ, Teresa (1983) *Los arabismos en el léxico andaluz* (Según los datos del Atlas lingüístico y Etnográfico de Andalucía), *Colección de Estudios Cordobeses*, Diputación de Córdoba, Instituto Hispano-Árabe de Cultura.

GUTIÉRREZ ORDÓÑEZ, S. (1989), *Introducción a la semántica funcional*, Madrid, Síntesis.

HARPER, Douglas (2013) Online Etymology Dictionary. http://dictionary.reference.com/browse/bazaar

HOLMES, J. & RAMOS, R. (1993) *False friends and reckless guessers: Observing cognate recognition strategies.* T. Huckin, M. Haynes, & J. Coady. *Second Language Reading and Vocabulary Learning*. Norwood, NJ: Ablex.

JOHNSTON, M.C. (1939). *Cognate Relationships Between English and Spanish Vocabularies as a Basis for Instruction*. Austin, TX: Unpublished Doctoral Thesis.

LOTTO, L. and A. DEGROOT. (1998) *Effects of learning method and word type on acquiring vocabulary in an unfamiliar language" Language Learning* 48 (1); 31-69.

LÓPEZ GARCÍA, A. (1990). *Sinonimia intralingüística y sinonimia interlingüística*, en *Actas del Primer Coloquio Internacional de Traductología*. Valencia.

MARCOS MARÍN, F. (1980) *Curso de Gramática Española*. Madrid (Cincel-Kapelusz).

PORUSHANI, I. (1995) Bazaar. Encyclopedia of the Islamic World.

SUNDERMAN, G., & SCHWARTZ, A. (2008*) Using Cognates to Investigate Cross-Language Competition In Second Language Processing*. Tesol Quarterly, 47.

Urdu Dictionary. (2011) *Urdu Encyclopedia*. Ministry of Science and Technology. Islamabad. Pakistan. http://www.urduencyclopedia.org/urdudictionary/

VALPY, F.E.J. (1828) *An Etymological Dictionary of the Latin Language*. Trinity College. Printed by A.J. Valpy. Cambridge.

VILLALÓN, Cristóbal. (1555) *Viaje de Turquia.* Calpe. Madrid. Edición de 1919.

WHITLEY, M.S. (2002). *Spanish/English contrasts: A course in Spanish Linguistics*. Second Edition. Washington, DC. Georgetown University Press.

ESPAÑOL Y PORTUGUÉS EN CONTRASTE: FALSOS AMIGOS, HISTORIA DE LA LENGUA, CAMPO LÉXICO, SEMÁNTICA DE PROTOTIPOS

José Alberto Miranda Poza
Universidade Federal de Pernambuco

Resumen: Este trabajo trata sobre una vieja cuestión en la interface portugués-español: el problema de los *falsos cognados*. En efecto, en los estudios más celebrados publicados hasta ahora –entre otros: Lado (1973), Leiva (1994), Bugueño Miranda (1998) o Durão (2002; 2004; 2005)– casi siempre este tema se enfocaba desde la perspectiva de la lingüística aplicada, en sus vertientes comunicativa e interaccionista, o sea, los autores se esforzaban en explicar cuáles eran las consecuencias, en el habla, de la existencia de este fenómeno, especialmente incidente en la *interlengua* de los hablantes de los respectivos idiomas enfrentados. Con todo, en nuestro trabajo destacamos la importancia que adquiere para el profesor de lengua extranjera una comprensión totalizadora del problema, lo que nos conduce a un análisis de la historia de las respectivas lenguas en busca de la etimología de los vocablos objeto de estudio. Solo de esta forma es como llegaremos al porqué del fenómeno al no contentarnos con la mera descripción del hecho. El análisis histórico participa de dos caras que explican el curso y la suerte de cada vocablo en las respectivas lenguas: la forma y la semántica. A partir de estas premisa se lleva a cabo a lo largo del artículo el estudio de los pares *suceso* (esp.) / *sucesso* (port.), *quesito* (esp.) / *quesito* (port.), *distinto* (esp.) / *distinto* (port.), *rato* (esp.) / *rato* (port.), *bodega* (esp.) / *bodega* (port.), lo que, a su vez, nos llevará hasta otros vocablos y no solo en las dos lenguas de referencia, sino también en otras lenguas románicas. Más adelante, cuando abordamos el par *vaso* (esp.) / *vaso* (port.) proponemos un nuevo enfoque para afrontar este tipo de estudio, recuperando una de las contribuciones de la semántica estructural: el concepto de campo léxico. De este modo, para comprobar de forma panorámica la dimensión del problema de los falsos cognados, debemos relacionar en cada una de las lenguas enfrentadas sus respectivos sistemas léxicos –Coseriu (1991), Salvador (1985), Justo Gil (1990)– estableciendo los rasgos mínimos de significado (*semas*), lo que nos lleva a la conclusión de que desde el punto de vista del significado los vocablos son cuasi sinónimos y que es solo

el uso, el contexto en el cual se especializaron en cada lengua, el que provoca la diferencia de sentido, a veces de forma sutil, por más que no siempre se perciba de esta forma por parte del hablante. Por último, proponemos la validez del análisis del significado de los lexemas en semas basándonos en las contribuciones de la semántica de prototipos, en especial, las oriundas de los trabajos de Labov (1973), Rosch (1975) y Alonso-Cortés (1993), quienes defienden que desde la perspectiva cognitiva el hablante tiende a identificar objetos con relación a las palabras que los designan según rasgos semánticos relacionados con la realidad, es decir, conforme características específicas del objeto en cuestión.

Palabras clave: falsos cognados; historia de la lengua; campo léxico; semántica de prototipos; interfaces portugués-español.

Abstract: On this paper it is analyzed an old matter on Spanish-Portuguese interface: the issue on false cognates. In fact, in the most renowned studies that have been published so far - amongst others: Lado (1973), Leiva (1994), Bugueño Miranda (1998), Durão (2002; 2004; 2005) – this subject was nearly always approached from the perspective of applied linguistics, in its communicative and interactional approaches, namely, the authors endeavored to explain what consequences, in speech, for this phenomenon existence, which is especially incident in the *interlingua* of speakers of the respective faced languages. Nevertheless, in our study we emphasize the importance that it acquires for foreign language teachers as a totalizing understanding of the problem which leads us to the analysis of the history of their languages in the search for the etymology of the words which are objects of study. It is only in this way that we will get to why the phenomenon by not settling for the mere description of this fact. Historical analysis participates in two phases explaining the course and the fate of every word in the language: the formal one and the semantic one. This is how it is analyzed throughout this article the pairs *suceso* (spa) / *sucesso* (por), *quesito* (spa) / *quesito* (por), *distinto* (spa) / *distinto* (por), *rato* (spa) / *rato* (por), *bodega* (spa) / *bodega* (por), which, in turn, will lead to other words not only in the two languages of reference, but also in other Romance languages. Then, when we approached the pair *vaso* (spa) / *vaso* (por), we propose a new bias to face this type of study, returning to a contribution from structural semantics: the concept of lexical field. Indeed, in order to prove so as an overview the extent of the issue on false cognates, we must relate in each of the languages their respective systems lexical in contrast – Coseriu (1991), Salvador (1985), Justo Gil (1990) – by establishing minimal features of meaning (*semes*), which leads us to the conclusion that from the point of view of the meaning, the words are almost synonyms and

it is only the use, the context in which they are specialized in each language which causes the difference in meaning, sometimes in a subtle way, although not always perceived in that way by the speaker. Finally, we propose the validity of the analysis of the meaning of the lexemes on semes based on the contributions of prototype semantics, especially those arising from the works of Labov (1973), Rosch (1975) and Alonso-Cortés (1993), who claim that from the cognitive perspective the speaker tends to identify objects concerning the words that designate them according to semantic features related to reality - that is, specific features of the concerned object.

Keywords: false cognates; history of language; lexical field; prototype semantics; Portuguese-Spanish interfaces.

1. Introducción

No son pocas las informaciones que, oriundas de múltiples y celebrados trabajos, se documentan a propósito del problema que ahora traemos a colación en el presente artículo: los llamados *heterosemánticos*, *falsos cognados* o, como calco de la expresión francesa "faux amis" –acuñada hace ya un tiempo por Koessler & Derocquigny (1928) con referencia a las lenguas francesa e inglesa–, *falsos amigos*, entre dos lenguas formalmente tan próximas y genéticamente emparentadas como lo son español y portugués. Baste con mencionar aquí y ahora, sólo a modo de ejemplo, las aportaciones de Beato & Durão (2001), Bugueño Miranda (1998), Camorlinga Alcaraz (2005), Durão (2002; 2004; 2005), Feijóo Hoyos & Hoyos Andrade (1992), Lado (1973), Leiva (1994), Sabino (2006), Schmitz (2001) o Werner (1997; 2006), que, directa o indirectamente, abordan desde perspectivas y objetivos diversos este mismo asunto.

Pues bien, en uno de esos primeros trabajos a los que hacíamos alusión y que inauguraban la saga en el ámbito de los estudios hispánicos en Brasil (Bugueño Miranda, 1998) se incidía en un aspecto que vamos a destacar a continuación como esencial para nuestro propósito y que, tal vez, no despertó el interés de la crítica como, a nuestro entender, hubiera merecido –nos referimos a la perspectiva histórica–, a pesar de que el objetivo del citado investigador se centrase más en destacar las diferencias de uso, consecuencia del paso

del tiempo, y no –como es nuestra pretensión demostrar–, en explicar cómo el origen común de ambas lenguas nos llevaría a preguntarnos si, en efecto, esas supuestas semejanzas que se dan entre palabras próximas lo son realmente, es decir, si el *significado* de las palabras en las dos lenguas viene a ser el mismo, por más que su *sentido*, por motivos de índole diversa, ya sean achacables a fenómenos onomasiológicos, ya a semasiológicos, haya provocado designaciones particulares y distintas en el uso, esto es, en la aplicación específica que a cada término se le ha venido a asignar en las respectivas lenguas en un determinado corte sincrónico.

Lo mismo podemos decir de otro extenso trabajo (Leiva, 1994) –una tesis de maestría, léase en portugués *dissertação de mestrado*–, que en sus primeras páginas se detiene también en consideraciones históricas con relación al problema que abordamos, si bien como marco generalizador, para después apartarse de su propósito inicial y centrarse en justificar el concepto de transferencia en el aprendizaje de lenguas extranjeras como factor desencadenante de la *interlengua*, en especial, cuando se analizan lenguas como el español y el portugués, llevando así su estudio al ámbito del habla y de la lingüística aplicada. A pesar de ello, establece una interesante clasificación de los falsos cognados en cuatro tipologías, lo que puede servir de guía, pero no es menos cierto que se haría necesario establecer algunas precisiones al respecto de algunos de los ejemplos que ilustraban cada una de ellas, en especial en aspectos tocantes a la determinación del étimo de algunas de las palabras.

Lo cierto es que, para nosotros, y aún más si nos paramos a considerar la formación que debe suponérsele al profesor de lenguas extranjeras, de modo muy señalado cuando en la interacción profesor-alumno, en la clase o fuera de ella, flotan en el ambiente, entremezcladas, lenguas como el portugués y el español, explicaciones y análisis de índole etimológica se hacen inexcusables para la consecución del éxito en el proceso de enseñanza-aprendizaje de la lengua meta, y también a la hora de entender, comprender y asimilar en su justa medida la verdadera dimensión del problema en cuestión. Queremos decir que, más allá de la obviedad –sentidos diferentes

en dos lenguas enfrentadas para palabras formalmente próximas o semejantes–, del equívoco o del error de interpretación / comprensión o, en ocasiones, como bien saben aprovechar en su propaganda las empresas dedicadas a comercializar cursillos de idiomas, de la situación humorística o ridícula que provoca el equívoco mencionado, existen otras formas de abordar este asunto desde perspectivas distintas, cosa que pocas veces o casi nunca se ha hecho, y, cuando lo fue, podemos decir que sólo de pasada o sin el debido rigor.

En este sentido, pretendemos demostrar que, por el hecho de que portugués y español son lenguas románicas, etimológicamente (genéticamente) provenientes del latín, muchas veces el significado (etimológico) de la palabra original de la que provienen los pares enfrentados realmente se mantiene en las dos lenguas, si bien, a lo largo de la historia respectiva de cada una de ellas, ha especializado su sentido, lo ha generalizado o incluso, en una de las lenguas, o en ambas, ha aparecido algún otro término que ha hecho variar el sistema, siquiera de forma leve.

Estas consideraciones nos conducen, por un lado, a la necesidad de retomar algunos de los presupuestos que estableció Michel Bréal ([1897]/1982) –precursor de la *semántica tradicional* que, en el siglo XX, habría de encontrar su cima en los trabajos de Ullmann (1965; 1986)– cuando creyó encontrar "dos leyes o principios que, a su juicio, le permitirían clasificar los cambios semánticos que se dan en el ámbito de la palabra, analizar las causas que los producen y explicar las razones por las que aparecen o desaparecen las palabras: *ley de especialización* y *ley de repartición*" (Salvador, 1985: 52).

Por otro lado, la aparición o desaparición de las palabras a lo largo de la historia de una lengua no se agota en el propio hecho. Por ello, encontraremos en la aplicación del concepto de *campo léxico* al problema de los falsos cognados una forma de explicación idónea para algunos de los ejemplos que nos disponemos a analizar. En virtud de tal hipótesis, el léxico de cada lengua enfrentada está organizado en una serie de estructuras que reúnen en torno a ellas un grupo de lexemas

emparentados por algún o algunos rasgos de significado común. Cada una de esas estructuras –de carácter paradigmático– recibe el nombre de *campo léxico* (Justo Gil, 1990: 32) y la disciplina lingüística que se encarga de su estudio, *lexemática* (*semántica léxica*).

Por último, con relación a los *rasgos de significado* que caracterizan semánticamente cada *lexema* (palabra), nos serán de utilidad algunas de las observaciones que, desde la psicología cognitiva, desarrolló la *semántica de prototipos*, que trabaja en el ámbito de la percepción cognitiva del hablante con relación a los rasgos de significado que necesariamente conforman el concepto que representa la palabra, y que tuvo entre sus primeros iniciadores a Koffka (1926), con su *ley de constancia de la figura*, cuestión que fue posteriormente retomada y puesta en práctica con el experimento de las figuras de Labov (1973) –que llegó a cuestionar parcialmente la discreción de los rasgos distintivos de significado en la percepción real de los objetos mostrados a los informantes–, lo que llevó a Rosch (1975) a proponer que en toda categorización / conceptualización hay un núcleo o *prototipo fijo* y unas propiedades o rasgos graduales, lo que le lleva a afirmar que la categorización es en ciertos casos borrosa (Alonso-Cortés, 1993).

Una vez expuestos los fundamentos teóricos que presiden nuestro trabajo, así como el objetivo de incorporar, como necesaria, la perspectiva histórica al estudio de los *falsos amigos*, aplicaremos una metodología en virtud de la cual analizaremos grupos de palabras que, más allá, del par enfrentado, se encuentran afectadas directa o indirectamente por la situación de conflicto, ya sea por la forma (*onomasiología*), ya por el significado (*semasiología*).

2. *Suceso* (español) / suceso (portugués)

Tanto la palabra española *suceso* como la portuguesa *sucesso* provienen del sustantivo latino *successus, -ūs, ´*secuencia, sucesión`, ´éxito` (Corominas y Pascual, 1983, II: 12-13), derivado, a su vez, del participio pasivo *successus, a, um*, del verbo *succēdo, is, ĕre, cessum*. Este último verbo es también palabra

derivada, proveniente de otro verbo, *cēdĕre*, ´retirarse, marcharse`, ´no resistir` (Corominas y Pascual, 1983, V: 324). Por más que, al hacer referencia al vocablo portugués, Machado (1995, V: 240) ofrezca un sustancioso ejemplo del siglo XVI: "Em todas as obras humanas há bom e mau sucesso, culpa e louvor, glória e pena...", donde parece claro el sentido neutro que encerraba la palabra, esto es, ni negativo, ni positivo, Blánquez Fraile (1978: 478) aboga por la teoría de que ya en el propio latín la palabra *successus, -ūs* poseía un doble valor: a) ´aproximación, venida, llegada`; b) evento próspero, favorable`. Y, de hecho, el propio Machado, a pesar del ejemplo aducido, así parece confirmarlo cuando, un poco antes, define *sucesso* como "acto de penetrar no interior de [...], aproximação, chegada, marcha para frente, resultado, triunfo" (Machado, 1995, V: 240).

Ocurre, por otro lado, que tanto en portugués como en español se documenta la palabra *éxito* / *êxito* y, en ambas, con el mismo significado: ´resultado feliz de un negocio, actuación, etc.`; ´buena aceptación que tiene alguien o algo` (DRAE, 2010: 690). Y el problema está en que en portugués conviven las formas *sucesso* y *êxito* con el mismo significado, siendo, al menos en Brasil, más común en el uso la primera. El conflicto entre las lenguas próximas se produce cuando en español *suceso* significa ´cosa que sucede, especialmente cuando es de alguna importancia` y además –sólo como cuarta y quita acepción respectivamente–, ´hecho delictivo`, ´accidente desgraciado` (DRAE, 2001: 1428). Considerando estas dos últimas acepciones del español *suceso*, por oposición a la acepción más actual de *sucesso* en portugués, cabe hablar del fenómeno de *falsos amigos*. Pero, no olvidemos la acepción más genérica del término, que es la primera que aparece en el DRAE, ni perdamos de vista el ejemplo del portugués del siglo XVI que ofrecía Machado. Y, para hacer aún más curiosa la cuestión, el propio diccionario académico nos puede dar alguna sorpresa cuando leemos la tercera acepción que atribuye al término español: ´éxito, resultado, término de un negocio`, con lo que los *falsos amigos* no serían tan falsos como podría parecer en un principio.

Porque, en efecto, estos últimos datos que revelamos nos invitan a considerar que, en realidad, el poder de la etimo-

logía es mayor de lo que en principio pudiera parecer y, desde luego, mucho mayor que aquellos que rechazan, por innecesaria, la perspectiva histórica en los estudios lingüísticos actuales. Así, en español –y ahí está el diccionario para atestiguarlo–, *suceso* también puede aplicarse, como en portugués, a un hecho positivo o feliz, si bien, no es lo más usual hoy en día, porque históricamente para ese significado de prosperidad se escogió, o mejor, se especializó otra palabra, *éxito*, que posee una particular historia.

En efecto, la palabra *éxito* (también la forma portuguesa *êxito*) proviene del verbo latino *exīre*, derivado, a su vez, de *īre* ´ir`, de donde, *exīre,* que significaba literalmente ´ir hacia fuera, ir hacia delante` (Corominas y Pascual, 1983, II: 824). De este verbo *exīre*, se formó el verbo castellano antiguo *exir*, que aún existe en valenciano actualmente, así como sus derivados: *exida* (cast.ant.; val.) ´salida` y, por supuesto, *éxito*. También la palabra *éxito* poseía un sentido genérico, ni positivo, ni negativo aún en el siglo XVIII, según documentan Corominas y Pascual (1983, III: 462-3): "Moratín habla de un éxito infeliz", en el sentido de ´salida o resolución de algo´, aunque ya en el *Diccionario de Autoridades,* en época contemporánea a la del citado autor, se dice de una empresa ardua que "no tendrá éxito" (con todo, entiéndase aquí como ´salida´, en sentido abstracto; y sólo a partir de él, ´perspectivas o resultados favorables`). Precisamente este es el significado que recoge como acepción del término portugués *êxito* (Machado, 1995, II: 514): ´acto de sair, saída; morte, fim; resultado; fim, termo`.

Pero, aún hay más. Porque, al definir *éxito* / *êxito*, han surgido las palabras *salida* / *saída*, que poseen su propia etimología. Estas últimas derivan del verbo latino *sălĭo, is, īre, sălŭi, saltum,* cuyo significado era ´saltar, brincar, dar saltos` (Blánquez Fraile, 1978: 437). Ya desde la época de orígenes, en el *Mío Cid*, se documenta el verbo *salir* con el sentido de ´pasar de dentro afuera`, de donde su derivado *salida*, que se generaliza en castellano, dando lugar a la posterior desaparición en la lengua española de uno de los lexemas en conflicto, *exida* (Corominas y Pascual, 1983, V: 139-40). Por su parte, y con relación al portugués, Machado (1995, V: 142) documenta la parcial desaparición de *exīre* cuando redacta el

artículo de la palabra *sair:* "do lat. *salīre*, ´saltar, pular, vibrar [...], sair`, *que, portanto, se encarregou de substituir parcialmente exīre*"(subrayado nuestro).

Ahora bien, hemos mencionado, a propósito de la definición del latín *sălĭo*, el verbo español *brincar,* con su sentido actual de ´dar saltos, dar brincos`. Pero, *brinco*, en portugués, no es un ´salto`. Es más, se da el caso de que el español *brincar* proviene del portugués *brincar*, cuyo significado era –y continúa siendo– ´jugar, retozar`. El vocablo portugués es un derivado del sustantivo *brinco*, ´anillo, sortija, juguete para los niños`, procedente, a su vez, del latín *vincŭlum*, *ī* ´atadura, lazo, ligadura` (Blánquez Fraile, 1978: 537).[1] Relacionada con esta última etimología, se encuentra la palabra española *vinco* que el DRAE (2001: 1564) documenta como leonesismo –téngase en cuenta la vecindad con Portugal–, con dos acepciones: una en singular, ´anillo de alambre que se pone en el hocico a los cerdos para evitar que hocen`; y otra en plural, ´pendientes que usan las mujeres formados por un aro de plata`. La forma española *brinco* con el sentido de ´saltar como los niños` puede deberse, según Corominas y Pascual (1983, I: 655-66) a que el término se aplicó, a partir de 1589, a un aro u otro objeto de forma anular, como los que rematan los chupetes de los niños de teta actualmente. De forma que en Cervantes, Lope de Vega y autores contemporáneos, ya se documenta *brinco* derivado de *brincar* como ´salto`.

Entonces, si *brinco* (<*brincar*) es *salto* (<*saltar*), a su vez, va perdiendo paulatinamente su sentido de ´joya`; por tanto, el español ha de buscar otra denominación, otra palabra, para significar ese valor desviado o metafórico que adquirió la palabra original que lo poseía: *pendiente*, del latín *pendens, tis,* participio de presente del verbo *pendĕo, es, ēre, pĕpendi* latino, que explica el origen del castellano *pender* <*pĕndēre*, ´suspendido, estar colgado de` (Blánquez Fraile, 1978: 348). En español, la forma *pendiente*, presenta dos valores: uno, como sustantivo, equivaldría al *brinco* del portugués, e incluso a una de las acepciones del leonesismo *vinco*; la otra, como adjetivo, presenta un significado derivado de su sentido original ´no resuelto`, pero, con este valor, entra en conflicto con *pendencia*, palabra formalmente próxima pero no directamente relacionada con ella desde el punto

de vista etimológico, por más que compartiesen significados fronterizos, de manera que, una vez establecido el conflicto, *pendencia* especializó su uso, en sentido negativo, primero manteniendo un cierto vínculo con el significado etimológico ´situación apurada, alarma`, para definitivamente desembocar en el sentido actual ´riña, pelea` (Corominas y Pascual, 1983, IV: 475-6). Porque, en efecto, *pendencia* no proviene de *pĕndēre* –étimo de *pendiente*–, sino de *paenitentia* ´pesar`, derivada de *paenitere* ´arrepentirse`, en portugués antiguo, *arrepender-se*.

Tenemos, por lo tanto, que *pendiente*, como sustantivo, sólo existe en español, con significado equivalente a *brinco* en portugués. Que el adjetivo *pendiente* (*pendente* en portugués) posee el mismo significado en ambas lenguas, derivado del original, con el sentido de ´no concluido, no acabado o no resuelto`. A partir de éste último, el español otorga al sustantivo *pendiente* el significado de ´cuesta o declive de un terreno` (DRAE, 2001: 1720). La palabra *pendencia* en español restringió su significado a ´contienda, riña` a partir del sentido etimológico derivado de *paenitentia*, ´arrepentimiento, disgusto`, mientras que el portugués especializa el significado de *pendência* como ´litígio` al lado de valores semejantes ´briga, contenda`, valores todos ellos derivados de otro que entra en conflicto con *pendente,* por más que su etimología sea dispar*:* ´tempo durante o qual uma causa ou um recurso está pendente ou correndo` (Ferreira, 1999: 1534).

¿Qué fenómenos históricos se han producido realmente? Como anticipamos más arriba, hemos de volver a Bréal ([1892]/1982) en busca de auxilio para nuestra explicación. Habíamos referido dos leyes a propósito de los trabajos de dicho autor, la *ley de especialización* de significado y la *ley de repartición* de significado. Antes, en un trabajo fundacional, el propio autor había hablado de una tercera, la *ley de especialidad*, que consiste en otorgar a una única palabra la función que con anterioridad era realizada por toda una serie de la misma especie. Por su parte, la *ley de especialización* afirma que términos que en la lengua común tienen un significado general tienden a ser empleados por cada grupo profesional, clase social, etc., con otro más específico que comporta la marca de aquello que los representa (Muñoz Núñez, 1999: 48-9). Por últi-

mo, la *ley de repartición* sostiene que cada vez que se produce sinonimia en una lengua las voces implicadas en ella tienden a repartir su contenido y acaban por diferenciarse semánticamente (Salvador, 1985: 53).

Mediante dichos conceptos vamos a conseguir explicar la compleja serie de procesos que se han llevado a cabo, en el sentido de que, lejos de lo que habitualmente se suele mostrar, la existencia de esos pares de supuestos *falsos amigos* no se da por un azar del lenguaje, ni mucho menos de forma aislada con relación a las dos lenguas enfrentadas en general o al resto de vocablos de cada una de ellas, en particular.

Pues bien, partíamos inicialmente del par *suceso* (esp.) / *sucesso* (port.), palabras que comparten un mismo étimo, el sustantivo latino *successus*, derivado del verbo *succedĕre*, cuyo significado era ´aproximación, llegada`, si bien se apunta, aún en época latina, un sentido próximo a ´favorable`. De forma paralela, reparamos en el par *éxito* (esp.) / *êxito* (port.), términos relacionados etimológicamente con *exida*, ´salida, fin, término`, documentada en castellano antiguo y en valenciano actual, derivadas todas ellas del verbo latino *exīre.* Encontramos, además, documentación que nos informa de que, en época antigua, el portugués *sucesso* podía ser *bom e mau* (Machado, 1995, V: 240) y el español *éxito*, podía ser *infeliz* (Corominas y Pascual, 1983, III: 462-3). Nos encontramos con dos palabras que poseen dos etimologías diferentes y que originalmente en latín no presentaban el mismo significado, que en su evolución histórica fueron perfilando sus matices significativos, llegando a colisionar en época romance. Entonces, aplicando las tesis de Bréal ([1892]/1982), podemos decir que, por un lado, para mantener la oposición preexistente y, al mismo tiempo, para evitar un fenómeno de sinonimia, los respectivos sistemas lingüísticos reaccionaron, no de forma impositiva, sino a través de su realización, en el uso.[2] Y ocurrió que en las respectivas culturas o comunidades que hablaban cada lengua se produjeron reacciones dispares, lo que explica la situación actual. Así, en español, una vez que se llegó a una situación de sinonimia, se produjo una *repartición de significado*, esto es, *suceso* se vinculó esencialmente al significado de ´hecho delictivo`, ´accidente desgraciado` (DRAE, 2001: 1428), por oposición a *éxito:*

´resultado feliz de un negocio, actuación, etc.`; ´buena aceptación que tiene alguien o algo` (DRAE, 2010: 690). Por el contrario, en portugués, una vez que se llegó a la situación de conflicto entre los dos términos, se mantuvo la sinonimia, pero el habla, al menos en Brasil, suele privilegiar como más usual la forma *sucesso* en detrimento de *êxito*. Ahora bien, tanto la *repartición de significado* en español como en la tendencia a privilegiar en el uso de una de las formas en portugués, se producen una vez que se ha dado una especialización en el significado adquirido. Así, se documenta tanto la posibilidad de asociar una valoración alternativa opuesta, positiva y negativa, al término *sucesso* en el portugués del siglo XVI, como al término *éxito* en el español del siglo XVIII.

Ahora bien, ¿por qué *éxito / êxito* se impregnan de nuevos matices de sentido, que los llevan a alejarse del significado originario y, por ende, a crear la situación de conflicto con *suceso / sucesso,* llegando a provocar, incluso, la desaparición del verbo del castellano antiguo *exir* <lat. *exīre*? Para explicar este fenómeno, debemos observar lo que ocurrió con el verbo que vino a sustituir semánticamente al desaparecido *exir* <lat. *exīre: salir / sair.* Tanto el verbo *salir / sair* como sus derivados *salida / saída* provienen del verbo latino *salīre* que, en principio, significaba ´saltar, brincar`, de donde se explica, con este mismo valor, el español *salto*, derivado del participio pasado *saltum,* del mismo verbo *salīre*, pero que, a su vez, en *salida* (cf. port. *saída*), comenzó a adquirir el significado de ´lugar por el que se pasa de dentro afuera`. No obstante, a pesar de este desplazamiento de significado, portugués y español mantienen aún palabras provenientes de *salīre* que responden al significado etimológico original. Así ocurre con el verbo portugués *salientar,* literalmente, ´poner en alto`, esto es, ´destacar, llamar la atención sobre`, y con los vocablos españoles *sobresalir* y su derivado *sobresaliente*, literalmente ´exceder en altura´, después usados en sentido figurado. Pues bien, el desplazamiento de significado producido en *salir* <lat. *salīre* y sus derivados en dirección a lo que originalmente significaba *exir* <lat. *exīre* y sus derivados, fue lo que condujo a que formas romances como *éxito* / *êxito* se contagiaran de nuevos valores, que, a

su vez, entraron en conflicto con los que originariamente poseían otras palabras, como *suceso / sucesso*.[3]

Una vez que *salir* ha adquirido un nuevo significado, ´dirigirse a un lugar de dentro afuera`, el significado etimológico ´dar saltos`, debe recaer en una nueva forma. En español, se ofrecen dos posibilidades: a) recuperar un verbo poco frecuente en latín clásico, alternativo de *salĭo, saltar*, proveniente de *salto, as, āre, āvi, ātum*, que realmente significaba en época clásica ´bailar, danzar con gestos y pantomimas`, es decir, realizar otro tipo específico de saltos (Blánquez Fraile, 1978: 437); b) paralelamente, acrecentar el acervo léxico a través de la apropiación –léase aquí préstamo léxico–, de la forma *brincar*, que el portugués adoptó y cuyo origen al parecer es germánico, si bien no se descarta un cruce etimológico entre *springan*, en el sentido de ´saltar`, y de *bli(n)kan* ´jugar` con el vocablo latino *uinc(u)lum*; y, entonces, en portugués, primero fue *brinco* ´juguete` y después llegó *brincar* cuyo significado, en cualquier caso, era ´jugar, retozar`, valor con el que permanece en esta lengua hasta hoy, pero que en español se traslada a la *casilla vacía* de significado que había dejado sin cubrir perfectamente el verbo *salir* cuando en castellano adoptó el nuevo sentido de *exīre*, olvidando el que poseía en latín *salīre*, entendiéndose como un tipo de juego o retozo específico consistente en ´dar saltos`, significado que se comparte con el nuevo valor adoptado para el verbo *saltar*.[4] Con todo, hoy en español, *saltar* y *brincar* no son exactamente sinónimos, sino que podemos decir, parafraseando, una vez más. los conceptos de Bréal ([1897]/1982), que han repartido su significado, especializándolo a la vez. En efecto, saltar hoy en español es ´salvar de un salto un espacio o distancia` (DRAE, 2001: 2015), mientras que *brincar* significa ´dar brincos o saltos´ (DRAE, 2001: 356). Se entiende, por tanto, que un *brinco* es un tipo de *salto*, ´súbito, leve`, pero que no todos los saltos son brincos, luego *brincar* es *saltar*, pero no se siempre que se salta se brinca.

Pero, *brinco* en portugués no significa ´salto`, sino que hace referencia a un tipo de joya o adorno que se coloca en el lóbulo de la oreja. Con ello, nos las tenemos con un nuevo par de *falsos amigos* y eso que, en este caso, parece que la historia del término nos empuja a pensar que el español lo tomó prestado

del portugués. Como en español, por las razones aducidas, tomó en préstamo el verbo *brincar* como solución parcial para cubrir el significado etimológico que poseía *salir*, una vez que había desplazado su significado, de nuevo se vio obligado a recurrir a otro término, que lógicamente, no va a compartir con el portugués, toda vez que este idioma, con relación a la palabra *brinco*, se limitó a precisar o especificar su significado original, como ya explicamos: de ´juego, juego de niños` pasó a designar un tipo de juguete en forma de aro que estos utilizaban para jugar, de donde, en virtud de su forma, se aplicó también a una argolla de tipo circular que se utilizaba como ornamento habitualmente en la oreja. Entra entonces en juego la forma *pendiente*, que en principio, igual que en portugués (*pendente*) era el participio presente del verbo *pender* <lat. *pĕndēre* y que funcionaba como adjetivo, con el sentido ´que cuelga, que pende de, que está suspenso`. De ahí, por semejanza, se metaforiza el sentido y se aplica a la joya u ornamento que efectivamente pendía de donde se fijaba, surgiendo así el sustantivo *pendiente* equivalente significativamente a *brinco* (portugués) y también, por otro lado, en este mismo sentido aquí apuntado, *pendiente* significa en español ´cuesta o declive de un terreno` (DRAE, 2001: 1720).

Por último considerando enfrentados los adjetivos *pendiente / pendente,* entra en discusión otro par: *pendencia / pendência.* La palabra *pendencia* en español restringió su significado a ´contienda, riña` a partir del sentido etimológico derivado de *paenitentia*, ´arrepentimiento, disgusto`, mientras que el portugués especializa el significado de *pendência* como ´litígio`, al lado de valores semejantes ´briga, contenda`, todos ellos derivados de otro, que entra en conflicto con *pendente*: ´não concluso`, por más que su etimología sea dispar*:* ´tempo durante o qual uma causa ou um recurso está pendente ou correndo` (Ferreira, 1999: 1534).

3. *Quesito* (español) / *quesito* (portugués)

En su *Diccionario*, María Moliner (2007, II: 2458) define *quesito* como "pequeña porción de queso cremoso, envuelta en papel de aluminio y vendida con otras iguales en una cajita de

cartón". De esta forma, se trata formalmente de un diminutivo de *queso* que, sin dejar de serlo, ha sufrido una cierta lexicalización especificando en parte su significado, en el sentido de que no a toda porción de queso puede aplicársele la designación *quesito,* sino solo a aquella que presenta ciertas características y cualidades.

Históricamente, la palabra española *queso*, como la portuguesa *queijo*, proceden del sustantivo latino *casĕus, i* (Corominas y Pascual, 1983, IV: 721; Blánquez Fraile, 1978: 87; Machado, 1995, V: 17)). De uso general en todas las épocas, el galorromance, el catalán y el italiano lo sustituyeron por *formatĭcus*, de donde las formas actuales *fromage* (fr.), *fromatge* (cat.), *formaggio* (it.). La palabra *formatĭcus* aparecía comúnmente en latín en combinación con *caseŭs: caseŭs formatĭcus,* constructo en el que *formatĭcus* hacía referencia al molde perforado en el que se hacía el queso.

Por su parte, la palabra portuguesa *quesito* proviene del latín *quaesītu* ´pergunta, questão` (Machado, 1995, V: 17), esto es, presenta una etimología completamente distinta a la palabra española, por más que ambos vocablos coincidan formalmente en la actualidad en las respectivas lenguas. Su significado específico es: "Ponto ou questão sobre que se pede resposta, opinião, juízo ou esclarecimento" (Ferreira, 1999: 1684). A su vez, la forma *quaesītus, a, um* era en principio el participio de pretérito del también latino verbo *quaero, is, ĕre, quaesīvi* o *sĭi, sĭtum* (Corominas y Pascual, 1983, IV: 717-21), cuyo significado era ´buscar, indagar` (Blánquez Fraile, 1978: 402). Así pues, no es de extrañar que tanto la palabra portuguesa *questão*, como la española *cuestión* provengan del sustantivo latino derivado de dicho verbo, *quaestĭo, onis* (Corominas y Pascual, 1983, II: 279), que si bien significaba ´busca`, también presentaba la noción de interrogatorio de los cautivos ´en busca` de una información. De aquí parte, por extensión, una de las acepciones del término portugués: "Demanda, litígio. Ponto em discussão que é levado à Justiça e submetido à decisão de um magistrado" (Ferreira, 1999: 1684), porque, además, ya en latín, en Cicerón, se halla la expresión *quaestionem instituere de morte alicuius* ´abrir un juicio para averiguar la muerte de alguien`. Lo que, a su vez, no obsta para que el término también posea acepciones próximas con el sentido

propio del étimo latino "pergunta para esclarecer ou para comprovar conhecimento; indagação; matéria a examinar; polémica; disputa" (Peixoto, 2007: 134).

Todas estas consideraciones nos obligan a comentar un cambio histórico que en español y en portugués se operó en el sentido que adquirieron los respectivos derivados del verbo *quaerĕre,* en origen, ´buscar, pedir`, pues *querer* (esp., port.) pasó a tener un sentido volitivo. En este caso, estamos ante una *restricción de significado* ´querer –léase ´buscar`– el bien de alguno`, ´amar`. Con todo, los derivados de *quaerĕre* latino desaparecieron en gran parte del resto de la Romania y como ejemplos próximos podemos aducir los casos del catalán y del francés donde, respectivamente, *voler, vouloir* mantienen este sentido, formas que en castellano, y con este mismo valor, solo se conservan en derivados, como *volitivo* o *voluntad (vontade,* en portugués). *Voluntad* (esp.), *vontade* (port.)[5] provienen del sustantivo latino *volŭntas, ātis,* (Machado, 1995, V: 406), con el significado de ´acto de voluntad, volición, facultad de querer`, palabra derivada, a su vez, del verbo *vŏlo, vis, vult, velle, vŏlŭi* ´querer, desear` (Blánquez Fraile, 1978: 541-2; Corominas y Pascual, V: 839).

Una última diferencia español-portugués, relacionada con el espectro significativo de *voluntad, vontade*: la palabra *gana,* tan presente en español en expresiones coloquiales como *porque me da la gana, tengo gana(s) de…* El origen de esta palabra es oscuro. Se trata de un vocablo propio del castellano, no común con otras lenguas románicas, pero que comparte el catalán (Corominas y Pascual, 1983, III, pp. 62-65). Su étimo remonta probablemente al gótico **ganô* (fem.) ´avidez`, emparentado con el escandinavo antiguo *gana* ´abrirse la boca, desear con avidez`, noruego *gana* ´quedarse boquiabierto`, ´mirar con ansia`. Tiene desde la Edad Media hasta hoy el mismo valor y usos parecidos en español y catalán.

Pues bien, por el hecho de que *querer* en español y portugués fueron perdiendo el sentido original de *quaerĕre* latino, impregnándose de un sentido volitivo –que otras lenguas románicas no adoptaron– y abandonando formalmente derivados de *vŏlo, velle* latino, se hacía necesario en estas dos

lenguas recubrir con otro verbo (vocablo) el significado original latino de ´buscar (saber), indagar, interrogar a los cautivos`. Y, así, surgen las formas *preguntar, pregunta* (esp.) y las correspondientes *perguntar, pergunta* (port.), que iban a competir con *cuestión*, *questão* con el significado de ´pregunta`.[6] Verbo y sustantivo derivan de formas que se desarrollaron en el latín vulgar en el momento de la reorganización de los sistemas lingüísticos de las respectivas lenguas románicas, una vez que la distancia con la lengua de referencia, el latín, era cada vez mayor: **praecunctare* < **percŏntari* ´someter a interrogatorio`; la forma vulgar está influida por *cunctari* ´dudar, vacilar` (Corominas y Pascual, 1983, IV: 635). Además de en español y en portugués, solo se conservan derivados de este vocablo en sardo: *percontare*. La forma catalana *preguntar* es castellanismo, usándose *demanar*, más próxima al francés *demander*.

En realidad, *percontari*, verbo latino, formado por el prefijo *per-* ´mediante, por medio de` y *contus* ´palo largo, percha`, significaba ´sondear, buscar en el fondo del mar`. Más tarde, por una clara metáfora, significó también ´sondear a una persona`, uso que se mantiene aún en español, es decir, someterla a un interrogatorio. La contaminación del latín vulgar con el verbo *cunctari* ´dudar, vacilar` fue reflejo, sin duda, de la actitud de quien respondía en el momento del interrogatorio (Buitrago y Torijano, 2007: 198).

Con todo, otras lenguas, como ya hemos dicho, adoptaron *demander* (fr.), *demanar* (cat.) para recubrir este significado, lo que no fue óbice para que también en español y portugués existan *demandar, demanda*, si bien con otros matices significativos diferentes de *preguntar, pregunta* (*perguntar, pergunta)*. El étimo común latino para todas las lenguas romances es el verbo *demando, as, āre, āvi, ātum*, cuyo significado era ´encomendar, confiar, poner a buen seguro` (Blánquez Fraile, 1978: 146; Machado, 1995, II: 298). Este verbo, por su parte, era derivado de *mando, as, āre, āvi, ātum*, que poseía el significado de ´encargar de, dar una misión´, ´confiar, encomendar` (Corominas y Pascual, 1983, II: 441; III: 800-1; Machado, 1995, II: 298). Como en español y en portugués las formas derivadas de este verbo latino entraron en conflicto con los nuevos vocablos aparecidos en el latín vulgar, hubieron de *repartir su sig-*

nificado, de manera que *demanda, demandar*, a través de la cuarta acepción que de la palabra portuguesa en cuestión ofrece Ferreira (1999: 619): "Reclamar, pedir, requerer, exigir: *A tarefa demanda muito tempo e competência*", se llega al sentido jurídico que, por excelencia, hoy presenta, en dichas lenguas –sexta acepción de Ferreira–: "Intentar ação judicial ou demanda contra; procesar, acionar".

Pero aún no hemos acabado con la serie de palabras emparentadas históricamente y que debemos mencionar como consecuencia de la toma en consideración de la pareja de falsos amigos *quesito* ´porción de queso de determinada cualidad` (esp.) / *quesito* ´punto o cuestión sobra la que se pide una respuesta u opinión` (port.). Porque, en efecto, relacionadas con las anteriores, no podemos olvidar otro par de términos que, por su parte, constituyen un nuevo ejemplo de falsos cognados: *exquisito* (esp.) / *esquisito* (port.). Ambas formas provienen de una misma etimología, el participio de pasado *exquīsītus, a, um* ´rebuscado, escogido, selecto, distinguido, refinado` (Corominas y Pascual, 1983, IV: 719), del verbo *exquīro, is, ĕre, sīvi, sītum,* derivado morfológicamente, a su vez, del prefijo *ex-* y del verbo *quaero*, cuyo significado era ´buscar con cuidado, indagar, escudriñar` (Blánquez Fraile, 1978: 190-1).

En este caso, aunque el significado y las respectivas formas se mantuvieron fieles a la del vocablo original, con el tiempo, español y portugués *restringieron el significado del término*, si bien en sentido opuesto.[7] Así, en español, lo ´rebuscado y escogido` se atribuye a un valor positivo, esto es, "de singular y extraordinaria calidad, primor o gusto en su especie" (DRAE, 2001: 693); por su parte, en portugués lo ´especial o escogido` se entiende en sentido negativo: "excêntrico, estranho, extravagante; de mau aspecto; feio e/ou malvisto; lugar deserto, ermo; caminho difícil e escabroso" (Ferreira, 1999: 823).

4. *DISTINTO* (ESPAÑOL) / *DISTINTO* (PORTUGUÉS)

La forma *distinto* en español y portugués posee una misma etimología: se trata del participio del verbo latino *distinguo, is, ĕre, stinxi, stinctum* ´separar, dividir`, verbo utilizado más

tarde en la lógica medieval para establecer distinciones (Corominas y Pascual, 1983, II: 506), vocablo compuesto, a su vez, por el prefijo *dis-* y el verbo *stinguo*, que significaba ´hacer desaparecer`, ´analizar, pesquisar, escudriñar` (Blánquez Fraile, 1978: 161). Entonces, lo que se separa, se divide, se corta, conlleva una diferenciación de cada una de las partes separadas de un todo (Machado, 1995, II: 347), que es el sentido que hoy posee el verbo *distinguir* y, consecuentemente, su correspondiente participio (adjetivo) *distinto*, como se define en Ferreira (1999: 695): "Separado, isolado", lo que no es igual en el sentido en que ya no forma parte de la unidad. Ahora bien, en portugués, aún restringe más su significado, siendo este el más común en el uso: "Que sobresai; notável, ilustre, eminente, preeminente: *orador distinto*" (Ferreira, 1999: 695). Es justo en este sentido, como decimos, el más usual del portugués para este vocablo, en el que entra en conflicto semántico con *distinto* español, lengua en la que para referirse a lo sobresaliente o ilustre se utiliza *distinguido*, participio débil de formación románica, formado a partir del fuerte *distinctum* latino. Entonces, en la lengua española, habiendo dos participios (léase adjetivos) derivados del mismo verbo *dĭstĭnguĕre*, reparte el significado entre ambas palabras, reservando para *distinguido* el valor ponderativo, y *distinto* para un valor neutro.

Ahora bien, *distinto*, tanto en portugués como en español, entra en conflicto con el significado de otra palabra, también adjetivo derivado del participio de pasado de un verbo: *diferente*. Así, el verbo *diferir* y sus derivados provienen del latín *differo, fers, differre, distŭli, dīlatum* ´dispersar, diseminar, llevar en diversos sentidos`, lo que implica diversidad, lo que no es igual (Blánquez Fraile, 1978: 155), de forma paralela a como *preferente* proviene del verbo latino *praefĕro, fers, ferre, tŭli, lătum* (Corominas y Pascual, II: 494; Blánquez Fraile, 1978: 375), con el significado de ´presentar, ofrecer, mostrar, llevar delante, poner de manifiesto` (Corominas y Pascual, IV: 632-4). Tanto *preferente* como *diferente* se corresponden con el participio de presente de los respectivos verbos antes mencionados. Pues bien, hoy *diferente* significa en portugués "que não é igual; que não coincide; que diverge; diverso, desigual" (Ferreira, 1999: 680), frente a *distinto* en esta misma lengua, que, preferentemente, se usa en el sentido de "que sobressai; notável,

ilustre, eminente, preeminente" (Ferreira, 1999: 695). Esto es, en portugués se da un perfecto ejemplo de *repartición de significado*, pues cuando dos palabras convergen con sentidos parecidos en un determinado estado de lengua, tienden a repartir sus significados –por más que aún no sean extrañas las acepciones en las que se mantiene la sinonimia, reflejo del común o, cuando menos, próximo al significado etimológico–. Por el contrario, como el español desarrolló el participio (adjetivo) *distinguido*, éste se especializó para expresar el sentido de ´sobresaliente, notable, ilustre`, manteniéndose una situación sinonímica entre *distinto* y *diferente*. Así, *distinto*, en el DRAE (2001: 567) se define como: "que no es lo mismo, que tiene una realidad o existencia diferente de aquello otro de que se trata. Que no es parecido, que tiene diferentes cualidades. Y *diferente* como "diverso, distinto" (DRAE, 2001: 556).

5. *Rato* (español) / *rato* (portugués)

La palabra española *rato* proviene del étimo latino *raptus, ūs* ´arrebatamiento, rapto` (Corominas y Pascual, 1983, IV: 795-7), vocablo este último que constituye un doblete por cultismo, *rapto,* con la acepción de "impulso súbito y violento de alguien provocado por un estado pasional, como la ira o los celos, o de locura" (Moliner, 2007, II: 2843). Obsérvese que en la acepción "acción de raptar" el sentido instantáneo se mantiene, pues *raptar* es secuestrar a alguien, particularmente un niño, para obtener un rescate, por ejemplo –pero, de forma rápida, por sorpresa, sin posibilidad de reacción–. Con todo, para la forma portuguesa *raptar* < *raptāre*, Machado (1995, V: 40) otorga el valor de "levar, arrastrar com violência."

Del sentido originario latino se derivan los significados ´tirón, arranque`, y figuradamente ´instante, espacio de tiempo (más o menos prolongado)`. Corominas discute la desviación etimológica que supone el sentido de espacio prolongado de tiempo que ya se documenta en la Edad Media en Berceo, señalando que solo recientemente se entiende *rato* como un período corto de tiempo.

Ahora bien, más allá de la discusión acerca de la duración del espacio de tiempo significado por *rato*, lo que más llama la atención en la interface español-portugués es que la palabra portuguesa *rato* no se refiere al espacio de tiempo, sino al "gênero de mamíferos roedores, murídeos, que apresentam os molares sempre cuspidados, sendo as cúspides dispostas em três séries em relação ao eixo longitudinal" (Ferreira, 1999: 1709). La etimología de esta palabra es oscura, según Machado (1995, V: 42) "vocábulo existente em diversos romances, mas de origem desconhecida." Esto significa, por tanto, que en realidad *rato* en español es una palabra distinta de *rato* en portugués, a pesar de que en el decurso histórico hayan confluido formalmente, tanto en la grafía como prácticamente en la pronunciación.

¿Y cómo se denomina en español al mamífero roedor? Para responder debemos consultar la entrada *rata,* que posee una forma masculina que en realidad es un falso aumentativo, *ratón*, si bien diferenciada no solo por el género, sino por la cualidad, pues se refiere a un tipo o especie de *rata*. En cualquier caso, vale decir que el *rato* portugués es el *ratón* español, toda vez que *rato* significaba en español ya ´espacio de tiempo –fuera este prolongado o no–`. Cuentan Corominas y Pascual (1983, IV: 792-4) que la palabra *rata* tiene origen oscuro –y en esto coinciden con la apreciación de Machado–, pues apareciendo en la mayor parte de las lenguas románicas, el latín tenía *mus*, término del que deriva en el vocablo *mur* que, con ese mismo sentido, aparece en diversos textos medievales. Blánquez Fraile (1978: 311) da noticia de la palabra latina *mus, mūris,* procedente del griego *mys*. Algunas palabras del español mantienen tímidamente el parentesco con la palabra latina: un tipo de *mus*, la *mustela* (lat. *mustēla*), también conocida como *comadreja*, da nombre a una familia o especie, los *mustélidos*.

Con todo, no debemos olvidar algunas de las observaciones de Machado (1995) en su diccionario histórico del portugués, pues añade una segunda entrada para *rato*, con etimología diversa de la anterior, referida al latín *raptu-* y que documenta con el siguiente ejemplo datado en el siglo XIII: *"Mas a pouco rato ouve de fiar…"*, ejemplo que nos muestra cómo, en la época de orígenes, el sentido "espacio de tiempo" para la forma

rato (<*raptus*) no era ajena a la lengua portuguesa, por más que después desapareciera. Aún hay más: el propio Machado (1995, V: 42) otorga una entrada a la palabra *ratão*, que remite a *rato* (mamífero roedor), completamente equivalente al español *ratón*, por más que sea poco o nada común en la actualidad.

En resumen, y con relación a la pareja de falsos amigos propuesta inicialmente, podemos decir que en español *rato* como masculino de *rata* entraba en conflicto con *rato*, espacio de tiempo, lo que no ocurre en portugués, pues desechó ya desde muy antiguo esa posibilidad, solo documentada por Machado en el siglo XIII. De ahí que, en español, se formase un falso aumentativo, *ratón*, que nunca lo fue, pues más que un masculino genérico se refiere a una especie diferente de *rata*, de tamaño menor y no mayor. En verdad, estamos ante sustantivos epicenos, *rata, ratón* (esp.) / *rata, rato* (port.), en los cuales, como apunta Alonso-Cortés (1993: 156), el contraste de género no conlleva necesariamente la oposición macho-hembra, sino más bien una diferenciación de especie. Sin embargo, el español cuenta con este procedimiento de crear un falso aumentativo o despectivo –en realidad, lexicalizados– cuando analizamos otros pares enfrentados, como en el caso de *abeja*, frente a sus respectivos masculinos, *abejón, abejorro* o *abejorruco*:

> "El sufijo *–orro*, que posee un valor despectivo, aparece lexicalizado como ya hemos comentado en formas como *cachorro*. Pudiera muy bien estarlo en la palabra *abejorro*, como masculino de *abeja*, al no contar con la existencia de un hipotético **abejo*." (Miranda Poza, 1991: 174)

De la misma forma que *rato* portugués equivale a *ratón* del español por los motivos aducidos, se producen expresiones y sentidos figurados comunes a ambas lenguas. Así, Ferreira (1999: 1709) documenta en la entrada *rato*, expresiones como, por ejemplo *rato de biblioteca* ´indivíduo maníaco por investigações em bibliotecas e arquivos` y en el DRAE (2001: 1291) encontramos la expresión *ratón de biblioteca* ´erudito que con asiduidad escudriña muchos libros (sentido peyorativo)`.

Por último en español se dan derivados de *rato* (<*raptare*) como *ratero*, que es el ladrón que, en principio, etimológica-

mente hablando, se lleva cosas de forma rápida, al descuido, y, por extensión, cosas de escaso valor, lo que le cualifica como ladrón de poca monta. Y, además, en el RAE (2001: 1290) la entrada *rata* incluye como sexta acepción (coloquial) de *ratero*, "ladrón que hurta cosas de poco valor". Para complicar aún un poco más las cosas, sólo a título descriptivo, Ferreira (1999: 1709) da como séptima acepción de la voz *rato* "ladrão" y como décima, "tratante, canalha."

6. *Bodega* (español) / *bodega* (portugués)

El vocablo *bodega*, tanto en español como en portugués, deriva del latín *apothēca*, término que, a su vez, proviene del griego *ἀποθήκη* ´depósito, despensa`, derivado del verbo *ἀποτιθέναι* ´depositar`, y este, a su vez, de τιθέναι, ´poner` (Corominas y Pascual, 1983, I: 609; Machado, 1995, I: 98). La aféresis que presenta la forma *bodega* es debida a confusión temprana con la *a* del artículo; de hecho, se documenta en castellano antiguo otra forma sin aféresis *abdega* (1104). El portugués, por su parte, presenta una forma *adega* "compartimento da casa, de temperatura baixa e constante, em geral subterráneo, onde se guardam azeite, vino e outras bebidas" (Ferreira, 1999: 49) que convive con *bodega*.

A partir de estos significados, fueron apareciendo otros, por extensión, comenzándose a matizar entre el lugar donde se guardan o almacenan mercancías en general, el lugar donde se comercializan tales mercancías y aún el tipo de mercancías. Así, por ejemplo, en español, la palabra *bodega* alude por igual al lugar de una casa donde se almacenan comestibles, despensa; al lugar interior de los buques, pero en especial, se hace hincapié, como primera acepción, en el tipo de género almacenado, restringiéndolo al vino (DRAE, 2001). Tal vez, la definición que ofrece María Moliner (2007: 390) resulte esclarecedora de los elementos que resultan pertinentes en la lengua española para llegar a la definición precisa de *bodega*: "sótano de las casas en donde se guarda el vino. Por extensión: almacén de vino guardado en toneles, botellas, etc., aunque no sea subterráneo."

Por su parte, el portugués, al contar con el doblete evolutivo *bodega / adega*, distribuye los significados, reservando el término *bodega* a la ´taberna`, esto es, al lugar en el que se comercializan fundamente bebidas, y, en Brasil, además, ´pequeno almazém de secos e molhados` (Ferreira, 1999: 312). A su vez, el término *adega* se especializó para designar el ´compartimento da casa, de temperatura baixa e constante…`. Aún se da, por extensión, otro sentido: ´o conjunto de bebidas guardadas nesse lugar` (Ferreira, 1999: 49). Por lo tanto, no siempre que decimos *bodega* en español y en portugués nos estamos refiriendo a la misma realidad, *falsos amigos*, cuya especificidad respectiva se explica por una doble evolución de la forma latina en portugués.

Pero, además de *bodega* –que comparten español y portugués– y *adega*, que solo tiene el portugués –pues la forma medieval *abdega* documentada en castellano en el siglo XII se perdió–, existe otro vocablo que posee idéntica etimología, tanto en español como en portugués: *botica* (Corominas y Pascual, 1983, I: 643). En español primero, durante la Edad Media, fue, en general, ´tienda y lugar de ventas`; pero, después se especializó en ´tienda de venta de medicamentos`: "Farmacia, establecimiento donde se hacen medicinas y se venden específicos" (Moliner, 2007: 406-7), lo que también vale para el portugués (Ferreira, 1999: 325). De *botica* derivaron *boticario* y también *botiquín –botequim* en portugués–.

Ahora bien, existen desavenencias entre los valores que se asignan respectivamente a *botiquín* y *botequim* en español y portugués. Para comenzar, Corominas y Pascual (1983, I: 643) lo dan como derivado de la forma griega *ἀποθήκη*, como *bodega* y *adega*, pero Machado (1995, I: 454) lo hace derivar del vocablo del italiano *botteghino* ´locale di vendita dei bighietti del teatro`, lo que nos parece poco probable, pues hay otras formas de llegar hasta él, como veremos.[8] Lo que interesa también, además de la etimología, es constatar que en portugués *botequim* es el "estabelecimento comercial onde se servem bebidas em geral (bebidas alcoólicas, refrigerantes, café, etc.) e pequenhos lanches; bar" (Ferreira, 1999: 325), mientras que en español, *botiquín* es, sobre todo, la "habitación, armario o caja portátil donde se tienen medicinas para casos de urgencias"

(Moliner, 2007: 407), si bien, circunscribiendo su uso a Venezuela, la Academia se hace eco de un valor genérico, también en español, que viene a coincidir con el del portugués: "establecimiento modesto en donde se expenden bebidas alcohólicas." Estamos, por tanto, ante un nuevo caso de *falsos amigos* en la interface español-portugués. Dado que históricamente tanto en una como en otra lengua se creó un vocablo *botiquín / botequim* derivado de *botica*, que, a su vez, provenía del término griego *ἀποθήκη*, origen de las palabras *bodega* (español y portugués) y *adega* (portugués), en español se empleó en el *sentido restringido* de ´mueble o caja que contiene medicamentos y similares`, esto es, lo que se vende en la *botica*, y en portugués se usó sin esta restricción, más en la línea de *adega, bodega*, ´local donde se venden bebidas`. Además, el portugués conoce un derivado regresivo de *botequim*, *boteco*, que en Brasil se da como sinónimo del primero, como recoge Ferreira (1999: 325).

Cabría aún preguntarse el porqué de la restricción de significado en el caso del término español *botiquín*. En este caso tenemos que recurrir a la palabra *bote* con el sentido preciso con que se define en el DRAE (2001): "recipiente pequeño, comúnmente cilíndrico, que sirve para guardar medicinas, aceites, pomadas, tabaco y conservas". Es decir, que en las farmacias o *boticas* había *botes* en los que se contenían esas materias y sustancias, y de ahí surgió la forma *botiquín*, cuando tales *botes* y sustancias no se hallaban en el local específico. Ahora bien, obsérvese que en *botes* se pueden guardar también otras materias que no son las propiamente farmacéuticas, en especial, bebidas y comidas, con lo que se entiende no ya la acepción hispanoamericana del término en español, sino la portuguesa.

Ahora bien, en portugués no existe la palabra *bote*, pero sí que posee otra, *pote*, que también existe en español junto a la primera, y con la cual comparte etimología. Etimología, además, oscura, pues se suele relacionar con el francés *pot* (Corominas y Pascual, 1983, I: 643; Machado, 1995, IV: 410), si bien los primeros –a diferencia de Machado, que lo relaciona con el sustrato celta– hacen derivar esta forma de *pōtus,* que en latín clásico tenía la acepción de ´bebida`, de donde el deri-

vado *potione* > *poción, pócima* (español), *poção* (portugués)*, poison* (fr.) ´[poción] venenosa`, ´veneno`.

Entonces, *pote* en portugués es "grande vaso de barro para líquidos […]; recipente de louça, vidro, etc. de largo bocal" (Ferreira, 1999: 1617). En español, se refiere a un "tiesto en forma de jarra en el que se plantan y tienen flores", una "vasija redonda, generalmente de hierro, con barriga y boca ancha y con tres pies, que suele tener dos asas pequeñas, una a cada lado, y otra grande en forma de semicírculo que se usa para guisar (en el mismo sentido que el *pot* francés)", "la comida que se hace en él (también como en francés)", y vaso de barro alto que se usaba para beber o guardar líquidos y prepararlos" (DRAE, 2001). El punto de inflexión entre unos y otros valores, lo encontramos en la última acepción que para *pote* ofrece María Moliner (2007: 816): "vasija de las que se tienen en las farmacias conteniendo las distintas sustancias".

7. *Vaso* (español) / *vaso* (portugués)

La palabra española *vaso*, de la misma forma que la portuguesa *vaso*, provienen etimológicamente de la voz *vasum* del latín vulgar, forma que procede, a su vez, del latín clásico *vas, is*, cuyo significado era en latín lo que en español actual se denomina *vasija*, esto es, ´utensilio destinado a contener algo; vaso, vasija, puchero` (Corominas y Pascual, 1983, V: 749-50; Blánquez Fraile, 1978: 526; Machado, V: 378). De esta palabra deriva *vasija*, que además de ser formalmente un diminutivo lexicalizado, como parece sugerir María Moliner (2007, II: 3007), "designación genérica aplicable a numerosos recipientes", que en la lengua española formaría parte de un hipotético *campo léxico* de los ´recipientes` en español. A su vez, *vasija*, que no es desconocido en la lengua portuguesa, que posee no ya *vasilha* –proveniente del latín *vasilia,* según Machado (1995, V: 378)– posee el colectivo *vajilla*, que en español alterna con la metáfora de la cualidad por la cosa, *loza*, término casi en desuso en la actualidad, siendo su correspondiente portugués *louça* más utilizado que en español. Tampoco debemos olvidar que de *vaso* deriva el término *envase* –formado, a su vez, por *derivación regresiva* a partir del verbo parasintético

envasar–, que posee el significado de ´recipiente dentro del cual se coloca alguna cosa`.

Y, antes de continuar, debemos retomar la información que ya adelantamos en la introducción a nuestro trabajo. No sólo los aspectos históricos resultan, a nuestro entender, fundamentales para un análisis adecuado del problema de los *falsos amigos* en lenguas emparentadas genéticamente como español y portugués, sino que sería el momento de proponer, además, retomar un viejo concepto de la semántica estructural, el de *campo léxico-semántico* (Coseriu, 1991; Salvador, 1985; Justo Gil, 1990).

En efecto, el estructuralismo lingüístico fue el primero en defender de forma más decidida la posibilidad de sistematización del nivel léxico del lenguaje. En virtud de tal hipótesis, el léxico de cada lengua estaría organizado en una serie de estructuras que reúnen en torno a ellas un grupo de lexemas emparentados por algún o algunos *rasgos de significado común*, denominados *semas*. Cada una de esas estructuras —de carácter paradigmático— recibe el nombre de *campo léxico:*

> "En el VIII Congreso Internacional de Lingüistas, celebrado en Oslo en 1957, se planteó como problema la posibilidad del estudio estructural del significado, tema tratado en la sesión plenaria del 8 de agosto, donde actuaron como ponentes el creador de la *Glosemática*, Louis Hjelmslev, y el profesor norteamericano Rulon Wells, de la Universidad de Yale. Puede decirse que, con aquel cambio de impresiones y aquel contraste de pareceres, se echaron las bases de una verdadera *semántica estructural*. El profesor Bernard Portier, que en aquella ocasión se pronunció claramente contra las especulaciones esencialmente teóricas, más preocupadas de su propia coherencia intelectual que de su eficacia, ha publicado un trabajo, *Vers une sémantique moderne*, que es toda una teoría estructural de la Semántica, pero ya ejemplarmente aplicada, y lista para nuevas y sucesivas aplicaciones" (Salvador, 1985: 13).

Pues bien, esta noción de *campo semántico*, en su versión estructural moderna, interesa en primer lugar a una teoría de la traducción, porque suministra las demostraciones más tangibles y más variadas del hecho de que todo sistema lingüísti-

co encierra un análisis del mundo exterior que le es propio, y que se diferencia de otras lenguas u otras etapas de la misma lengua (Mounin, 1971: 94).

Lo que pretendemos demostrar es que, más allá de las diferencias de sentido que las formas enfrentadas poseen en las respectivas lenguas, habríamos de preguntarnos si, efectivamente, se dan diferencias de *significado*, toda vez que provienen de un mismo étimo latino. Es decir, que lo que realmente ocurre es que significando en esencia –en el paradigma– lo mismo, las diferencias observables lo son a nivel de *habla*, como una posibilidad más en la caracterización específica de cada lengua.

Una vez establecida nuestra hipótesis de trabajo, vayamos a la raíz del supuesto problema que, para nosotros, contemplado desde nuestra perspectiva no lo es tanto o, simplemente, no lo es. En efecto, no son pocos los casos en los que se llama la atención acerca del hecho de que la voz *vaso* en español no se corresponde directamente con *vaso* portugués –baste aquí con hacer mención en nuestro entorno inmediato a las observaciones que al respecto elabora Vicente Masip (2003: 144)–, sino en realidad con el término *copo*, formalmente relacionado con el castellano *copa*, término que, a su vez, no tiene que ver con *copa* del portugués –mejor dicho, solo en ciertas ocasiones–, sino que en realidad tiene su correlato portugués en *taça*, que, por tanto, no tiene una correlación directa con el castellano *taza*, que remite, a su vez, al portugués *xícara,* término no completamente ajeno al español, *jícara*, si bien especializado en ser un tipo de *taza* exclusivo para determinados usos y de una forma específica. Aún habría otros elementos relacionados a los anteriores que irían conformando, según nuestra propuesta, el correspondiente *campo léxico-semántico* en cada lengua. Porque, *vaso* en portugués se corresponde con *jarro*, *jarrón* en español, sin que, por lo demás, falten lexemas con correspondencia biunívoca en los respectivos campos, como por ejemplo, *jarra / jarra*. O, incluso, tenemos el caso del par *cáliz / cálice*, que según los valores que posee en portugués coincide solo en parte con el que presenta en la lengua española.

Podemos, entonces, establecer, a modo de esquema, los siguientes cuadros, que representan los respectivos *campos*

léxico-semánticos de los ´recipientes que contienen líquido (u otro tipo de objetos)` en español y portugués:

	s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8	s_9	s_{10}	s_{11}	s_{12}	s_{13}	s_{14}	s_{15}	s_{16}	s_{17}	s_{18}	s_{19}	s_{20}	s_{21}
vasija	+	±	+																		+
*vaso*1	+	+	±	+	+																
*vaso*2	+		+								+	+									
taza	+	+	+	+			+														
tazón	+	+	+	+		+			+											+	
jícara	+	+	+	+										+	+						+
copa	+	+	+	+										+					+		
cáliz	+	+	+	+									+			+			+		
jarra	+	+	+					+		+							+				
jarro	+	+	+	+			+										+				
jarrón	+		+					+			+	+					+	+			

Tabla 1. Campo léxico ´recipiente` en español

Vasija. "Designación genérica aplicada a numerosos recipientes" (Moliner, 2007, II: 3007). / "Pieza cóncava y pequeña, de barro u otra materia y de forma común u ordinaria, que sirve

para contener especialmente líquidos o cosas destinadas a la alimentación" (DRAE, 2001: 543).

Vaso$_1$. "Vasija de vidrio u otro material, de forma generalmente cilíndrica o troncocónica, que se emplea para beber" (Moliner, 2007, II: 3007). / "Recipiente de metal, vidrio u otra materia, por lo común de forma cilíndrica, que sirve para beber" (DRAE, 2001: 1543).

Vaso$_2$. "Aplicado a las decorativas usadas para contener flores, etc., vasija" (Moliner, 2007, II: 3007). / "Obra de escultura, en forma de jarrón florero o pebetero, que, colocada sobre un zócalo, pedestal o peana, sirve para decorar edificios, jardines, etc.; pieza cóncava de mayor o menor tamaño, capaz de contener algo" (DRAE, 2001: 1543).

Taza. "Vasija profunda, de diversas formas, en que se toman líquidos" (MOLINER, II: 2830). / "Vasija pequeña, por lo común de loza o de metal y con asa, empleada generalmente para tomar líquidos" (DRAE, 2001: 1455).

Tazón. "Recipiente aproximadamente semiesférico, sin asa, usada igual que las tazas grandes o de desayuno" (Moliner, 2007, II: 2830). / "Recipiente comúnmente mayor que una taza, de contorno aproximadamente semiesférico, a veces con pie diferenciado y generalmente sin asa" (DRAE, 2001: 1455).

Jícara. "Tacita pequeña con el fondo muy grueso que se usaba particularmente para el chocolate" (Moliner, 2007, II: 1706). / "Vasija pequeña, generalmente de loza, que suele emplearse para tomar chocolate" (DRAE, 2001: 893).

Copa. "Vasija para beber, constituida por un cuenco, generalmente de forma acampanada, sostenido sobre un pie" (Moliner, I: 796). / "Vaso con pie para beber" (DRAE, 2001: 439).

Cáliz. "Vaso de forma de copa, donde se pone el vino que se consagra en la misa" (Moliner, 2007, I: 491). / "Vaso sagrado de oro o plato que sirve en la misa para echar el vino que se ha de consagrar" (DRAE, 2001: 273).

Jarra. "Vasija de loza, porcelana, cristal, etc., con asa y generalmente con un pico en el borde para echar bien el líquido" (Moliner, 2007, II: 1702). / "Vasija de barro, porcelana, loza, cristal, etc., con cuello y boca anchos y una o dos asas" (DRAE, 2001: 891).

Jarro. "Jarra" (Moliner, 2007, II: 1702). / "Vasija de barro, loza, vidrio o metal, a manera de jarra y con una sola asa" (DRAE, 2001: 891).

Jarrón. "Vasija de adorno, por ejemplo para sostener flores, de forma semejante a la de una jarra" (Moliner, II: 1702). /

"Vaso, por lo general de porcelana, artísticamente labrado, para adornar consolas, chimeneas, etc." (DRAE, 2001: 891).

	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}	S_{11}	S_{12}	S_{13}	S_{14}	S_{15}	S_{16}	S_{17}	S_{18}	S_{19}	S_{20}	S_{21}	S_{22}	S_{23}
vasilha$_1$	+	+	+																				
vasilha$_2$	+	±																					
vaso$_1$	+	±	+																				
vaso$_2$	+		+								+	+						+					
jarro$_1$	+	+	+				+			+										+			
jarro$_2$	+		+								+	+											
copo	+	+		+	+																		
xícara	+	+		+		+	+														+		
taça	+	+		+	+														+				
copa	+		+										+									+	
jarra	+	±					+			+		±											
cálice$_1$	+	+	+	+												+							
cálice$_2$	+	+	+	+															+				+

Tabla 2. Campo léxico ´recipiente` en portugués

Vasilha. "Vaso para líquidos"; *Bra.* "Recipiente de uso doméstico utilizado particularmente para guardar ou conter alimentos" (Ferreira, 1999: 2050) / "Qualquer vaso usado para guardar líquidos"; *Bra.* "Tipo de recipiente que serve para conter ou guardar líquidos ou sólidos, especialmente alimentos" (Houaiss & Villar, 2001: 2832).

Vaso. "Qualquer objeto côncavo próprio para conter substâncias líquidas ou sólidas"; "Peça análoga que se enche de terra e onde se plantam flores" (Ferreira, 1999: 2050) / "Recipiente côncavo, de vários formatos, próprio para conter líquidos ou sólidos"; "Esse tipo de objeto feito de material delicado, servindo de peça ornamental, gen. usado como floreira". "Jarro" (Houaiss & Villar, 2001: 2832).

Jarro. "Vaso alto e bojudo, com asa e bico, próprio para água, e us. em geral, para deitar água nas mãos ou na bacia onde se lavam as mãos" (Ferreira, 1999: 1157) / "Vaso alto, ger. com asa e bico, onde se põe água, vinho, etc."; "Vaso para decoração ou para conter flores" (Houaiss & Villar, 2001: 1676).

Copo. "Vaso, em geral cilíndrico, sem tampa, que serve para beber" (Ferreira, 1999: 550). / "Recipiente de vidro, cristal, plástico, etc. ger. cilíndrico, sem asa e sem tampa, pelo qual se bebe" (Houaiss & Villar, 2001: 831).

Xícara. "Pequena vasilha com asa para servir em especial bebidas quentes, como, p. ex., café, chá, leite, etc." (Ferreira, 1999: 2097). / "Pequeno recipiente us. es. para bebidas quentes com asa para facilitar sua manipulação" (Houaiss & Villar, 2001: 2896).

Taça. "Vaso largo de pouca profundidade, ger. provido de pé, para beber" (Ferreira, 1999: 1914). / "Copo cilíndrico com haste, us. para beber vinho, champanhe, conhaque, etc." (Houaiss & Villar, 2001: 2655).

Copa. "Vaso covo; taça; copo" (Ferreira, 1999: 550) / "Vaso fundo de dimensões e formas variáveis, para bebidas; taça"; "*Des* Peça artística com essa forma ou outra análoga, de metal nobre e com inscrições alusivas, que se dá como prêmio ao vencedor de uma competição" (Houaiss & Villar, 2001: 830).

Jarra. "Vaso para água ou para flores; jarro" (Ferreira, 1999: 1156). /"Recipiente para líquidos; vaso, ger. com asa

e bico para acondicionar ou beber água e vinho" (Houaiss & Villar, 2001: 1676).

Cálice. "Vaso empregado na missa para a consagração do vinho"; "Copo com pé, de pequena dimensão, para vinhos licores e/ou outras bebidas" (Ferreira, 1999: 372). / "Espécie de vaso, quase cilíndrico, us. durante a realização da santa missa, para a celebração do vinho"; "copo de forma aprox. semiesférica, alongada ou semelhante a um cone invertido, que tem um pé formado por uma haste mais ou menos comprida e uma base ger. circular; us. para certos tipos de vinho, licores, etc." (Houaiss & Villar, 2001: 575).

Los respectivos campos léxicos han sido establecidos a partir de 23 semas o rasgos mínimos de significado que, en cada caso, se revelan como pertinentes para establecer las respectivas oposiciones léxicas entre los miembros enfrentados, conforme informaciones de los repertorios léxicos (diccionarios) más relevantes de cada lengua. Tales rasgos significativos mínimos son los que integran el significado de un signo lingüístico y son los que presentan la cualidad de diferenciar una unidad semántica (lexema, palabra) de otra (Justo Gil, 1990: 32). La descripción de cada uno de ellos es la que se indica a continuación: S_1 = ´recipiente`; $S_2$ = ´que contiene líquidos`; S_3 = ´cóncavo`; $S_4$ = ´para beber`; S_5 = ´cilíndrico`; $S_6$ = ´semiesférico`; S_7 = ´con una asa`; $S_8$ = ´con una o dos asas`; S_9 = ´sin asa`; $S_{10}$ = ´con un pico en el borde`; S_{11} = ´decorativo`; $S_{12}$ = ´que contiene flores`; S_{13} = ´de oro, plato o material noble`; $S_{14}$ = ´de loza`; S_{15} = ´para tomar chocolate`; $S_{16}$ = ´sagrado, para consagración en misa`; S_{17} = ´con boca y cuello ancho`; $S_{18}$ = ´de porcelana o material delicado`; S_{19} = ´con pie`; $S_{20}$ = ´de tamaño grande o alto`; S_{21} = ´pequeño`; $S_{22}$ = ´trofeo de competición deportiva`; S_{23} = ´para ciertos licores`.

Ahora bien, por la particularidad idiosincrásica de cada lengua (idioma), algunos de esos rasgos mínimos de significado no son relevantes en esencia en los respectivos sistemas: obsérvese, en este sentido que los semas S_8 = ´con una o dos asas`, $S_9$ = ´sin asa`, S_{14} = ´de loza`, S_{15} = ´para

tomar chocolate`, y $S_{17}$ = ´con boca y cuello ancho` no son pertinentes o relevantes para el establecimiento del campo en los lexemas seleccionados para el portugués, del mismo modo que, en la selección de lexemas enfrentados correspondiente al mismo campo en español, no son relevantes los semas S_{22} = ´trofeo de competición deportiva`, $S_{23}$ = ´para ciertos licores`.

Así, por ejemplo, en español, S_{14} = ´de loza`, $S_{15}$ = ´para tomar chocolate` son pertinentes, esto es, relevantes, para definir lo que es (significa) *jícara*, que es un tipo de *taza* en español, aquel tipo que, siendo de pequeño tamaño (S_{21}), sirve para tomar chocolate y es de loza. Una *taza*, en términos generales en español, puede ser variable en cuanto a su tamaño –por lo tanto, los rasgos que se refieren al tamaño mayor o menor no son relevantes en su definición– y en cuanto al material del que está fabricada –a diferencia de la especificidad que presenta la *jícara*–. Sin embargo, el portugués emplea el término *xícara* –de la misma etimología que *jícara* en español, del náhuatl *xikallí* (Machado, 1995, V: 412) – como genérico, no existiendo un término específico para aquella *xícara* que se destina a tomar chocolate.[9] Entonces, cuando en portugués *xícara* posee el significado genérico de *taza* en español, la palabra *taça* del portugués se refiere fundamentalmente al recipiente con pie que sirve para tomar líquidos, esto es, la *copa* del español, término que aun existiendo en portugués se especializa o desplaza para otros significados, unos relacionados directamente con el campo, en especial la forma masculina *copo* –que, a su vez, equivale al español *vaso*–, otros derivados y que casi pierden el significado originario: a) ´trofeo deportivo`, sin duda influido por contacto con la lengua española en Hispanoamérica, aunque también compartiendo sentido con *taça*; b) "compartimento da casa onde se lavam e se guardam as louças e talheres de uso diario, certos géneros alimentícios, etc. e onde, em geral, há mesa de refeições" (Ferreira, 1999: 550), y por extensión, "compartimento em hospitais, hoteis, etc. que atende ao serviço dos quartos e ao preparo de alimentos leves" (Houaiss & Villar, 2001: 830). La no existencia de *copo* en español, hace que el genérico portugués *vaso* se especialice en el recipiente por excelencia

para beber líquido y, en español, comparta usos con el genérico, próximo también a *jarro,* que además presenta un aumentativo *jarrón* no habitual en portugués, lengua en la que *jarro* y *jarra* se refieren más a recipientes con líquido y para beber que al recipiente de carácter ornamental, para el cual se usa *vaso*. Por último, *cáliz* en español se especializó en el recipiente para consagrar el vino en misa, mientras que en portugués, al lado de dicho significado, se mantiene el de *taça* (léase *copa* en español) destinada a determinados licores –lo que justifica la presencia de un sema pertinente, S_{23} = ´para ciertos licores`, que no aparece, por no pertinente, en la descripción del campo léxico en español–.

Pues bien, confrontando los dos campos léxicos correspondientes a los respectivos sistemas lingüísticos (español y portugués), observamos, en primer lugar, las siguientes correspondencias lexemáticas:

Español	*Portugués*
vasija	$vasilha_2$
$vaso_1$	copo
$vaso_2$ jarro jarrón	vaso $jarro_2$
taza	xícara
cáliz	$cálice_1$
copa	taça $cálice_2$ $copa_2$
jarra	jarra $jarra_1$
jícara	-

Tabla 3 Correspondencias lexemáticas español-portugués en el campo léxico ´recipiente`

Ahora bien, obsérvese que, más allá del lexema que en cada lengua se especializó en expresar un sentido específico, las correspondencias no siempre coincidentes se deben

a que, en ambas, en realidad poseen el mismo significado. Y esto se produce por dos motivos: primero, porque ambas lenguas proceden históricamente del latín y cuando fueron apareciendo en el curso de la historia otros vocablos procedentes de otras lenguas –*jarra, jícara, taza* (*jarra, xícara, taça*)–, los respectivos sistemas lingüísticos se fueron reorganizando en cada caso, pues tal organización se hacía necesaria porque la aparición de un nuevo término entraba en conflicto con el *campo de dispersión* del vocablo patrimonial o preexistente; en segundo lugar, porque, como muestra el análisis componencial de cada uno de los lexemas descritos en los sistemas lingüísticos enfrentados, éste presenta una base común en los respectivos *sememas* o "conjunto de rasgos significativos mínimos que integran el significado de un lexema" (Justo Gil, 1990: 32). En efecto, tanto en portugués como en español casi en la totalidad de los lexemas seleccionados aparecen los rasgos S_1 = ´*recipiente*`, $S_2$ = ´*que contiene líquidos*` y S_3 = ´*cóncavo*`.

Entonces, la pregunta que surge es: ¿se trata realmente de términos heterosemánticos? Porque, en efecto, una cosa es el significado y otra el uso o el sentido inmediato que en cada lengua un determinado vocablo posee. Y es justamente esta visión la que explica el equívoco (a ello volveremos más adelante). Porque, además, en realidad, la posibilidad de deducir a partir de la distribución de una unidad léxica su significado, entendiendo, a su vez, por distribución el conjunto de todos los contextos o entornos lingüísticos en los que la unidad léxica se encuentra o pueda encontrarse, llevado al extremo, provocaría

> el considerar que el significado de una unidad léxica es su uso y las palabras solo tendrían significado en un contexto determinado, *olvidando, así, que las palabras tienen un núcleo significativo fijo y estable independiente del contexto en el que aparecen*." (Justo Gil, 1990: 28, el subrayado es nuestro)

En este mismo sentido, no deben sorprendernos en exceso las extensiones de significado –de carácter connotativo o metafórico– que se producen en cada una de las lenguas, como por ejemplo *vaso sanitário* en portugués, cuando en

español *vaso* se emplea para significar lo mismo que el portugués *copo*, ni, recíprocamente, la metaforización paralela del español: *taza del váter*. Piénsese que tanto en una como en otra expresión prevalece el significado original latino y árabe respectivamente (*vasum* < *vas, is, ṭássa*), más allá de las particularidades históricas de uso que fueron contornando la peculiaridad de cada una de las lenguas.

8. Conclusión

Hemos abordado en las páginas precedentes un problema que ha sido tratado en no pocas ocasiones y desde otras perspectivas cuando se analizan las relaciones entre el portugués y el español en el ámbito del vocabulario, y muy en especial, con relación a los problemas de enseñanza-aprendizaje de lenguas extranjeras, al desarrollo de la *interlengua* en el aprendiente, así como a cuestiones atinentes a la interferencia lingüística en general.

Las metodologías –o mejor, si se prefiere, las estrategias– que el profesor de lengua extranjera debe desarrollar en el ejercicio de su actividad docente tienden a fomentar en los últimos tiempos las cuestiones más relacionadas con la comunicación y la interacción, de carácter pragmático e inmediatista, con lo que se produce una, a nuestro entender, peligrosa tendencia a olvidar –incluso rechazar– las cuestiones gramaticales y aún más, si cabe, las referencias históricas.

Sin duda, en función de las necesidades específicas del profesor en el aula derivadas del tipo de aprendiente y de enseñanza que cada circunstancia exija, las metodologías y/o estrategias en boga pueden ser válidas, pero cuando hablamos de formación de profesores a nivel universitario, esto es, de formación de futuros formadores, no podemos dejarnos llevar por el inmediatismo o el pragmatismo, o lo que es lo mismo, dicho en otros términos, por la comodidad o el simplismo. Porque una cosa es enseñar una lengua –extranjera o no– a alguien que la quiere aprender y otra muy distinta enseñar a ser profesor de esa lengua.

En este sentido, un profesor debe ir más allá de lo meramente aparente y debe conocer en profundidad las estructuras de la lengua y no solo contentarse con la contemplación de su dinámica activa en el uso interactivo que proporciona la comunicación hodierna.

Los fenómenos equívocos a los que abocan, en la interacción comunicativa, los heterosemánticos o falsos amigos, muy en especial en lenguas tan próximas como el español y el portugués, tienen su porqué, su explicación, que va mucho más allá de la mera anécdota a la que conduce el propio equívoco y se deben a factores que solo se explican desde el conocimiento de la historia de la lengua y que remontan a las lenguas clásicas –latín y griego– y aún a otras –léase aquí árabe o lenguas indígenas de América del Sur– que, en el curso de la historia, por diferentes motivos, entraron en contacto con las lenguas en cuestión y provocaron la reorganización de los sistemas lingüísticos y de los respectivos valores que asignaban a los vocablos –antiguos e importados–.[10]

Descubrimos, por fin, que las aportaciones de la lexemática –como desarrollo de la semántica estructural– son de especial interés en lo que atañe al análisis del significado y que, además, el *cognitivismo* ha llegado a justificar, desde perspectivas radicalmente distantes, la pertinencia del *análisis componencial* del significado de las palabras, que precozmente proponía aquella, a través de experimentos de carácter psicológico de aplicación de la teoría conductista, cuando se pretende investigar los mecanismos que se producen en la mente humana cuando se trata de designar un determinado objeto –en mayor o menor medida conocido– mediante una palabra.

BIBLIOGRAFÍA

Alonso-Cortés, Ángel. 1993. *Lingüística General.* 3ª ed. corregida y aumentada. Madrid: Cátedra.

Beato, Ana Paula Marques; Durao, Adja Balbino Amorim Barbieri. 2001. A influência da língua materna na aprendizagem do léxico de uma segunda língua. *Entretextos. Revista da Pós-Graduação em Letras: Estudos da Linguagem* 2. 23-36.

Blánquez Fraile, Agustín. 1978. *Diccionario manual latino-español y español latino.* Barcelona: Ed. Ramón Sopena.

Breal, Michel [1897]1982. *Essai de Sémantique (Sicence des significations).* Saint-Pierre de Salerne, Brionne: Gérard Monfort Editeur.

Bugueño Miranda, Félix Valentín. 1998. Sobre algunos tipos de falsos cognados. *Anuario Brasileño de Estudios Hispánicos* VIII. 21-27.

Buitrago Jiménez, Alberto; Torijano, J. Agustín. 2007. *Diccionario del origen de las palabras.* Madrid: Espasa-Calpe.

Camorlinga Alacaraz, Rafael. 2005. Do português ao espanhol: os prós e os contras da proximidade. In: Sedycias, João (org.) *O ensino do espanhol no Brasil: passado, presente, futuro.* São Paulo: Parábola, 195-205.

Corominas, Joan; Pascual, José Antonio 1983. *Diccionario Crítico Etimológico Castellano e Hispánico.* Madrid: Gredos.

Coseriu, Eugenio 1991. *Principios de semántica estructural.* 2ª ed. Madrid: Gredos.

Durao, Adja Balbino Amorim Barbieri. 2005. Lenguas parecidas, no obstante, diferentes. El estado de la cuestión de los estudios de español en Brasil. *Fiape I Congreso Internacional: El español, lengua del futuro.* Toledo. s.p. [publicación electrónica]. Disponible en la dirección de la Internet: *www.mecd.gob.es/redele/Biblioteca-Virtual/2005/Numeros-Especiales/Mayo_I_FIAPE.htlm* Último acceso: 27 de mayo de 2013.

—. 2004. *Análisis de errores en la interlengua de brasileños aprendices de español y de españoles aprendices de portugués.* Londrina: Eduel.

—. 2002. ¡Por detrás de la transparencia existe una opacidad! La lectura por estudiantes brasileños de textos redactados en español. *Anuario Brasileño de Estudios Hispánicos* XII. 13-28.

Feijóo Hoyos, Balbina Lorenzo; Hoyos Andrade, Rafael Eugenio. 1992. *Dicionário de falsos amigos do espanhol e do português.* São Paulo: Embajada de España / Consejería de Educación / Scritta Editorial.

FERREIRA, Aurélio Buarque de Holanda. 1999. *Novo Aurélio Século XXI: O dicionário da língua portuguesa.* 3ª ed. totalmente revista e ampliada. Rio de Janeiro: Nova Fronteira.

HOUAISS, Antônio; VILLAR, Mauro de Salles. 2001. *Dicionário Houaiss da Língua Portuguesa.* Rio de Janeiro: Editora Objetiva.

JUSTO GIL, Manuel. 1990. *Fundamentos del análisis semántico.* Santiago de Compostela: Universidade. Servicio de Publicacións e Intercambio Científico.

KOESSLER, Maxime; DEROCQUIGNY, Jules. 1928. *Les faux amis ou Les trahisons Du vocabulaire anglais (conseils aux traductuers).* Paris : Librairie Vuibert.

KOFFKA, Kurt. 1926. *Bases de la evolución psíquica.* Traducción de J. Gaos. Madrid: Revista de Occidente.

LABOV, William 1973. The Boundaries of Words and their Meanings. In: BAILEY, Charles J.; SHUY, Roger W. (eds.) *New Ways of Analizing Variation in English.* Washington: Georgetown University Press. 340-73.

LADO, Robert. 1973 *Lingüística contrastiva.* Madrid: Ediciones Alcalá.

LEIVA, Myriam Jeannette Serey. 1994. *Falsos cognatos em português e espanhol.* Dissertação (Mestrado em Linguística). Orientadora: Eunice Ribeiro Henriques. Campinas: Instituto de Estudos da Linguagem / Universidade Estadual de Campinas. Disponible en la dirección de la Internet: *www.bibliotecadigital.unicamp.br/document/?code=vtls000082335&fd=y* Arquivo: Leiva, Myriam Jeanette Serey.pdf. Último acceso: 27 de mayo de 2013.

MACHADO, José Pedro. 1995. *Dicionário Etimológico da Língua Portuguesa. (Com a mais antiga documentação escrita e conhecida de muitos dos vocábulos estudados).* 7ª ed. Lisboa: Livros Horizonte.

MASIP, Vicente. 2003. *Semântica: Curso-oficina sobre sentido e referência; com um apêndice contrastivo português/espanhol.* São Paulo: EPU.

MIRANDA POZA, José Alberto. 2010. *Introdução à Linguística.* Recife: Editora Universitária da UFPE.

—. 2007. *España y América. Tres ensayos de lengua y literatura.* Recife: Bagaço.

—. 1991. Notas para un estudio de la sufijación nominal en andaluz y canario. *Notas y Estudios Filológicos* 6. 147-216.

MOLINER, María. 2007. *Diccionario de uso del español.* 3ª ed. Madrid: Gredos.

MOUNIN, Georges. 1971. *Los problemas teóricos de la traducción.* Madrid: Gredos.

MUÑOZ NUÑEZ, María Dolores. 1999. *La polisemia léxica.* Cádiz: Universidad de Cádiz, Servicio de Publicaciones.

PEIXOTO, Francisco Balthar. 2007. *O interior das palavras.* Recife: Editora Universitária da UFPE.

REAL ACADEMIA ESPAÑOLA. 2001. *Diccionario de la lengua española*. 22ª ed. Madrid: Espasa-Calpe.

ROBERT, Paul. 1977. *Le petit Robert. Dictionnaire alfabetique et analogique de la langue française.* Rédaction dirigée par Alain Rey et Josette Rey-Debove. Paris: Société du Nouveau Littré.

ROSCH, Eleanor. 1975. Human Categorization. In: WARREN, N. (ed.) *Advances in Cross Cultural Psychology.* London: Academic Press. 1: 1-72.

SABINO, Marilei Amadeo. 2006. Falsos cognatos, falsos amigos ou cognatos enganosos? Desfazendo a confusão teórica através da prática. *Alfa,* São Paulo, 50.2. 251-63.

SALVADOR, Gregorio. 1985. *Semántica y lexicología del español. Estudios y lecciones.* Madrid: Paraninfo.

SCHMITZ, John Robert. 2001. A problemática dos dicionários bilíngues. In: OLIVEIRA, Ana Maria Pinto Pires de; ISQUERDO, Aparecida Negri. (org.) *As ciências do léxico: lexicologia, lexicografia, terminologia.* 2ª ed. Campo Grande, MS: Editora da UFMS. 1: 161-70.

TRIER, Jost. 1931. *Der Deutsche Wordschatz in Sinnberzirk des Verstandes. Die Gesichte eine Sprachilchen Feldes.* Heildelberg: C. Winter.

ULLMANN, Stephen. 1986. *Introducción a la semántica francesa.* Traducción y anotación por Eugenio de Bustos Tovar. Madrid: Instituto de Filología / Publicaciones de la Revista de Filología Española.

—. 1965. *Semántica. Introducción a la ciencia del significado.* Madrid: Aguilar.

WERNER, Reinhold. 2006. El diccionario bilingüe y la enseñanza del español como lengua extranjera. *SIGNUM: Estudos Linguísticos*, 9.1. 205-38.

—. 1997. Algunos elementos de una teoría del diccionario bilingüe. *Cicle de conferències 95-96. Lèxic, corpus i diccionaris.* Barcelona: Institut Universitari de Lingüística Aplicada / Universitat Pompeu Fabra. 113-31.

NOTAS

[1] "Ou *brincar* provenha de *springan*, no sentido de ´pular`, e de *bli(n)kan*, no de ´gracejar, entreter-se`, sendo portanto formas convergentes; ou proceda de um só destes verbos germânicos, sendo a segunda acepção desenvolvimento da primeira; ou ainda, o substantivo *brinco* significando ´pingente` seja o latim *uinc(u)lum*, independente portanto de *brinco*: o que é certo é que este em português adquiriu significados em que o seu correspondente castelhano *brincar*, ´pular`, o não seguiu, pois se diz ali *jugar, juguetear. Brinco:* ´brinquedo`. Documentada no séc.XVI com o sentido de ´enfeite, ornamento` e mais concretamente ´enfeite de orelhas" (Machado, I: 465).

[2] "Todas las palabras que debieran ser sinónimas y que incluso lo han sido momentáneamente en el pasado, han adquirido significados diferentes y han dejado de ser intercambiables. La sinonimia sería, pues, un accidente una coincidencia pasajera a la que la lengua pone enseguida remedio" (Salvador, 1985: 53).

[3] Cf. la solución que ofrecieron otras lenguas románicas, en especial, las del grupo galorrománico. Así, en lugar de adoptar derivados del verbo latino *salīre* para recubrir el significado de *exīre* (fr.ant. *issir*), se recurrió a un verbo *sortir* y sus respectivos derivados, *sortida* (cat.), *sortie* (fr.). La etimología de *sortir* es oscura (Robert, 1977: 1837-8). Se sugiere, de un lado, la alteración del latín clásico *surrectus* en la forma vulgar **surctus*, participio del verbo *surgo, is ĕre, surrexi, surrectum*, ´poner en pie, erigir, levantar` (Blánquez Fraile, 1978: 485), a la que se asociaría el nuevo sentido. De otro, mucho más probable que la tesis anterior, se relaciona *sortir* con el verbo latino *sortĭor, īris, īri, ītus sum* ´sacar la suerte, echar la suerte`, derivado, a su vez, del sustantivo *sors, sortis* ´todo objeto que se ponía en una urna o cajita para sacar a suerte` (Blánquez Fraile, 1978: 463).

[4] Ello no quiere decir que, necesariamente, la estructuración que el sistema léxico de cada lengua va adoptando a lo largo de su evolución histórica haya de cubrir todos los espacios de significado a través de lexemas específicos. Ni siquiera que, en un determinado corte sincrónico, los sistemas léxicos de una lengua dada hayan de cubrir biunívocamente los diferentes significados (Salvador, 1985: 42-3). En efecto, Coseriu (1977), desdice la sugerencia que, en este sentido, parecía realizar Trier (1931) cuando, en los albores de la semántica, definió el *campo semántico* como un "conjunto de palabras no emparentadas etimológicamente en su mayor parte (ni tampoco unidas entre sí por asociaciones psicológicas, individuales, arbitrarias, contingentes) que, colocadas una al lado de otra como las piedras irregulares de un mosaico, recubren exactamente todo un terreno bien delimitado de significaciones, constituido, ora tradicionalmente, ora científicamente, por la experiencia humana" (Miranda Poza, 2010: 92).

[5] "VONTADE [Do lat. Voluntate] 1. Faculdade de representar mentalmente um ato que pode ou não ser praticado em obediência a um impulso ou a motivos ditados pela razão. 2. Sentimento que incita alguém a atingir o fim proposto por esta faculdade; aspiração; anseio; desejo" (Ferreira, 1999: 2087).

[6] "PERGUNTAR. Esta forma é bastante ant. em Port., se bem que na fase arcaica deste se documente mais raramente que PREGUNTAR. Essa maior vulgaridade nos nossos textos medievais, ao lado do castelhano PREGUNTAR, obriga-nos a preferir a hipótese que tira este voc. do lat. Vulgar *PRAECUNCTARE, *PRECUNTARE ou *PRECONTARE, alterações de qualquer destas formas do latim clássico: PERCONTARE e PERCUNCTARI ´inquirir, interrogar, questionar; sondar, no sentido moral`. PERGUNTA: deriv. regressivo de PERGUNTAR" (Machado, 1995, IV: 344).

[7] "O sentido moderno [de *esquisito*] data do séc. XIX, traduz as ideias de ´fendido, separado, dividido`, do Gr. schistós, ´separado, dividido` (Machado, 1995, II: 473).

[8] La palabra italiana *botteghino,* aludida por Machado con el sentido de ´local de pequeñas dimensiones´, parece más cercano a los procesos metafóricos a partir del francés *boite,* en principio, ´recipiente para guardar diversos materiales y cosas, esto es, caja`, que a partir del siglo XV se aplicó al ámbito de la residencia, para referirse a lugares de reducidas dimensiones, por ejemplo, *boiste aux cailloux*, ´prisión`. (*apud Centre National de Recherches Textuelles et Lexicales* < *www.cnrtl.fr/etymologie/boite* >). De aquí se deducen los valores actuales del galicismo en español y portugués: *boîte* ´sala de fiestas o discoteca`.

[9] Cuando el término *jícara* llega al español proveniente del náhuatl, *se especializa* con ese valor y sin duda con ese uso ´para tomar chocolate`, además de porque el chocolate era también un producto importado de América (Corominas y Pascual, 1983, III: 513-514), por más que en el español de América, más en concreto, en México, América Central y Cuba, *jícara* posea un sentido más genérico: ´vasija para depositar alimentos` o incluso ´vasija pintada para adorno de la casa`, evidentemente, una vez que perdió su valor originario, al chocar con el ya patrimonial *taza*, si bien procedente a su vez del árabe *ṭássa*, ´escudilla`, ´tazón`, ´caldero` (Corominas y Pascual, V: 443).

[10] Cf. a este respecto lo que dijimos en otro lugar a propósito de las palabras *almadía* y *canoa* en virtud de las informaciones contenidas en los textos de los cronistas de Indias (Miranda Poza, 2007: 68-91).

LÉXICO SOBRE UTENSILIOS DE LABRANZA DE CEREAL EN LA PROVINCIA DE SORIA

Elena Jiménez García
Universidad de Valladolid

Resumen: En este trabajo se han recopilado los términos relativos a los utensilios de labranza que se utilizaban antes y después de la mecanización del campo, pues este cambio ha repercutido en la terminología agraria, algunos términos han dejado de usarse aunque, por el contrario, también se han creado otros nuevos.

Posteriormente, se han definido para conocer y valorar las posibles diferencias en el uso de términos distintos en diversas localidades (sinonimia) o variantes de significado de un mismo término (polisemia) y, además se ha comprobado si las obras lexicográficas del español más utilizadas y otras obras más específicas recogen estas formas.

Palabras clave: léxico, utensilios de labranza, variantes léxicas, agricultura de cereal.

Abstract: This paper is about the compilation of terms related to carving tackles before and after the introduction of agricultural machinery. Farming terminology has undergone changes after the introduction of agricultural machinery, for example some terms have become obsolete and, on the other hand, others have been created.

Furthermore, we define these terms to know whether there are differences in the use in different boroughs (synonymy) or meaning variants (polysemy). In addition, we have checked these terms in the most used Spanish lexicographical works and in some more specific works.

Keywords: lexicon, carving tackles, lexical variants, cereal farming

1. Introducción

Hasta aproximadamente los años 50 del siglo veinte, en la provincia de Soria, las labores agrícolas se realizaban a mano, con utensilios y herramientas, y con la ayuda de animales, a los que preparaban ataviándolos con aparejos. Fue a partir de la segunda mitad del siglo cuando llegaron a la provincia los pri-

meros tractores y la mecanización del campo en las labores agrícolas, especialmente del cereal, que es la agricultura más productiva en la provincia soriana. A partir de entonces, comenzó a reducirse el esfuerzo físico del agricultor, puesto que se intensificó el uso de las máquinas y se produjeron cambios importantes en las técnicas de cultivo, en los sistemas de producción agraria y en la sociedad rural. Paralelamente, las tendencias migratorias hicieron que la población activa agraria se redujera considerablemente. A esto se añade que la población que realizó los trabajos agrícolas, utilizando otros utensilios diferentes a los actuales, actualmente tiene una edad avanzada, y con ellos, en un periodo relativamente corto de tiempo, se van a perder esas palabras y expresiones que, de hecho, ya no se utilizan y, en muchos casos, ni siquiera resultan familiares a los agricultores más jóvenes.

Ante esta situación, surge la necesidad de dar una visión general del léxico de aquellos recursos que en los últimos sesenta años, aproximadamente, han estado a disposición de las personas dedicadas al mundo agrícola del cereal, además de recuperar aquellas palabras en desuso y en peligro de desaparición y olvido, así como de entrever una pequeña muestra de la evolución lingüística en variantes y diferencias léxicas entre unos hablantes y otros, a veces separados por varios kilómetros y en otros casos separados por escasos metros. Todo ello forma parte de la cultura de una sociedad y, probablemente, resulte interesante para agricultores, investigadores y curiosos la recopilación de una parte de ella.

Este estudio es una pequeña sección del texto que en septiembre de 2011 se presentó en la Universidad de Valladolid como tesis doctoral con el título *Contribución al estudio del léxico agrícola de Soria*. Está centrada en la agricultura del cereal, puesto que es ésta la de mayor producción en la provincia soriana.

Ámbito geográfico

Soria está situada en el sector oriental de Castilla y León y cuenta con una superficie de 10.287 km² en la que

residen sólo 94.610 habitantes, lo que hacen de Soria la provincia española menos poblada, con 40.286 habitantes en la capital, y una de las menos habitadas de la Unión Europea, con una densidad de 9,2 habitantes/km². La capital cuenta con el 42,58 % de la población total.

Como dato relevante, es notable destacar la media de edad que presenta la población de Soria, que es de 47 años, una población relativamente mayor, según los datos que se recogen en la página web del Instituto Nacional de Estadística (http://www.ine.es), correspondientes a 2011, de lo que se deduce que la población joven emigra a otras ciudades, en las que encuentran más oportunidades de formación y empleo. Según los datos del Instituto Nacional de Estadística, de 2001, Madrid es la ciudad donde más sorianos habitan (20.516), seguida de Zaragoza (15.393) y Barcelona (14.463), ciudades donde se encuentran más oportunidades. El descenso demográfico parece haber tocado fondo con el cambio de siglo, ya que se han registrado más corrientes migratorias en los últimos años procedentes del exterior, principalmente de Hispanoamérica, África o los países del este de Europa. El número de nacidos en Soria que viven en esta provincia es de 68.290 y en el resto de España es de 147.964.

La provincia cuenta con más de 500 núcleos de población, agrupados en 183 municipios, de los que la mayoría son pueblos de menos de 100 habitantes, pues únicamente Almazán y El Burgo de Osma tienen más de 5.000 habitantes y trece municipios más superan los 1.000.

Para otorgar la dimensión geográfica al trabajo, hemos tomado como referencia las comarcas agrarias que contempla el mapa publicado por el Ministerio de Agricultura, Pesca y Alimentación en 1996.

Todos estos datos geográficos nos aportan unas características concretas que van a ser punto de inflexión a la hora de llevar a cabo la metodología ya que, al disponer de una población tan reducida, también se ve aminorada la población agrícola y, por lo tanto, los posibles informantes. Además hay que tener en cuenta la despoblación de algunas localidades durante el invierno, o los movimientos migratorios, que son causa de influencias lingüísticas en los hablantes. También, como veremos, es razón

de riqueza léxica las migraciones de personas que durante los años 50 y 60 del siglo pasado venían a esta provincia para realizar las labores agrícolas, especialmente y entre otras, de Andalucía y Extremadura.

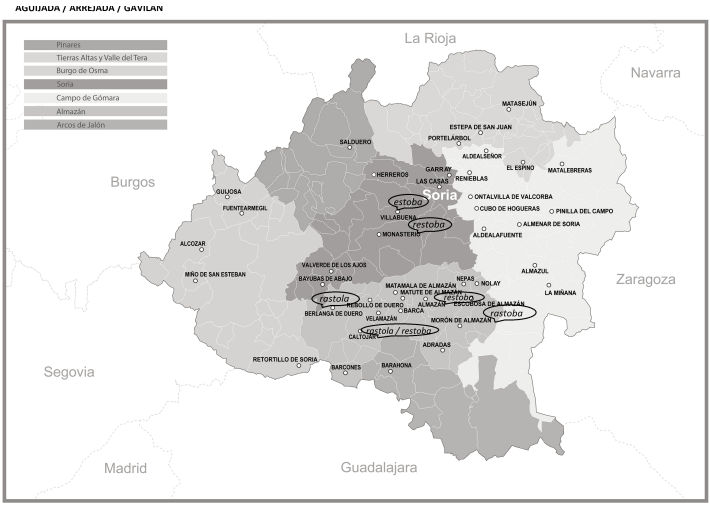
2. Metodología

Partiendo de la lectura de una bibliografía general y específica sobre agricultura y léxico agrícola (García Malmierca, 2007) (González Torices y Díez Barrio, 1991) (Herrera, 1994) (Martín Criado, 1999) (Mingote Calderón, 1990) (Soldevilla, 1994) hemos extraído términos relativos a utensilios que se utilizaban para el cultivo de cereal. Posteriormente se ha elaborado un cuestionario que ha sido respondido por una muestra de cien informantes, en su mayoría agricultores profesionales (68%) y, el resto, familiares directos y personas relacionadas con la agricultura. Para obtener la información se han realizado visitas a diversas localidades sorianas y allí hemos seleccionado las personas dedicadas a la agricultura o que tengan relación con ella (familiares directos), quienes accedieron a ayudarnos.

Algunos cuestionarios han sido registrados digitalmente con una grabadora de voz, pero no todos ellos, ya que algunos informantes han preferido permanecer en el total anonimato.

Los cuestionarios constan de tres partes: una sección correspondiente a los datos personales, a partir de los cuales podremos hacer un estudio estadístico posterior de los usos de los términos; una sección de cuestiones sobre los utensilios, su definición y utilidad, relacionadas con el cultivo de cereales; y otra sección que consta de un dossier de fotografías, algunas extraídas de manuales y catálogos y otras de elaboración propia, mediante las que el informante nos explica lo que representan, su uso, e incluso anécdotas o historias que nos han proporcionado nuevas unidades léxicas relacionadas con el tema de nuestra investigación.

Las localidades donde se han realizado los cuestionarios están indicadas en el siguiente mapa:



Comarcalización agraria de Soria y localidades en las que se ha realizado el cuestionario (elaboración propia).

3. Análisis de los datos

La información ha sido recogida sin tener en cuenta únicamente los términos de forma individual, sino que se han enmarcado en el contexto, lo que nos ha aportado mayor concreción de significado.

El uso del cuestionario ha favorecido la obtención de algunas formas que no se habían extraído de los corpus previamente, como *arcial, bielo, cernedor, estola, estoba* o *rastola* entre otras.

Los términos obtenidos han sido ordenados alfabéticamente para definirlos y facilitar la búsqueda de información al usuario. Después se han indicado las variantes y las localidades donde se utilizan. En las definiciones aparecen algunas palabras marcadas con un asterisco (*), que indican que también están definidas y analizadas en este trabajo. Únicamente en las entradas correspondientes a variantes, se incluye el enlace al diccionario online de la Real Academia Española, porque los términos genéricos están incorporados. Se ha utilizado el signo ✚ en los enlaces para términos recogidos con la misma acepción que hemos indicado, el signo ● si aparece con distinta acepción y un guión - si no está recogido. Hemos elegido el DRAE como referencia para incluir esta información, por ser la

obra lexicográfica a la que los hablantes de español le otorgamos el valor normativo. Por ejemplo, *jalma* aparece definido, *manilla* aparece definido con otra acepción y *matabueyes* no está incluido en el DRAE.

Pero también se ha comprobado en obras lexicográficas de referencia si los términos aparecen o no, si se recogen con la misma o con distinta acepción, o si no están incluidos. Esto se debe a que sentí la curiosidad de comprobar la pérdida de los términos en los diccionarios más utilizados en español y en otros más específicos: el *Diccionario de la Real Academia Española (DRAE)* y el *Diccionario de uso del español (DUE)* de María Moliner. Posteriormente, también comprobé si los términos que yo recogí aparecían en otros dos diccionarios más especializados: el *Diccionario del Castellano Tradicional (DCT)* y el *Diccionario de habla soriana (DHS).* Así mismo, consulté también el *Corpus de referencia del español actual (CREA),* el *Nuevo Tesoro Lexicográfico de la Lengua Española (NTLLE)* y el *Atlas Varilex*.

Este es un ejemplo de una entrada completa:

aguijada: Instrumento formado por una hoja metálica semicircular unida a un tubo por el que se introduce un astil de madera. Su extremo acaba en una punta de hierro. Tiene dos finalidades: por un lado, sirve para limpiar la tierra que se pega al arado, y por otro, sirve para animar al ganado que está remiso al trabajo. (Barcones)

- *arrejada*, *estoba* (Villabuena), *gavilán*, *hijada*, *rastoba* (Morón de Almazán), *rastola* (Caltojar, Berlanga de Duero), *rejada, rejón, restoba* (Almazán, Caltojar, Monasterio), *restola*

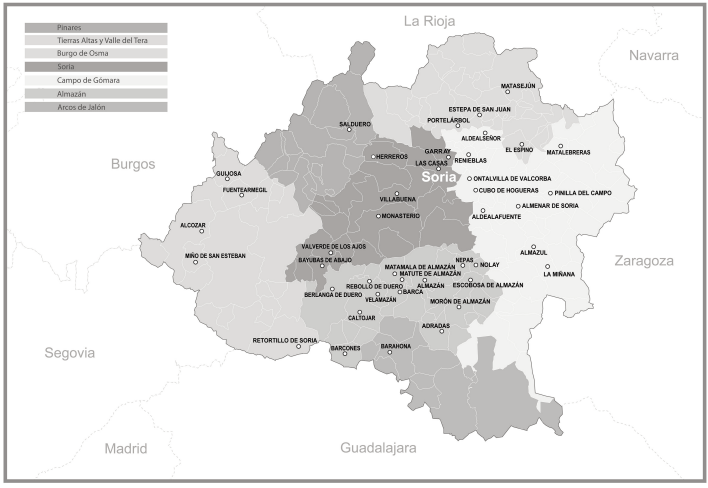


Aguijada. Elaboración propia

DRAE	DUE	DCT	DHS	CREA	NTLLE	VARILEX
✚	✚	✚	✚	✚	✚	▬

Definición del DRAE en línea:
http://buscon.rae.es/draeI/SrvltConsulta?TIPO_BUS=3&LEMA=aguijada

Al final y a modo de anexo se incluye un mapa en el que se señalan las variantes en las localidades en las que se han recogido.



Comarcalización agraria de Soria y localidades en las que se ha realizado el cuestionario (elaboración propia).

4. Conclusiones

La realización de esta investigación ha dado lugar a la confirmación de algunas de las hipótesis con las que partimos:

1. Efectivamente, los cambios sustanciales que se introdujeron en la producción agrícola, en la mecanización, en los sistemas de cultivo, etc., han supuesto grandes transformaciones en la sociedad rural, y han tenido también repercusión en diferentes denominaciones. Como hemos comprobado, todavía quedan rastros lingüísticos de términos referentes a utensilios y, con este tipo de trabajos, se ayuda a su conservación, aunque sea de una manera escrita, y se facilita a investigadores y curiosos una información que puede serles de gran utilidad.

Se recogen las formas orales y también escritas, ya que algunos informantes han preferido realizar el cuestionario por escrito. De esta manera también apreciamos, en algunos casos, su nivel cultural, ya que determinados errores fonéticos se ven reflejados en la escritura. Es el caso de *harnero, arnero;* o *biela, viela, bielda*.

2. Los más mayores recuerdan unidades léxicas que están en peligro de desaparición, mientras que los más jóvenes conocen y utilizan terminología relativa a la maquinaria, abonos químicos, etc. que los mayores prácticamente desconocen, o simplemente han escuchado su nombre.

Existen numerosos términos desconocidos para la población local, que sólo se conservan en la mente de la población más vieja y que están en vías de extinción. No hablamos de sustitución total en el léxico, ya que no se utilizan unos términos en lugar de otros, sino que, en muchos casos, se trata de la pérdida de palabras en el uso y la nueva creación de otras para nuevos instrumentos, herramientas, máquinas, labores, etc.

3. Un referente recibe términos diferentes por diversas razones, principalmente geográficas y culturales. Algunos de ellos son sinónimos, pero otros son variantes lingüísticas deformadas por etimología popular, metátesis, asimilación, hipercorrección, etc.

4. Los usos de este léxico en distintas localidades se deben, en parte, al movimiento pecuario, que fue una de las vías de comunicación e influencia en el uso de las palabras. En los años 40-50 del siglo pasado, los pastores fueron una fuente de migración de palabras de unas localidades a otras.

5. Las obras lexicográficas consultadas no recogen todas las variantes, sobre todo las no canónicas, dado que no es un hecho general de la lengua.

6. Algunas variantes reflejan el nivel cultural de los hablantes. Los informantes más mayores carecen en su mayoría de estudios y solo tienen estudios primarios, mientras que existe un índice progresivo de estudios universitarios en los más jóvenes.

Se trata de un estudio sincrónico, que puede ser de utilidad en investigaciones que profundicen en aspectos de la dialectología, la terminología y la lexicografía de cada campo de especialidad.

BIBLIOGRAFÍA

ALVAR, M. (1979): *Atlas lingüístico y etnográfico de Aragón, Navarra y Rioja*. Tomo I. Zaragoza: Departamento de Geografía Lingüística. Institución Fernando el Católico de la Excma. Diputación Provincial de Zaragoza. CSIC.

ALVAR, M. (1995): *Atlas lingüístico de Castilla y León (tomo I).* Salamanca: Junta de Castilla y León.

Atlas Varilex. PHP-Atlas VARILEX. Disponible en: <http://lingua2.cc.sophia.ac.jp/varilex/php-atlas/lista3_search.php?> [Consulta: 17 de mayo de 2012]

BARTOL HERNÁNDEZ, J. A. (2004). *Léxico disponible de Soria*. Burgos: Fundación Instituto Castellano y Leonés de la Lengua. [Colección Beltenebros, nº 7].

BORREGO NIETO, J. (1981). *Sociolingüística rural. Investigación en Villadepeña de Sayago*. Salamanca: Universidad de Salamanca.

BORREGO NIETO, J. (1988). «Desajustes semántico-sintácticos y variación dialectal». *La Torre/Tercera época*. enero-junio, vol. 3(7-8): p. 89-97.

CABALLERO COTILLAS, M. Y. (2009). *Léxico pasiego burgalés*. [Tesis Doctoral inédita]. Valladolid: Universidad de Valladolid. Director: César Hernández Alonso.

DE LA CRUZ, J. (1961). «Algunas voces de Hornillayuso (Burgos)», *Revista de Dialectología y Tradiciones Populares,* XVII, 174-180 y 557-560.

GARCÍA MAMIERCA, Mª V. (2007). *Agricultura tradicional y moderna. El léxico Agrícola en La Bóveda de Toro (Zamora)*. Salamanca: Impreso en Kadmos S.C.L.

GARRABAU, R. y FERNÁNDEZ, J. (1985). *El fin de la agricultura tradicional*, vol. III. Barcelona: Crítica.

GIMENO FERNÁNDEZ, F. (1993). Dialectología y sociolingüística españolas. Alicante: Universidad de Alicante.

GOICOECHEA, C. (1961). *Vocabulario riojano*, Anejo VI del *Boletin de la Real Academia Española.* Madrid: RAE.

GOIG SOLER, Mª I. y GOIG SOLER, Mª L. (2000). *Diccionario de habla soriana*. Soria: Asociación Cultural SAAS/2.

GONZÁLEZ BACHILLER, F. (2002). *Hablar riojano. Comentario sobre voces riojanas*. Logroño: Radio Rioja y Caja Navarra.

GONZÁLEZ OLLÉ, F. (1964). *El habla de Bureba. Introducción al castellano actual de Burgos*. Madrid: CSIC. Instituto Miguel de Cervantes. *Revista de filología española*. Anejo LXXVIII.

HAENSCH, G., y HABERKAMP, G. (1987). *Diccionario de Agricultura*. Madrid: Ediciones Prensa-Mundi.

HERNÁNDEZ ALONSO, C. (Coord.) (2001). *Diccionario del castellano tradicional.* Valladolid: Ámbito.

HERRERA, G. A. de. (1994). *Agricultura general: que trata de la labranza del campo, y sus particularidades, crianza de animales, propiedades de las plantas que en ella se contienen, y virtudes provechosas a la salud humana.* Madrid: Fundación General de la Universidad Politécnica

HOYOS SANCHO, N. de (2006). *Refranero agrícola español*. Madrid: Ministerio de Agricultura.

IBÁÑEZ RODRÍGUEZ, M. y SÁNCHEZ NIETO, Mª T. (2006). *El lenguaje de la vid y el vino*. Valladolid: Universidad de Valladolid.

JIMÉNEZ GARCÍA, E. (2011). *Contribución al estudio del léxico de Soria* [Tesis doctoral]. Valladolid: Universidad de Valladolid. Directores: Manuel Ramiro Valderrama y Jesús María Bachiller Martínez. Disponible en: <http://uvadoc.uva.es/handle/10324/893> [Consulta: 17 de mayo de 2012]

MARTÍN CRIADO, M. (1999). *Vocabulario de la Ribera del Duero*, [Biblioteca 14 Estudio e Investigación]. Aranda de Duero (Burgos): Ayuntamiento de Aranda de Duero.

MINGOTE CALDERÓN, J. M. (1990). *Catálogo de aperos agrícolas del Museo del Pueblo Español*. Madrid: Ministerio de Agricultura, Pesca y Alimentación, Servicio de Extensión Agraria: Ministerio de Cultura, Dirección de Bellas Artes y Archivos.

MOLINER, M. (1996). *Diccionario de uso del español*. Madrid: Gredos.

REAL ACADEMIA ESPAÑOLA. Banco de datos (CREA) [en línea]. *Corpus de referencia del español actual.* Madrid: Real Academia Española. Disponible en: <http://www.rae.es> [Consulta: 16 y 17 de noviembre de 2010]

REAL ACADEMIA ESPAÑOLA. Diccionario de la lengua española [en línea]. Madrid: Real Academia Española. Disponible en: <http://buscon.rae.es/diccionario/drae.htm> [Consulta: 10 y 11 de noviembre de 2010]

SOLVEDILLA, M.C. (1994). *Contribución al estudio del léxico agrícola riojano.* Logroño: Gobierno de la Rioja, Instituto de Estudios Riojanos.

SOROA Y PINEDA, J. M. (Dir.) (1968). *Diccionario de agricultura.* Barcelona: Labor.

Anexo 1. Presencia de las palabras en las obras lexicográficas consultadas

	DRAE	DUE	DCT	DHS	CREA	NTLLE	VARILEX
acial	+	+	+	+	o	+	–
aguijada	+	+	+	+	+	+	–
arcial	+	–	+	–	–	+	–
arnero	–	–	–	+	+	–	–
arrejada	+	+	–	–	–	+	–
biela	o	o	+	–	o	o	–
bielo	–	–	–	–	–	–	–
bielda	+	+	+	+	–	+	–
bieldo	+	+	+	+	+	+	–
ceazo	–	–	+	+	–	–	–
cedazo	+	+	+	+	+	+	–
celemín	o	o	–	–	+	o	–
cernedor	+	+	+	+	+	+	–
colodra	+	+	+	+	o	+	–
criba	+	+	+	+	+	+	–
enjalmadera	–	–	–	+	–	–	–
estoba	–	–	–	+	–	–	–
estola	o	o	o	–	o	o	o
fanega	o	o	–	–	o	o	–
gavilán	+	+	o	–	o	+	–
harnero	+	+	+	+	–	+	–
hijada	–	–	–	+	o	o	–
hijón	–	–	–	+	–	–	–
horca	+	+	+	+	+	+	–
horcajo	+	+	+	–	–	+	–
horcón	+	+	+	+	–	+	–
horquilla	+	+	+	+	–	+	o
horquillo	–	–	+	+	–	–	–
laya	+	+	+	–	o	+	–
maceta	o	+	o	+	–	o	o
mayal	+	+	+	+	–	+	–
mazo	+	+	+	–	+	+	–
orca	o	o	–	–	o	o	–
rastoba	–	–	+	+	–	–	–
rastola	–	–	+	–	–	–	–
rastrillo	+	+	+	+	+	+	–
rejada	+	+	+	–	o	+	–
rejón	o	+	+	–	o	o	–
restoba	–	–	+	+	–	–	–
restola	–	–	+	–	–	–	–
torcedor	+	+	+	+	o	+	–
tralla	+	+	+	+	+	+	–
trillo	+	+	+	+	+	+	–
trillón	o	o	–	–	+	o	–
viela	–	–	–	–	–	–	–
zoqueta	+	+	+	+	o	+	–

TESTIMONIOS

LOS DISCURSOS DE LA PRIMAVERA ÁRABE, UNA APROXIMACIÓN LINGÜÍSTICA

Mohamed El-Madkouri Maataoui
Universidad Autónoma de Madrid

Resumen: Entre las particularidades de las distintas manifestaciones de la Primavera Árabe está la aparición no sólo de nuevos actores sociales, sino también de distintos discursos de movilización. La irrupción en escena del sexo femenino (de todo tipo de tendencias, orientaciones ideológicas y confesionales) obliga a la reconsideración de la imagen de la mujer como actor social de movilización. Con esta hipótesis intentaremos poner de relieve, desde el Análisis del Discurso, cómo se construye el discurso femenino de la movilización, cuáles son sus constantes y variables con respecto a las convenciones tradicionales del discurso árabe y, finalmente, cómo se percibe este tipo de discurso constructor de una nueva/antigua identidad árabe.

Para ello nos centraremos en el análisis de una breve proyección de una joven egipcia, Asmae Mahfouz, alojada en Youtube desde el 19 de enero de 2011 y con 59563 reproducciones desde entonces, en el momento de la redacción de estas líneas, y con más de 30.000 seguidores en Facebook y otros tantos en Twiter.

Palabras clave: Primavera árabe. Feminismo árabe. Mujer árabe. Análisis del Discurso.

Abstract: The Arab Spring has given rise not only to new social agents, but also to innovative mobilising discourses. The increasing role of women (from different social, ideological and religious backgrounds and positions) in Arab Spring events calls for a reconsideration of the image of women as mobilising social agents. Using a Critical Discourse Analysis approach, this paper analyses how Arab women's mobilising discourse is constructed, how its constant and variable features compare with the conventions of traditional Arabic discourse, and how it is perceived as a discourse (re)creating an old/new Arab identity.

The data consists of a brief videoclip by Asmae Mahfouz, a young Egyptian woman, which has attracted 59563 views, more than 30,000 followers on Facebook and many others on Twitter since it was posted on Youtube on 19 January, 2011.

Keywords: Arab Spring, Arab feminism, Arab women, Critical Discourse Analysis.

Las movilizaciones de los dos últimos años en los países árabes, especialmente en Túnez, Egipto y Yemen, han sorprendido tanto a los medios de comunicación, como a la Universidad y a los investigadores de numerosas zonas del mundo. En este caso la sorpresa es unánime y no afecta sólo al Mundo Occidental, sino también a las élites políticas e intelectuales de los propios países afectados. La sorpresa y la expectación iniciales se deben a que se trata de unas movilizaciones poco ortodoxas, que no se corresponden con la experiencia libresca acumulada sobre movilizaciones y revoluciones.

Entre las particularidades de las distintas manifestaciones de la Primavera Árabe está la aparición no sólo de nuevos actores sociales, sino también de distintos discursos de movilización. La irrupción en escena de la mujer (de todo tipo de tendencias, orientaciones ideológicas y confesionales) obliga a la reconsideración de la imagen de ésta como actor social de movilización. Con esta hipótesis intentaremos poner de relieve, desde el Análisis del Discurso, cómo se construye el discurso femenino de la movilización, cuáles son sus constantes y variables con relación a las convenciones tradicionales del discurso árabe y, finalmente, cómo se percibe este tipo de discurso constructor de una nueva/antigua identidad árabe.

Para ello nos centraremos en el análisis de una breve proyección de una joven egipcia, Asmae Mahfouz, alojada en Youtube desde el 19 de enero de 2011, con 59.563 reproducciones desde entonces, en el momento de la redacción de estas líneas, y con más de 30.000 seguidores en Facebook y otros tantos en Twiter. Reproducimos aquí la transcripción del vídeo analizado y comentado:

> Cuatro egipcios se han quemado a lo bonzo a causa del menosprecio, del hambre, de la pobreza y de la humillación en que venían viviendo desde hace treinta años (…). Cuatro egipcios, lo hicieron pensando que pudiera pasar igual que lo que pasó en Túnez… para que podamos convertirnos en un país libre, en un país con justicia, en un país con dignidad, en un país en el que el ser humano sea un ser humano de verdad… y para que la persona no siga viviendo como un animal… Hoy uno de ellos ha muerto y han anunciado su muer-

te. Yo bajé, encontré a toda la gente parada comentando "pobrecito… éste murió como un infiel [el suicido es ilícito en el Islam]… éste murió buscando la fama" ¡Pero, esto no se puede decir! Yo bajé (a la calle), escribí que soy una chica… que voy a bajar a la Plaza de Tahrir [lit. Plaza de la Liberación]… Lo haré sola y llevaré un letrero… puede que la gente sienta… escribí mi número para que la gente bajara… Nadie lo hizo excepto tres jóvenes… había tres patrullas, las fuerzas antidisturbios en la calle y decenas de *Baltaguiya* (matones) y oficiales… venían con extremo terror. Se quedaron hablando con nosotros y nos alejaron de la gente… Cuando nos apartaron, nos dijeron que ya bastaba… que ellos son como nosotros… que son del pueblo… que la gente que murió es porque tenían problemas psicológicos… Eso nos dijeron ¡Problemas psicológicos! ¡Ya es el colmo! Todos los medios informativos y periódicos del Gobierno dicen que cualquier persona que muere por la humillación y el asco de vida que llevamos, murió por trastornos psicológicos… Problemas psicológicos y se va delante del Parlamento.

Estoy grabando este vídeo para decir una sola cosa, para enviaros un solo mensaje… tenemos la intención de salir a la calle el día 25 [de enero]… Si todavía nos queda algo de dignidad… y con el deseo de vivir como personas y como seres humanos en este país… Si es así, debemos salir a la calle el día 25… saldremos a reivindicar nuestro derecho, nuestro derecho como seres humanos… No os digo a reivindicar nuestros derechos políticos, no se trata de ninguna reivindicación política… ni del Parlamento… ni de ninguna presidencia… ni de ninguna porquería de toda esa palabrería… Queremos nuestros derechos… no queremos otra cosa… Ese gobierno que se ha quedado obsoleto… el presidente es corrupto, las fuerzas de seguridad son corruptas… esa gente no temió a la muerte, pero teme a la política… tiene miedo de la corrupción… imaginaos… ¡Vosotros sois así! Os vais a matar solitos, ¿O no os dais cuenta de ello? Yo saldré el día 25… A partir de hoy y hasta el 25 bajaré a repartir panfletos a cada persona que encuentre en la calle… Yo no me voy a quemar a lo bonzo, si el gobierno lo quiere hacer que venga y me queme…

El que se considere hombre en este país que baje a la calle… el que diga que las chicas que salen a la calle se rebajan a sí mismas y que no deberían hacerlo y que

es ilícito, *Haram*… que se arme de hombría y de dignidad y que baje a la calle el día 25.

Cualquiera que diga que el número de los que están en la calle es muy pequeñito y que no va a cambiar nada, le digo que tu eres responsable de lo que nos pasa, sí, tu eres el responsable… y eres tan culpable como el Presidente… como cualquier corrupto… como cualquier oficial que nos pega y que nos maltrata. Tú eres el responsable… tu salida con nosotros es beneficiosa y no es poco… tu palabras dirigidas a gente de tu edad, a tus compañeros, amigos y familiares para animarles a salir a la calle… aunque salgáis en cualquier sitio, en cualquier calle o plaza, y no necesariamente en la Plaza del Tahrir, a hacer constar vuestra postura… y que digáis que sois personas…es suficiente. Que te quedes en tu casa mirando y viendo el Facebook y siguiendo las noticias, nos humilla a nosotros, me humilla a mí. Si tienes dignidad, eres persona y hombre en este país debes salir a la calle… salir para protegerme y proteger a cualquier chica que salga. Pero si prefieres quedarte en tu casa, te mereces todo lo que ocurre… y no solo tú… serás culpable y responsable… cargarás con la culpa de este país y de cualquier persona que viva en él… serás responsable de lo que ocurra a cualquier persona que haya salido a la calle para reivindicar sus derechos mientras tú estés sentado en tu casa. Sal de tu casa, envía mensajes y todo lo que grabes a tus amigos, escribe en Internet en cualquier sitio… ¡Conciencia!. Tú conoces tu círculo, con solo tu edificio, tus amigos, tus familiares, diles que te acompañen y salgan contigo… hazles bajar, a cinco o diez personas… Si cada uno de nosotros logra hacer salir a cinco o diez personas y llevarlos a la plaza de Tahrir, a la plaza de Lazogbi (sede de la policía política) o a cualquier sitio aunque sea anónimo y hablar con la gente diciéndoles que basta ya… y que en vez de quemarnos a nosotros mismos, hablemos, que hagamos constar nuestra postura será muy provechoso.

No digas nunca que no hay esperanza… Si tú dices que no hay esperanza, no la habrá. Pero si sales a la calle para hacer constar tu postura, habrá esperanza. Sé valiente y no tengas miedo del gobierno, ten miedo de Dios. Dios dice que Él no cambia a un pueblo mientras no se cambie a sí mismo. Y no es para que tú te quedes tranquilito sin dar palo al agua, (apoyado en la pared: es una expresión popular que significa no dar

> palo al agua. Aquí significaría no meterse en política). Si así te quedas, la pared acabará cayéndote encima… y nos aplastará a todos. Sal a la calle y reivindica tus derechos, mis derechos, los de tu familia y los derechos de todos nosotros. Yo saldré a la calle el día 25 [Pancarta: Saldré a la calle el 25 de enero por mi dignidad como egipcio]. Diré NO a la corrupción, y NO al régimen.

1. Singularidad de la Primavera árabe

Las movilizaciones de la *Primavera Árabe*, *Primavera Democrática*, *Revolución del Jazmín*, o cualquiera de las denominaciones con las que se conocen, suponen un caso revolucionario insólito por varias razones, entre las cuales se pueden citar las siguientes:

1.1 La juventud de los movilizadores

Lo que más sobresale en esta Primavera Árabe es la juventud de sus primeros inspiradores, tanto en Túnez como en Egipto. Los que entregaron sus vidas al fuego (en realidad sus cuerpos), inmolándose para incendiar la indignación y la lucha de estos países son todos jóvenes. Sin embargo, el fuego es sólo la llama, la gasolina estaba ya servida por varios lustros de opresión, animalización y, como consecuencia de ello, indignación. Es una Primavera Árabe joven en que varias generaciones de árabes lo han entregado todo, en términos de dignidad, tanto con el colonialismo, la mal llamada independencia y las recientes intervenciones y guerras en Oriente Medio. Era demasiada carga emocional la que albergaban las sociedades arabo-musulmanas, hasta que unos jóvenes –Bouazizi en Túnez y Abdou Abdel-Moneim Jaafar en Egipto- prendieron la mecha que incendió un fuego liberador que ni los regímenes locales ni sus apoyos extranjeros supieron cómo apagar.

1.2 La novedad aspectual de las manifestaciones

El terreno de la movilización y de la lucha es un terreno novedoso y sin datos acumulados de otras experiencias

históricas anteriores. Si bien se observa cierto parentesco con las sublevaciones populares árabes contra los ocupantes ingleses y franceses de finales del siglo XIX y principios del XX, la Primavera Árabe ha sorprendido porque ha sabido llevar al contrincante (las fuerzas de seguridad) a un terreno poco transitado por ellas. Así, los sistemas de comunicaciones son novedosos y variados, como se verá más tarde; al perder las poblaciones el miedo, la lucha se regenera automáticamente. De hecho, el principal factor es precisamente la pérdida del miedo, tanto a la tortura como a la muerte. Los castigos ejemplares, en vez de calmar la situación en favor del gobernador, lo que hacían era incendiar más los ánimos. Sin embargo, y a pesar del fuego instigador y del incendio de los ánimos, los protestantes no se dejaban caer en la provocación de las fuerzas de seguridad.

1.3 El discurso empático con las fuerzas de seguridad

Es clásico que el policía y el soldado sean considerados como los enemigos primarios, inmediatos y más cercanos a los revolucionarios. De hecho, los primeros en caer en las batallas campales por lo general suelen ser tanto los manifestantes como miembros de las fuerzas de seguridad. Sin embargo, dicho modelo no se ha reproducido con tanta intensidad ni se generalizó en las últimas revueltas del mundo árabe, por lo menos desde el lado de los manifestantes. En este sentido, los jóvenes árabes sublevados contra la tiranía de sus gobernantes, no consideraban a esas fuerzas como enemigas. De hecho, las consignas eran claras: los bienes deben conservarse porque si no son privados y pertenecen a sus dueños, son bienes comunes del Estado, es decir, de todos, y deben conservarse. Las cámaras de los aficionados no dejaban de transmitir imágenes de manifestantes que en vez de tirar piedras, ofrecían flores a los soldados agazapados a sus tanques y tanquetas. Sin embargo, y a pesar de ello, llama poderosamente la atención la postura tanto del ejército tunecino como del egipcio con respecto a las movilizaciones. La decisión de posicionarse del lado del pueblo, o con tan sólo mantenerse imparcial, no deja de sorprender en estos dos países que no sobresalen precisamente por su democracia en tiempos de Zine El Abidine Ben Ali y Hosni Mubarak.

1.4 El discurso pacífico de los manifestantes

Lo que más sobresale como inherente a la primavera árabe es el hecho de que sus movilizaciones fueron pacíficas. Los manifes-

tantes no se dejaban provocar a pesar de la irrupción en escena de los contra manifestantes parapoliciales y paramilitares, movilizados para crear el caos y dar más legitimación a la intervención policial y militar. Los jóvenes delincuentes, *Baltaguiya*, sacados de las cárceles y pagados por los regímenes para sabotear las manifestaciones, irrumpiendo incluso con caballos y camellos en la Plaza del Tahrir en Egipto, por ejemplo, no consiguieron ni asustar ni amedrentar a los manifestantes. De hecho, uno de los eslóganes más vitoreados tanto en Túnez como en Egipto fue: *¡Silmiya, Silmiya!,* ¡Pacífica, pacífica! El tipo de lucha que llega al seguidor de estos procesos es el desafío y la resistencia a los distintos tipos de violencia ejercida sobre los manifestantes.

1.5 El efecto sorpresa

La Primavera Árabe asombró al mundo, siendo los primeros en sorprenderse los regímenes árabes, acostumbrados como estaban y siguen estando aún algunos, a métodos clásicos de gestión de crisis. Al presentarse estas movilizaciones como ajenas a los manuales policiales y militares, especialmente a los organismos de información, han hecho que su canalización, gestión y resolución al interés del gobernante sea imposible. Ni siquiera los incondicionales aliados occidentales de los regímenes de Túnez o de Egipto, como son Estados Unidos y Francia, pudieron auxiliar a sus defendidos. De hecho, lo único que pudo hacer Francia por el régimen tunecino es lo que se sabe por la prensa: cargamentos de porras y gases lacrimógenos que no pudieron contener a poblaciones que ya habían perdido el miedo no sólo a las fuerzas de seguridad, sino también a la muerte. No debe olvidarse que estas manifestaciones se iniciaron por una inmolación.

1.6 El apoyo exterior

Al ser pacíficas, multitudinarias, resistentes y perseverantes, estas manifestaciones se han ganado no sólo la simpatía y la militancia de la diáspora árabe en el mundo, sino también la comprensión de la opinión pública internacional. De hecho, las tradicionales potencias legitimadoras de los regímenes tunecino y egipcio se quedaron sin argumentos ante la actitud no violenta de los manifestantes y ante la ausencia, en las manifestaciones, de lemas o

pancartas políticos propios de orientaciones políticas y partidistas de la región. Es el pueblo en su totalidad el que se levanta, ajeno a los partidos tradicionales y a los sindicatos, para revindicar su dignidad, palabra muy reiterada en estas manifestaciones. El apoyo de la opinión pública de muchos países occidentales,...

1.7 La aparente espontaneidad

Todas estas manifestaciones se han presentado como descabezadas: estas movilizaciones no tienen una cabeza visible ni una organización piramidal, como es el caso de las revoluciones tradicionales. Son sublevaciones populares que se visibilizan como populares y generalizadas, muy competitivas entre las distintas regiones y ciudades de cada país, pero sin ningún mando aparente, ni visible. Este hecho positivo lleva en sí la semilla de sus dificultades posteriores. El que estas movilizaciones no llevaran una cabeza visible, les ha permitido aguantar las embestidas tanto del régimen tunecino, como del egipcio. Sin embargo, esa espontaneidad y falta de una dirección ideológica o política (en el sentido clásico de los términos) ha dificultado la gestión de la etapa posterior a la caída de los totalitarismos combatidos.

1.8 Uso de medios de comunicación modernos

Aquí las juventudes han sabido sacar un provecho sustancial a las TICs. De hecho, el uso de los medios de comunicación basados esencialmente en las nuevas tecnologías es quizá la característica más destacada de las movilizaciones de la Primavera Árabe. Las juventudes árabes han sabido sobrepasar los medios de comunicación tradicionales, consistentes primordialmente en la televisión, la radio y la prensa escrita, para sortear el control de los gobiernos. En este sentido, podría decirse que los gobiernos, objeto de las protestas, gobernaban una juventud postmoderna con medios a lo sumo modernos, por no decir pre-modernos. La oligarquía tunecina y egipcia no fue capaz de entender y, sobre todo, reaccionar ante la noción del cambio del tiempo y del espacio. Hasta ahora, la conceptualización tanto del tiempo como del espacio

de acción eran clásicas. Los eventos y acontecimientos o se desarrollaban de día o de noche, y en una lugar concreto de un mapa físico representativo del país o de la ciudad. Para pasar información de una ciudad a otra y de un país a otro se necesitaba de las dos coordenadas, espacio y tiempo. Y esto facilitaba el control policial y militar de las poblaciones. Ahora, estas conceptualizaciones han cambiado: uno puede estar en distintos espacios (cibernéticos) a la vez, comunicarse (sin hacer ruido telefónico) a cualquier hora del día. Y esto es lo que hizo que la Primavera Árabe llevase la iniciativa informativa sobre los métodos terrenales y clásicos de los gobernantes locales. Las juventudes de todas las orientaciones y adscripciones políticas y confesionales han roto el cerco mediático de los gobiernos; han burlado su estrategias de silenciamiento informativo y han hecho llegar su voz más allá de las fronteras físicas tradicionales. Las juventudes árabes han sabido llevar la iniciativa de la información. De hecho, son informantes, más que informadas. Los gobiernos del Mundo árabe, sin excepción, han llevado un escrito control sobre la prensa desde prácticamente su independencia militar de algunas potencias europeas como Inglaterra, Francia o Italia. Es decir, que hubo y sigue habiendo (aunque con cierta relajación y con muchas diferencias entre países) cierto control sobre las publicaciones escritas. Al sustituirse el soporte de papel por la pantalla de los ordenadores, también se ha intentado bloquear algunas páginas web como hacía sistemáticamente el gobierno tunecino de Ben Ali. De lo que estos gobiernos no fueron conscientes es de que la comunicación también se hace por otros medios como la redes. Quizá interpretaron que Facebook es un *libro de rostros* donde la gente cuelga sus fotos y que el espacio virtual del mundo cibernético es un terreno de juegos adolescentes. Fue un error. No se puede gobernar una mentalidad postmoderna con métodos heredados de los colonizadores franceses e ingleses.

1.9 Organización y presentación de la información

No sólo han cambiado los medios y soportes, sino también la forma de la organización y presentación de la información y, sobre todo, que mucha información presentada es abierta. Es decir, que un artículo de prensa electrónica no es una

información cerrada con el único propósito de influir sobre el lector, sino que permite a éste reaccionar y publicar sus comentarios y puntos de vista. Así, esta suerte de democratización virtual no se corresponde con los que estos jóvenes y no tan jóvenes se encuentran cuando se apean del mundo cibernético. Había una contradicción que debía resolverse y es lo que intentaron.

1.10 La multifocalidad

Gracias a la nuevas tecnologías y a las redes sociales, además de las posibilidades comunicativas que ofrecen las nuevas tecnologías de los teléfonos móviles, se han podido organizar y convocar manifestaciones a lo largo y ancho de los países afectados y en varias zonas de la misma ciudad. Si aplicamos la metáfora del fuego que lo inflamó todo desde Túnez, podríamos decir que si las fuerzas de seguridad del gobernador podían controlar un fuego localizado en una zona determinada, no pudieron hacer lo mismo cuando los focos aparecieron a la vez en todas partes. Y esto es lo que ocurrió tanto en Túnez como en Egipto: hubo una competencia entre ciudades y barrios para demostrar quien protesta más. De hecho, en los primeros días de las manifestaciones tunecinas, los vecinos de una ciudad gritaban que dónde estaban los de la ciudad rival, que no se movilizaban. Pero junto con ello, hay determinados espacios que son simbólicos y que se han convertido en emblemáticos, como el Bulevar Bourguiba en Túnez o la Plaza del Tahrir (Liberación) en El Cairo. De hecho, esta plaza de la Liberación, nombre anterior a la Primavera Árabe, es el kilómetro cero del vigor de las manifestaciones en Egipto y es lo que reclama Asmae Mahfouz, objeto de esta investigación.

1.11 Nuevos actores

El factor humano es esencial y sigue siendo el motor del cambio y de la acción. El pensar, como se ha dicho en alguna que otra tertulia radiofónica española, que las revoluciones árabes las han llevado a cabo el Facebook y el Twiter es una trivialidad. Aquí, como en otras ocasiones, se da

protagonismo al medio (supuestamente nuestro) sobre el factor humano (el Otro). Todos estos medios son herramientas y es una de las primeras veces que han sido utilizadas por los jóvenes para algo más que subir fotos, canciones favoritas, películas y demás usos convencionales. De hecho, se ha hecho un uso poco convencional de esos medios. Junto con las redes sociales arriba mencionadas, y otras, se ha recurrido a otros medios como la telefonía móvil, los blogs, la prensa electrónica y la televisión por satélite. Aunque los gobiernos hayan bloqueado y dificultado el acceso a Internet, o perturbado las señales de las televisiones por satélite, se ha utilizado el teléfono móvil. Cabe señalar que el canal qatarí de *Al-Jazeera* se posicionó desde el principio de lado de los manifestantes y ha supuesto todo un cambio en la zona, especialmente en aquellos hogares alejados de los centros urbanos o sin medios para acceder a las nuevas tecnologías. Podría afirmarse en este sentido que tanto la postura de los ejércitos tunecino y egipcio como el papel mediático de Al-Jazeera han facilitado la transición y el cambio, sin demasiado derramamiento de sangre, como en el caso de Libia, Yemen o Siria. Cabe señalar con relación a la Primavera Árabe que los cuatro casos visibles: Túnez, Egipto, Yemen, Libia y Siria son tipológicamente distintos. Existen más similitudes entre Túnez y Egipto que entre estos dos y Libia y Siria por ejemplo. De hecho, podríamos decir que Túnez y Egipto firman un bloque, Libia y Siria otro, quedando Yemen como una bisagra entre los dos. El pacifismo del cambio iba decreciendo hasta que desemboco en los dramas humanitarios que estamos observando en Siria (noviembre de 2012).

1.12 La pérdida del miedo

La situación anímica del Mundo árabe ha llegado hasta el punto de entregar la vida cremándose. Los ciudadanos del Mundo árabe venían sufriendo todo tipo de vejaciones y humillaciones tanto desde el interior como desde el exterior. Además, como se ha dicho anteriormente la noción del espacio ha cambiado y todo se ha internacionalizado. Los jóvenes se ven como si estuvieran encerrados en corrales

cuando el mundo a su alrededor está desarrollándose. De hecho, el mundo árabe en general, venía interiorizando humillaciones y ofensas en su modo de vida, en su cultura y en sus creencias. Las élites intelectuales y políticas formadas a gusto de las ex potencias colonizadoras ya quedan obsoletas para las nuevas realidades. Han oprimido y vejado a sus pueblos y éstos ya han llegado al límite del aguante. De hecho, las palabras más usadas en las movilizaciones eran dignidad y humanidad. Además, como en el discurso de Asmae Mahfouz se equipara la situación actual con el posible riesgo que se pueda asumir y no encuentra diferencia alguna. Cuando afirma que si uno no sale a la calle quedándose arrimado a la pared, ésta acabará cayéndole encima. Con lo cual ya daba igual manifestar que no manifestarse porque ya no quedaba nada por lo que uno pudiera preocuparse. Esta pérdida del miedo ha sido señalada por la mayoría de los escritores que se han acercado al tema de la Primavera Árabe para estudiarlo y analizarlo. Así leemos por ejemplo en la contraportada del libro de Tahar Ben Jelloune (2011) que «millones de personas han perdido el miedo y se han lanzado a las calles para pedir libertad y democracia, dignidad e igualdad». En algunos casos no sólo se trata de perder el miedo, sino de salir a la calle para *entregar el cuerpo* y no volver a casa; y es eso precisamente lo que han hecho todas las víctimas del fuego y muchas personas anónimas que iban directamente en dirección a las balas de la policía, en vez de retroceder. Las veintenas de víctimas diarias de las manifestaciones de Túnez han entregado lo único que les quedaba, el cuerpo, una vez convencidos, tal y como afirma Asmae Mahfouz, de que el poder los ha animalizado desde hace treinta años, les ha despreciado, menospreciado y privado de dignidad. Esto mismo es lo que se denuncia en el vídeo, reclamando la consideración de los ciudadanos como personas (y no como animales) con plena dignidad.

1.13 El apoliticismo aparente

El aparente apoliticismo de los movimientos contestatarios y reivindicativos de la Primavera Árabe es quizá la característica más llamativa desde el punto de vista de las prácticas discursivas del lenguaje político tradicional en el mundo árabe. Contrariamente a lo que se ha observado en la movida política protes-

tante en Marruecos, las pancartas que ondeaban encima de las cabezas, eran las banderas nacionales de Túnez y Egipto. Hasta en esto han sabido evadir la posible deslegitimación, muy tópica, que les hubiese vinculado con una determinada orientación política o ideológica. La única baza que le quedaba al poder dominante era vincularles con el "espanta pájaros", con el cual se legitimaban ante los sectores de poder occidentales. De hecho, muchas dictaduras del mundo árabe adquieren cierta legitimidad y apoyo político occidental porque se presentan como un dique contra el islamismo, integrismo, fundamentalismo, terrorismo (y más denominaciones) que ellos mismos alentaban a veces para dar posibles pruebas "reales" de la amenaza contenida. Este argumento ha sido hábilmente esquivado por los pueblos tunecino y egipcio tanto en los símbolos ostentados como en los discursos usados. Aquí la construcción discursiva de la "amenaza" ficticia tradicional del discurso político de Mubarak o de Zine El Abidine Ben Ali en su alianza con el discurso mediático divulgador, tanto interior como exterior, no ha encontrado nada en la realidad en que puedan anclarse. Las revueltas no eran de ninguna confesión en concreto, ni de ningún color político determinado. Y cuando aparecen expresiones de tipo religioso como en el discurso de Asmae Mahfouz se plantean desde la cultura islámica del pueblo, y no desde un hipotético discurso militante islamista. De hecho, los islamistas, tanto egipcios como tunecinos, acosados quizá por los periódicos y los duros golpes recibidos durante los últimos 60 años, estaban inicialmente calladitos y como si estuvieran ajenos a las movilizaciones. Por lo que la represión contra unas poblaciones sin color político, sin banderas específicas de sus orientaciones ideológicas, confesionales o políticas, no podía legitimarse con el método clásico del pretexto de la contención del islamismo. Por ello, el discurso rebelde, aparentemente apolítico, ha dado un vigor y una legitimación tanto interior como exterior a las sublevaciones árabes en tanto que revoluciones populares sin adscripción política concreta.

1.14 La universalización de los derechos humanos

Otro caso relacionado también con los medios de comunicación y con la internacionalización de la información y que tuvo igualmente cierto impacto en el mundo árabe es el concepto de derechos humanos y democracia. Los ciuda-

danos árabes ya no pueden vivir en una isla de despotismo, rodeada de países que garantizan las libertadas fundamentales en donde las personas son consideradas ciudadanos de pleno derecho, cuya voz y voto son tenidos muy en cuenta y que, además, viven libres y dignos. A esto añadimos que existe una fuerte comunidad árabe y musulmana afincada en países que garantizan incluso una libertad religiosa musulmana que muchas veces, aún siendo oficial, está coaccionada en el mundo árabe. Hay más libertad para la barba y el pañuelo en el mundo occidental, salvo raras excepciones, que en muchos países modernos del mundo árabe. El sistema judicial democrático y su independencia del poder ejecutivo es algo que se percibe como un deseo en el mundo árabe. El ciudadano árabe tiene sed de justicia que ve, por los canales vía satélite, que existe en otras latitudes.

1.15 La colectivización y universalización de las protestas

Todas las razones mencionadas anteriormente han participado de un modo u otro en que las protestas de la Primavera Árabe fueran colectivas y universales. No se presentan como específicas de un grupo social sin el otro, ni de una franja de edad sin otra. Si bien es verdad que los primeros en movilizarse fueron los jóvenes, por el grado de maestría que tienen en el uso de las nueves tecnologías y por el ímpetu e inconformismo propio de esta edad, pronto lograron convencer a prácticamente todos los sectores de la sociedad. Todos han visto como madres y abuelas se acercaban a los soldados agazapados encima de sus carros de combate y tanquetas ofreciéndoles flores, lo que ha hecho que muchos miembros de las fuerzas de seguridad –generalmente de las clases más humildes de la sociedad- se solidarizaran con el pueblo, dejando sus armas.

2. La feminización de la protesta en el mundo árabe

Para entender este nuevo fenómeno simbólico en el Mundo árabe (y no árabe) cabe revisar el concepto de

héroe y luchador no sólo en este contexto, sino en prácticamente toda la cuenca del mediterráneo. El héroe es tradicionalmente un hombre, un gladiador, un Fernán González, un Mío Cid, un Carlomagno, un César, un Tarik; pero muy pocas veces una Cleopatra. En la cultura árabe clásica islámica, aunque se da cierta visibilidad a la mujer en los primeros días del Islam, la cultura árabe no tardaría mucho en reasignarle el papel que siempre ha ocupado en la mente de muchos árabes. Frente a la presencia, incluso guerrera, en los primeros tiempos del Islam (pensemos, entre otras, en Ashifaa Bent Abdelah Al Majzumiyya quien asumió, durante el califato de Omar ibn al-Khattaab, la responsabilidad de control del comercio interno, los impuestos y la calidad en los mercados, además era consejera personal del califa; en Umm Salamah, quien juega un rol decisivo durante el tratado de paz de Hudaibiya; o en Rufaidah bint Sa'ad y Amina bint Qais, quienes con una sólida formación técnica dirigieron equipos médicos en las primera batallas en época del Profeta Muhammad (Insgram 1983) o, muy posteriormente Shayarat ad-Durr (d. 1250), que se declaró sultana en Egipto después de la muerte de su esposo y de su hijo, haciendo frente a los cruzados),se impuso un silenciamiento de lo femenino hasta bien entrado el siglo XXI. Sin embargo, en la cultura árabe moderna, la mujer ha perdido este protagonismo que le dieron los primeros siglos del Islam. Su papel, aunque lo tuvo bien notorio en las resistencias a la ocupación como en Argelia y Marruecos, ha sido silenciado. No estamos hablando aquí de la mujer árabe del siglo XX que reivindica ciertos valores considerados –por sus opositores- como ajenos a su sociedad, sino de una nueva mujer con poder de cambio desde dentro de su propia sociedad y con un discurso que encaja perfectamente con los valores de su cultura. Por ello, nos ha sorprendido como seguidores del desarrollo de las movilizaciones árabes y nos ha llamado la atención como investigadores el papel de la mujer en estas movilizaciones. En este caso nos hemos centrado en exclusiva con un caso concreto que hemos tomado como ejemplo de varios otros, muy presentes en los nuevos medios tecnológicos de información. El que nos ha ocupado es el caso de Asmae Mahfouz que tuvo

mucho protagonismo en los primeros días de las movilizaciones egipcias. Esta joven veinteañera ha colgado una breve presentación en Youtube con un discurso, aparentemente espontáneo, pero sofisticadamente construido. Un discurso puramente árabe (sin ostentación de ningún préstamo ideológico o conceptual ajeno a la cultura en la cual se mueve), pero eso ha sido visualizado casi *medio millón* de veces sólo en Youtube, sin contar los hipotéticos envíos mediante cualquier programa de mensajería o las grabaciones de teléfono a teléfono mediante *bluetooth* u otros programas parecidos.

3 La construcción del discurso femenino.

El discurso de Asmae Mahfouz combina algunos aspectos discursivos que hacen de él un discurso sofisticado y eficiente.

3.1 Aspectos formales

Mahfouz utiliza expresiones sencillas, palabras cotidianas y tópicas, sin ningún rasgo formal que pueda inspirar que pertenece a una clase social determina, ni ningún tipo de ostentación de cultura o de grado de instrucción. Es un lenguaje social juvenil común que puede emplear cualquier persona de su edad. Las únicas palabras que quizá llamen la atención es: net, Facebook y otras, procedentes todas ellas del ámbito de la nuevas tecnologías. El tono usado es también el propio de una joven segura de sí misma, tan segura en su tono que transmite seguridad y hace sentirse incómodos a los inseguros.

3.2 Aspectos del contenido.

Esta joven militante ha construido temáticamente su discurso de tal modo que no se muestra a sí misma como heroína, ni más valiente que los destinatarios del discurso, sino que se presenta como voluntariosa y pide compañía y protección. De hecho, en este discurso hay un juego de valentía (ella no lo dice en ningún momento, pero

es algo que llega al oyente escuchándola), vulnerabilidad (soy chica y necesito, dentro del marco cultural árabe, protección), no quiero protagonismo, sino que entre todos vamos a hacer algo. De hecho, no sólo se aleja de cualquier pretensión de protagonismo, sino que da pistas a los destinatarios (jóvenes) de su mensaje sobre cómo ellos se pueden convertir en los verdaderos protagonistas (hablar con los amigos, compañeros, familiares y vecinos del inmueble; escribir en la Web, publicar grabaciones de vídeo, salir a protestar a cualquier calle, plaza y lugar, no quedarse en casa). Con respecto a la información que transmite de las fuerzas de seguridad, informa de que éstas les separaron de la gente (aparentemente sin violencia porque en el vídeo no se dice los contrario) llevándoles a un sitio apartado para explicarles que los que se habían quemado a sí mismos no estaban en lo cierto y que detrás de su fechoría no había ninguna reivindicación legítima.

3.3 La argumentación

Llama poderosamente la atención también la instrumentación que se hace en esta grabación de la noticia de los cuatro jóvenes egipcios que se han quemado a lo bonzo. Primero desmonta la explicación y contra-información oficial de que son jóvenes desequilibrados en busca de protagonismo y de fama. Ella argumenta que un desequilibrado buscador de la fama no hace lo que ha hecho, ni en el lugar emblemático en que lo ha hecho. Además, en un contexto en el cual todavía está vivo el recuerdo de Bouazizi que desencadenó la revolución tunecina, Mahfouz ha hecho que la versión oficial de las autoridades egipcias se presentara no solo como insostenible, sino también como patética. Sin embargo, sorprende igualmente el que ella no legitima la muerte a lo bonzo, sino que hace una interpretación negativa de este fenómeno: nosotros no nos vamos a quemar, si el gobierno quiere que nos quememos, que lo haga él. Este hecho, a su vez, vacía la justificación y la contra-información de los medios de comunicación e información oficiales del Egipto de Mubarak de todo su contenido. Es decir, si estos medios tuvieran preparado el argumentario para justificar cualquier muerte pasada o

futura de jóvenes egipcios, no se les va a dar la oportunidad de usarlo. Es como si un ejército estuviera armado y preparado para la lucha, pero el adversario no le hiciese el juego que se esperaba, quedándose el primero con su armamento sin usar. Este discurso no legitima, ni justifica, por tanto, la auto-cremación, sino que explica sus razones: los cuatro protagonistas de la inmolación o de su intento, los habían protagonizado, a juicio de la militante, por las siguientes causas:

3.3.1 El sentimiento de desprecio:

> El ciudadano egipcio (y por extensión árabe en general), simbolizado por los inmolados y caracterizado por el discurso analizado, se siente despreciado por el poder político. De hecho, a ojos de Asmae Mahfoud, el despotismo de los gobernantes ha rebajado a las personas a la categoría de animales. Es decir, que al ser humano egipcio ya no se le consideraba como persona bajo el régimen de Hosni Mubarak, por eso estos jóvenes se inmolaron para que la persona no siga viviendo como un animal en Egipto. De hecho, el desprecio del pueblo por los regímenes políticos árabes es un sentimiento generalizado entre los ciudadanos. Ciudadanos a los que, en determinados países como es el caso de las monarquías árabes, ni siquiera se les considera como tales… En el mejor caso no son más que súbditos. Algo que nos remite a la contradicción mencionada arriba, entre el tiempo "real" de los regímenes árabes y el tiempo tecnológico y "virtual" de sus poblaciones, especialmente los jóvenes.

3.3.2 El hambre y la pobreza:

> La falta de recursos de la mayoría de las poblaciones árabes incluidas, en algunos casos, las de las petromonarquías, la carestía de la vida y el paro son razones objetivas y suficientes para salir a la calle a protestar. Sin embargo, el hambre por sí sola nunca es una razón única para movilizarse y protestar (hay países en peores condiciones, pero no protestan, como algunos países africanos, surasiáticos y centroamericanos). Lo que agudiza el hambre es el sentimiento de injusticia y de

las abismales diferencias entre ricos y pobres. El ciudadano no siente el hambre como una realidad objetiva, sino como una situación inducida por otros. Es decir que uno es pobre porque el otro es rico por las arbitrariedades del poder ejecutivo, la corrupción del poder judicial, etc. El pobre protesta no por sentirse pobre, sino por convencerse de que le han empobrecido. La pobreza se esgrime como razón para protestar, como hemos observado con la Primavera Árabe, porque el que sale a la calle culpabiliza de su situación de pobreza al gobernador por favorecer que algunos sean ricos a costa de otros, desembocando así en la injusticia social. La gente, por ello, protesta no por sentirse pobre, sino empobrecida; es decir que percibe cierta injusticia (por redimir) en el hecho de ser pobre. No consiente que a él se le empobrezca para enriquecer a otros.

3.3.3 El sentimiento de humillación:

La palabra humillación junto con su campo léxico semántico es muy reiterada en boca de los manifestantes árabes. Así observamos que Mahfouz afirma que el ciudadano egipcio inmolado se siente humillado desde hace treinta años. Esto es lo que esta joven egipcia revela al afirmar que:

> Todos los medios informativos y periódicos del Gobierno dicen que cualquier persona que muere por la *humillación* y el *asco de vida* que llevamos, murió por trastornos psicológicos…

Con ello puede afirmarse, sin lugar a dudas, que el sentimiento de humillación de las poblaciones árabes es una razón suficiente para que salgan a la calle ¿Por qué se sentirían humilladas? Varias razones pueden barajarse:

1. Falta de sentimiento de ciudadanía.

Los ciudadanos, como se desprende del discurso de Mahfouz, no son considerados como ciudadanos, ni siquiera como personas.

> […] debemos salir a la calle el día 25… saldremos a reivindicar nuestro derecho, nuestro derecho como seres

> humanos… No os digo a reivindicar nuestros derechos políticos, no se trata de ninguna reivindicación política… ni del Parlamento… ni de ninguna presidencia… ni de ninguna porquería de toda esa palabrería…

Aquí se pide a los ciudadanos que reivindican lo más elemental que se les reconozca antes que ciudadanos, personas. Es como si estuviera diciendo a sus destinatarios que ni siquiera son ciudadanos, y que antes de reivindicar este derecho, deben ser primero personas a ojos de sus gobernantes y esto no sería posible sin las movilizaciones y las manifestaciones.

2. Falta de justicia social..

La reivindicación de la justicia social no es ninguna novedad introducida por los movimientos de la Primavera Árabe, sino que ha sido esgrimida desde prácticamente el día siguiente de la independencia de estos países, desde cuando descubrieron que, como en el caso de Marruecos, "los musulmanes de Rabat" eran peores que los franceses de los cuales se pensaba que se habían independizado.

3. Falta de respeto a la persona por parte de las instituciones públicas.

Los ciudadanos árabes no consideran que las instituciones públicas les respeten como ciudadanos, ni siquiera como personas. Esto lo que se desprende de los eslóganes vitoreados por los manifestantes, y es lo que hemos observado en el discurso analizado. El ciudadano árabe es ninguneado y silenciada su opinión. De hecho en Marruecos, por citar un ejemplo, los manifestantes del 20 de febrero pedían el alejamiento del entorno real a Mounir Majidi y Fouad Ali El Himma, secretario personal y fundador del PAM (Partido de la Autenticidad y de la Modernidad) y amigo personal del rey Mohamed VI, respectivamente; junto con otras figuras como el Clan Fassi Fihri, etc. Estas personas, consideradas corruptas por los manifestantes, no sólo no han sido apartadas de la cúpula del Estado, sino que se han metido más en Palacio al ser designadas como consejeros reales. De

todos modos, las poblaciones árabes han sido desde siempre severamente reprimidas y agredidas hasta que llegó el momento en que estas mismas poblaciones se agredieron a sí mismas de la peor forma posible (quemándose) para decirle al gobernador de forma simbólica que lo único que les quedaba era ese cuerpo que le entregan quemado, para que no sigan viviendo como animales. De hecho Asmae Mahfouz insiste en que la gente se inmola y protesta porque quieren vivir como personas dignas y no como animales:

> Para que podamos convertirnos en un país libre, en un país con justicia, en un país con dignidad, en un país en el que el ser humano sea un ser humano de verdad… y para que la persona no siga viviendo como un animal…

4. La humillación extranjera.

Frente a la humillación interior ejercida sistemáticamente por los gobernantes locales, viene a sumarse, y con más transparencia y claridad en las últimas décadas, la extranjera, especialmente la occidental. Los distintos embargos sobre la población civil en Irak, Sudán y Libia, la invasión del primero (como todo el mundo sabe en cuanto a violaciones, asesinatos indiscriminados, torturas y humillaciones en Abu Gharib), la escisión del segundo... Todo esto se ha vivido en el mundo árabe como actos de ultraje. A esto han venido a sumarse en los últimos años los ataques contra los símbolos de la cultura árabe-islámica, especialmente en materia religiosa: caricaturas, quemas de ejemplares del Corán, meadas de soldados americanos sobre cadáveres de musulmanes, ataques a los símbolos y representaciones religiosas islámicas en Guantánamo. La burla y humillación de los símbolos culturales y religiosos, sin que hubiese ninguna respuesta por parte de los poderes políticos oficiales de los países árabes, convierte, muy a menudo al gobernador y a las instituciones políticas árabes clásicas en cómplices.

Los pueblos árabes culpabilizan a sus gobernantes de todo lo que les ocurre y les acusan no sólo de cómplices, sino de ser puestos a la cabeza de los estados árabes por esas fuerzas extranjeras, precisamente. Además, todas las dudas sobre la connivencia y la corrupción de los gobernantes árabes han sido despejadas por Wikileaks. De hecho, se confirmó, por ejemplo, que tanto el estancamiento de la cuestión palestina como las distintas invasiones americanas y europeas del mundo árabe e islámico no hubiesen sido posibles sin la connivencia de los gobernantes árabes como Hosni Mubarak, por ejemplo. Sin embargo, más allá de Wikileaks, el ciudadano árabe es consciente de su historia reciente. Observemos por ejemplo este mensaje escrito a modo de comentario a una noticia, aparecido en el periódico digital marroquí Lakome.com el 28 de noviembre de 2012 con el título de *El hijo de Allal ben Abdellah vive con una pensión inferior a 600 Dirhams al mes (- de 60 Euros).* Dice el comentarista anónimo*: Lokome.com 28.11.12 // 6# comentario: el hijo de Allal ben Abdellah vive con una pensión de menos de 600 Dirhams (60 Euros) al mes. Madarach.*

> En aplicación de la política majzeniana de la economía de la renta, legalizada a medida, siguiendo los antojos de los que detentan las riendas y capacidades de la gente y del país según su particular resistencia a la ocupación, el hijo del mártir y luchador Allal Ben Abdellah debe buscarse la vida porque al Majzen la colonización francesa no le entregó el poder, sino después de comprometerse a recompensar a los traidores y a sus hijos (y por consiguiente castigar a los resistentes, patriotas y a sus hijos y nietos). Si alguien no se lo cree, que eche un vistazo a la lista de los cercanos [a las jerarquías superiores (afortunados y agraciados con las riquezas del país) encontrará que la mayoría son traidores o sus hijos y familiares.
>
> Pero el hijo de nuestro luchador que entregó la vida por el país, tendrá que aguantarse e invocar a Dios para que libere a Marruecos y a los marroquíes de las zarpas de las águilas a las cuales la colonización francesa confió tanto la continuidad de la ocupación de Marrue-

> cos como la vigilancia de sus intereses y los de los traidores colaboracionistas con ella… No hace falta citar nombres aquí porque todo el mundo les conoce… No os olvidéis de que la ocupación sigue apoderándose de Marruecos y de los marroquíes, porque nunca salió.

De lo anterior se desprende que son varias las causas que pueden sacar a las poblaciones árabes a la calle para reivindicar el cambio e impulsar una nueva dinámica social en estas sociedades. Lo que denuncia Asmae Mahfouz en su vídeo no es todo lo que se ha intentado explicar anteriormente, sino sus consecuencias. De hecho, todo se resume en dos conceptos: dignidad y humanización:

> Estoy grabando este vídeo para decir una sola cosa, para enviaros un solo mensaje… tenemos la intención de salir a la calle el día 25 [de enero]… Si todavía nos queda algo de dignidad… y con el deseo de vivir como personas y como seres humanos en este país… Si es así, debemos salir a la calle el día 25… saldremos a reivindicar nuestro derecho, nuestro derecho como seres humanos…

Es decir que este discurso quiere marcar una frontera entre dos periodos en la historia actual de Egipto: uno con Mubarak (supuestamente) en que al egipcio se le hecho perder la dignidad y se le ha condenado a vivir en condiciones infrahumanas (como animal en el discurso); y otro periodo en el que se aboga por una mejora de dichas condiciones ¿Qué lo que reclama Mahfouz, entonces? Las proclamas más reiterativas son: la libertad, la justicia, la dignidad y que el ser humano sea considerado como tal, y no como una animal.

3.4 Construcción del discurso de la movilización:

3.4.1 El uso de la feminidad en el discurso.

Mahfouz es consciente de la imagen de la mujer en la sociedad egipcia y de los roles sociales asignados tanto al sexo femenino como al masculino. Y son estos roles precisamente lo que esta militante instrumentaliza en la construcción de su discurso para llegar a su auditorio e influir sobre él.

a) Desconstrucción del modelo del héroe árabe:

La cultura egipcia, y árabe en general, asigna el rol de defensor tanto del territorio como de los espacios simbólicos del entorno familiar al varón. Éste tiene la obligación moral de entregar la vida para defender su tierra, su honor y su familia. La sociedad no asigna esta obligatoriedad defensiva a la mujer. De hecho, la literatura árabe, tanto preislámica como islámica, está llena de casos representativos de este ideal simbólico. Esta obligatoriedad moral ha hecho que el varón árabe –en muchas ocasiones- convierta la obligación en derecho, instaurándose simbólicamente como un gladiador ficticio y verbal, sin asunción alguna de la responsabilidad defensiva que su cultura le asigna. El varón árabe debe ser protector, sin embargo, observamos con el discurso de Asmae Mahfouz, que ni siquiera puede protege su propia dignidad. Aún así, rechaza y reprocha a la mujer que salga a la calle.

b) Cuestionamiento del varón protector:

El varón árabe tiene la obligación moral y cultural de defender a la mujer. Es decir que antes de verla a ella exponerse al peligro hace entrega de su propio cuerpo, si hiciera falta, un escudo de protección. Sin embargo, no es lo que observamos en la proyección analizada. El "héroe" egipcio no sólo no hace ningún esfuerzo para defender a la mujer, sino que ni siquiera defiende su propia dignidad. Aún más, reprocha a la mujer el salir a la calle para defender la dignidad (palabra muy repetida) de todos tanto del hombre como de la mujer e intentar que los poderes políticos consideren al hombre no como un animal, sino como una persona. Asmae, mujer, afirma en su video "que escribió (en la web) que es una chica,… que iba a bajar a la Plaza de Tahrir, que lo haré sola llevando un letrero".

El adjetivo *sola* es enigmático… porque significa a la vez sin compañía de más manifestantes, pero también –y he aquí la crítica mordaz- sin compañía y sin protección. Tanto en un caso como en otro, lo silen-

ciado es que el varón, quien debe estar en el Tahrir, se ha desembarazado de su rol social dejando que su protegida asuma el papel no solo de su propia protección, sino de la lucha por la libertad de todos, mujeres y hombres. La crítica no podía ser más severa y más explícita: "el que se considere hombre en este país que baje a la calle"

c) Reavivación de modelos femenino legitimadores:

Aunque la mujer tiene cierta relevancia en la literatura teológica islámica clásica, esta visibilidad de la mujer como persona activa en la sociedad no tardó mucho en "corregirse" sometiéndose nuevamente a los avatares de la cultura beduina árabe preislámica. De hecho, el papel de la mujer activa a penas pudo sobrevivir un siglo tras la muerte del Profeta del Islam. Las costumbres árabes pudieron aclimatar y "localizar" el Islam para compatibilizarlo con las tradiciones árabes. Lo que se hizo, con respecto a la mujer, es elevar la moral beduina y tribal a la altura de la religión creando una amalgama entre el universalismo religioso y el localismo tribunal. De hecho, se lee muy a menudo en los textos biográficos e históricos de los primeros tiempos de los musulmanes y del Islam que la mujer era activa no solo en el ámbito civil, sino también en el militar: enfermera de guerra en el frente, agente de la logística, retaguardia,… Desde los Omeyas, la mujer volvió a visibilizarse poco a poco a pesar de ciertos nombres relucientes a lo largo y ancho del mundo árabe como Fatima Fihriya, Sayyida Hurra en Marruecos, entre otras. Pero estos casos, se presentan como la excepción que confirma la regla. A este respecto, la voz femenina de Asmae Mahfouz, entre otras, sacude la conciencia varonil árabe. No se trata aquí de voces portadoras de un discurso que podría considerarse como ajeno y de importación desde occidente (Fatima El Mernissi, Naoual Saadaoui….), sino de una mujer que habla desde el paradigma cultural árabe, y por tanto, difícilmente exteriorizada (rechazada…):

> El que diga que las chicas que salen a la calle se rebajan a sí mismas y que no deberían hacerlo y que es ilícito, *Haram,* que se arme de hombría y de dignidad y que baje a la calle el día 25.

Más claro y más directo no podía ser este discurso dirigido a los hombres. La oradora utiliza las mismas bases y convenciones culturales para criticar la dejadez del hombre árabe y para animarlo a movilizarse.

Conclusión

El discurso femenino analizado revela que su autora ha partido desde los postulados y convenciones de su propia cultura para construir y articular un discurso encaminado a movilizar a sus compatriotas para el cambio del régimen político. El discurso refleja el dominio del aparato argumentario con el cual el poder político egipcio de Hosni Mubarak deslegitima a sus adversarios. Por eso esta militante elude sabiamente a las formulaciones clásicas de la formulación de la reivindicaciones para no dejar pistas a sus adversarios. Mahfouz no enmarca su discurso dentro de ningún color político o ideológico, presentándose como parte del pueblo egipcio en general, sin distinción alguna. Las reivindicaciones discursivas de esta mujer se distancian mucho de las demandas partidistas y sindicalistas, haciendo hincapié en lo más elemental y básico para las personas: la dignidad y su consideración como personas. Es decir, que reclama que se deje de animalizar a la persona egipcia y que, en cambio, se la considere como un ser humano. Con ello elude las demandas tradicionales muy trivializadas:

> Saldremos a reivindicar nuestro derecho, nuestro derecho como seres humanos… No os digo a reivindicar nuestros derechos políticos, no se trata de ninguna reivindicación política… ni del Parlamento… ni de ninguna presidencia… ni de ninguna porquería, *Zift*, de toda esa palabrería…

En este sentido, el discurso político anterior a la Primavera Árabe no era, a ojos de las nuevas juventudes, más que palabrería y lo que se reclama ahora no son derechos políticos, sino algo más básico como la dignidad de la persona y la consideración de los egipcios como seres humanos. Es un nuevo discurso para unas reivindicaciones básicas y asumibles por todos los egipcios. Para hacerlas efectivas y movilizar a los egipcios, Asmae Mahfouz procede, entre otras estrategias, mediante la subversión del discurso masculino acostumbrado como estaba a relegar a la voz femenina a un segundo plano. Para ello, Mahfouz denuncia que el varón falla en el cumplimento de sus responsabilidades y falla en su rol cultural y social. Al ser así, la mujer anuncia que ella misma tomará la iniciativa y que la responsabilidad moral de lo que le ocurriera será imputable al varón. Éste ya no tiene ningún derecho como varón ya que ha demostrado que no cumple con su responsabilidad, y por consiguiente no tiene derecho a reprocharle a la mujer que se expone al peligro manifestándose en la calle.

Observamos igualmente en este discurso, igual que la subversión del discurso masculino, una subversión del discurso político tradicional. Asmae Mahfouz no plantea sus reivindicaciones en términos de política tradicional: igualdad, derechos políticos y laborales, participación y transparencia, sino que plantea algo más básico y universal: dignidad, humanización, justicia y libertad. Estas nuevas reivindicaciones llegan directamente al egipcio medio en boca de una joven firme, valiente y segura de sí misma. Una mujer que asume por sí misma el rol que la cultura ha otorgado tradicionalmente al varón, pero que éste, por desidia o por miedo, ha dejado de ejercer. Por eso le espeta: “Sal a la calle y reivindica tus derechos, mis derechos, los de tu familia y los derechos de todos nosotros”. En este sentido, Asmae Mahfouz no es una excepción, sino que observamos una participación activa de la joven mujer árabe en las movilizaciones de Túnez, Yemen, Siria, Marruecos,…

Bibliografía

Anónimo: "Las más famosas musulmanas a lo largo de la historia", en http://www.eternalegypt.org/EternalEgyptWebsiteWeb/HomeServlet?ee_website_action_key=action.display.module&language_id=3&story_id=21&module_id=158&ee_messages=0001.flashrequired.text (2.12.2012)

Bechir Ayari M., Geisser V. (2011): Renaissances Arabes: 7 questions clés sur des révolutions en marche. Paris: Éditions de l'Atelier.

Ben Jelloune, T. (2011): *Primavera árabe: el despertar de la dignidad*. Barcelona: Alianza Editorial.

Insgram, D. (1983): *The Awakened: Women in Iraq*. El Libano: Third World Centre for Research and Publishing Ltd, .

Mehmet Caner, E. (2005): *Voces detrás del velo: el mundo del islam y la mujer musulmana. Michigan: Editorial Portavoz.*

Morales Lezcano, V. (2012): *Norte de África: rebeliones sociales y opciones políticas*. Madrid: Diwan.

Núñez Sarmiento, L. P (2012)*: Túnez a la luz de la primavera árabe: Incidencia de las identidades transnacionales en las comunidades de origen.* Alemania: EAE.

Urdiales Viedma, M. E. (2007): *Geopolítica y desigualidades*. Granada: Editorial Universidad de Granada.

Zaccara L. (2012). *"*Elecciones en el Golfo: resistir `la primavera árabe*'"*, en *Política exterior,* Vol. 26, Nº 148. 2.

NOTAS

[1] Esta contribución se enmarca en los resultados del proyecto de investigación I+D "Juventudes, cambio social, política y sociedades en red en el Mediterráneo: el caso del Magreb", financiado por el Ministerio de Ciencia e Innovación (CSO2011-29438-C05-04).

DOCUMENTOS

CONFERENCIAS DE AMÉRICO CASTRO EN CHILE: 1923

Francisco A. Marcos Marín
University of Texas at San Antonio

La Universidad de Buenos Aires creó el 21 de junio de 1922 un Instituto de Lingüística, dependiente de la Facultad de Filosofía y Letras. Aunque ése fuera su nombre oficial, el que se impuso en la práctica fue el de "Instituto de Filología". El decano de la Facultad era a la sazón Ricardo Rojas, quien preparó el proyecto junto con Emilio Ravignani. De este modo se introdujeron la filología y la lingüística en el ámbito académico universitario argentino, siguiendo un modelo en el que Rojas y Ramón Menéndez Pidal estaban de acuerdo y en el que el modelo era el Centro de Estudios Históricos de Madrid, dirigido desde 1910 por Menéndez Pidal. Esta elección de modelo se debe en buena medida a la acción de la Institución Cultural Española, que apoyaba, desde 1914, las ideas de la Junta para Ampliación de Estudios e Investigaciones Científicas de España.

En 1923 Don Ramón propuso a Américo Castro como director del Instituto, en una dirección alternada. La idea original era que en 1924 el director fuese Tomás Navarro Tomás y que en 1925 regresara don Américo a la Argentina. En realidad, el director en 1924 fue Agustín Millares Carlo, seguido por Manuel de Montolíu en 1925. El regreso de Castro a la Argentina se produjo tras la guerra española, pero ya no al Instituto, bien dirigido desde 1927 por Amado Alonso, y sólo por un breve tiempo, antes de la marcha definitiva a la universidad norteamericana. A Amado Alonso se debe, por cierto, el cambio oficial del nombre a "Instituto de Filología", en 1940.

Ese año de 1923 Américo Castro desarrolló en la Argentina una enorme tarea. Para conocer lo que pensaba entonces hay que en cuenta las cartas de éste a la dirección del Centro de Estudios Históricos publicadas por Fernando Degiovanni y Guillermo Toscano y García (2010). La postura de Castro era muy delicada, entre la situación argentina, que incluía la amabilísima reacción del público,

y la discrepancia entre la dirección de la Universidad de Buenos Aires y el Centro de Estudios Históricos de Madrid. Esa tarea, muy en sintonía con la realidad porteña, ha quedado oscurecida, en parte e injustamente, por la reacción antiespañolista, muy posterior a esa época, de J. L. Borges a *La peculiaridad lingüística rioplatense y su sentido histórico,* un libro de 1941. Lo que aquí importa es que, dentro de la actividad del Instituto en Buenos Aires estaban las conferencias multitudinarias que pronunciaba su director. Estas conferencias no se limitaron a la Argentina.

A fines de 1923, don Américo se desplazó a Chile, invitado por la Universidad de Chile. Allí impartió siete conferencias, destinadas a profesores y estudiantes de "Castellano" en el Instituto Pedagógico:

1 Épocas principales de la historia de la lengua española.
2 Influencia del Renacimiento en la evolución de la lengua española.
3 "La Celestina" de Fernando de Rojas, como representación del concepto renacentista de la vida.
4 Lope de Vega.- Su vida.
5 Lope de Vega.- Sus obras.
6 Cervantes.- Su filosofía de la naturaleza y su técnica literaria.
7 Metodología de la enseñanza de la lengua y literatura españolas.

Las conferencias se publicaron en los *Anales de la Universidad de Chile* (ISSN: 0365-7779). La Vicerrectoría de Extensión y Comunicaciones de la Universidad las ha puesto a disposición del público en su portal de internet, un sitio web desarrollado por SISIB como una adaptación de *Open Journal Systems.* En los índices de búsqueda, adviértase, la conferencia sobre "La Celestina" aparece con la autoría de Fernando de Rojas y no de Américo Castro, lo que puede haber dificultado su localización. Se trata de estudios de gran interés, porque corresponden a la etapa filológica de Castro, en la que centra su investigación en los textos. Las dos primeras permiten comprender cómo evolucionaba la *Historia de la Lengua* que elaboraba Ramón Menéndez Pidal y que nunca terminó, y anticipa futuras publicaciones de don Américo que fueron esencia-

les en la interpretación de Cervantes y Lope. En el caso de la conferencia metodológica, ya se habían publicado sus libritos esenciales sobre la reforma de la enseñanza de la lengua y la literatura españolas, por lo que importa es la difusión de sus tesis en un medio tan influyente en el campo de la lengua española de la época como el chileno. Don Américo aprovechó también para dar noticia de las colecciones de textos que se estaban publicando entonces en España, especialmente los clásicos de *La Lectura,* que más tarde serían *Clásicos Castellanos.* En esta colección aparecieron ediciones castrianas notables de textos teatrales del Siglo de Oro, así como de varias obras clásicas por otros miembros distinguidos del Centro de Estudios Históricos que dirigía Menéndez Pidal en Madrid. La publicación se basa en transcripciones de las conferencias, por lo que se pueden encontrar algunos errores en la interpretación de nombres propios, desconocidos para el transcriptor o confundidos por éste y algunos otros desajustes menores. En general, constituyen un ejemplo de claridad expositiva y desarrollo de una presentación oral con admirable dominio de los recursos oratorios. Para los estudiosos de la ortografía, tienen además la peculiaridad de estar escritas en la ortografía chilena modernizada de Andrés Bello, oficial en Chile desde 1844, que en 1927 fue reemplazada por la ortografía académica general, según la Real Academia Española.

Parece útil, por ello, dar la referencia a las páginas de internet de donde pueden descargarse las transcripciones de las conferencias, para incrementar el conocimiento de esta etapa de la filología española.

Castro, A.. Conferencias del filólogo español *Anales de la Universidad de Chile*, 0 7 12 2012. Consultado el jun 18, 2013, de http://www.revistas.uchile.cl/index.php/ANUC/article/viewArticle/25318/28137

Rojas, F.. "La Celestina" de Fernando Rojas, como representación del concepto renacentista de la vida *Anales de la Universidad de Chile*, 0 7 12 2012. Consultado el jun 18, 2013, de http://www.revistas.uchile.cl/index.php/ANUC/article/viewArticle/25319/28138

Castro, A.. Conferencias del filólogo español Don Américo Castro. Lope de Vega: Su vida *Anales de la Universidad de*

Chile, 010 12 2012. Consultado el jun 18, 2013, de http://www.revistas.uchile.cl/index.php/ANUC/article/viewArticle/25532/28147

Castro, A.. Conferencias del filólogo español Don Américo Castro. Cervantes: Su filosofía de la naturaleza i su técnica literaria *Anales de la Universidad de Chile*, 010 12 2012. Consultado el jun 18, 2013, de http://www.revistas.uchile.cl/index.php/ANUC/article/viewArticle/25534/28148

Castro, A.. Metodolojía de la enseñanza de la lengua i literatura españolas *Anales de la Universidad de Chile*, 010 12 2012. Consultado el jun 18, 2013, de http://www.revistas.uchile.cl/index.php/ANUC/article/viewArticle/25536/28149

REFERENCIAS

CASTRO, Américo (1924) *Conferencias dadas en el salón de honor de la Universidad [de Chile] en 1923 por el profesor de filolojía de la Universidad Central de Madrid, don Américo Castro,* Santiago de Chile: Soc. Imprenta i Litografía Universo.

DEGIOVANNI, Fernando y Guillermo TOSCANO Y GARCÍA (2010) "Disputas de origen: Américo Castro y la institucionalización de la Filología en Argentina", *Nueva Revista de Filología Hispánica,* 58, 1, 2010, 191-213.

IN MEMORIAM

SAMUEL ARMISTEAD (1927-2013)

Francisco A. Marcos Marín
University of Texas at San Antonio

Mis primeras y vagas imágenes de Sam remontan a finales de los años sesenta, en el seminario de don Rafael Lapesa en la Universidad Complutense, como una figura que pasó por allí y con quien entonces no hablé; pero el primer recuerdo nítido está relacionado con el volumen de homenaje dedicado en Madrid por Ediciones Taurus a Américo Castro, en 1971. Después del acto, ambos nos interesamos simultáneamente por el ejemplar sobrante de la edición y Jesús Aguirre aprovechó la circunstancia para presentarnos. Como es natural, Sam se llevó el libro. Desde entonces mantuvimos una cordialísima relación, con intercambio de publicaciones, coincidencias en congresos, seminarios en Madrid, reuniones con Diego Catalán o visitas mías a Davis o suyas a Berkeley, con la común amistad de Arthur Askins o Charles Faulhaber. Especialmente en los últimos años las coincidencias presenciales eran más complicadas, por su horario completamente nocturno; pero nos las arreglamos para mantener contacto telefónico.

La pérdida que lamentamos hoy es la de uno de los más destacados investigadores de la transmisión tradicional de la lengua y la literatura españolas, tanto en las versiones cristianas como judaicas, autor de treinta libros y quinientos artículos, en números redondos. Su interés arranca, muy en la línea neotradicionalista de Ramón Menéndez Pidal, de las crónicas medievales y su relación con la épica y el romancero. Recuerdo alguna llamada telefónica para expresarme, muy amablemente, su discrepancia porque en mi edición del *Cantar de Mio Cid* no había seguido decididamente alguna interpolación de don Ramón, procedente de las crónicas. Tuve que explicarle que mis convicciones sobre el neotradicionalismo seguían siendo las mismas y que se trataba sólo de una exigencia del tipo de edición; pero no lo convencí. Para él la continuidad entre historia, épica y romance estaba más que probada y no se debía dejar resquicio a ninguna hipotética discrepancia.

Siendo sobre todo medievalista, su condición de folclorista tenía que llevarlo directamente a interesarse por los restos de variantes lingüísticas españolas en los Estados Unidos, especialmente las mejor conservadas, las canarias de los isleños de Luisiana. Sus trabajos y entrevistas (muchas todavía inéditas) son decisivos a la hora de conocer la historia de esas hablas, salvándolas de su lento proceso de desaparición, que fenómenos naturales como el huracán Katrina de 2005 tanto han agravado.

En la realidad de las encuestas cuenta mucho la empatía entre encuestador y encuestado. El cuestionario es necesario; pero limita y coarta. La conversación con Sam, su capacidad de provocación de los interlocutores, hacía aflorar historias, palabras, expresiones, canciones que los propios interlocutores creían olvidadas o que su timidez inicial les había impedido presentar. Este proceso ha resultado fundamental no sólo para la tradición peninsular o canaria, en España y América, sino muy especialmente para la judeo-española. Además de preservar los textos sefardíes, su aportación ha sido valiosísima para rescatar la música. Era frecuente que incluyese en sus clases o presentaciones fragmentos entonados por él mismo, una tradición con la que me siento muy gratamente identificado y que, en mi opinión, nada objetiva, los alumnos aprecian.

Discípulo de Américo Castro, heredó directamente su vinculación a la escuela española de Filología y la llevó a cimas magistrales. Correspondiente de la Real Academia Española y doctor honoris causa por la Universidad de Alcalá de Henares, recibió también, entre otros muchos premios, el de Antonio de Nebrija. Y, sobre todo, será recordado mientras vivamos por quienes tuvimos el honor de ser sus amigos y nos consideramos (aunque él pondría alguna de sus típicas objeciones) sus discípulos. ¿Qué nos vincula y qué permanece? Nos vinculan nombres, de personas e instituciones: Menéndez Pidal, Castro, Lapesa, Catalán, Silverman, Faulhaber, Askins. Nos une también la convicción de que la poesía y las tradiciones permanecen, a veces en primer plano, a veces latentes, de que el pueblo tiene interés por conservar y el proceso de conservación, que puede ser extraordinariamente fiel,

puede tener una versión creativa. Cuando encuentro romances españoles en el sur de Tejas, en Nuevo México o en el norte de México tengo que percibir su transformación hacia el corrido; pero es innegable el elemento que subsiste. Este elemento va más allá del núcleo argumental, para incluir aspectos retóricos tradicionales que siguen firmes. Sevilla puede haber dado paso a una localidad mexicana; pero las hijas siguen siendo siete, sin hermano varón y sigue siendo la pequeña la que quiere ir a la guerra vestida de hombre, igual con la casada infiel, o con Gerineldo. Cuando, en medio de una conversación en la que tratamos de hacer aflorar un viejo texto, uno de los interlocutores (no siempre mujer) se arranca con unos versos de romance, con una música anterior a la presencia de los europeos en este continente, se siente ese escalofrío delicioso por la espalda y se refuerza la conciencia de lo que debemos a maestros que dieron el ejemplo de sus vidas.

ÍNDICE

ARTÍCULOS

Leonardo Campillos Llanos, «Análisis de la producción y de errores en un corpus oral de español como lengua extranjera»..... 5

Antonio Moreno Sandoval y Marta Garrote Salazar, «La anotación de la negación en un corpus escrito etiquetado sintácticamente»..... 45

María Isabel Maldonado García, «Estudio etimológico de cuatro pares de cognados en español y urdu»..... 61

José Alberto Miranda Poza, «Español y portugués en contraste: falsos amigos, historia de la lengua, campo léxico, semántica de prototipos»..... 75

Elena Jiménez García, «Léxico sobre utensilios de labranza de cereal en la provincia de Soria»..... 119

TESTIMONIOS

Mohamed El-Madkouri Maataoui, «Los discursos de la Primavera Árabe, una aproximación lingüística»..... 133

DOCUMENTOS

Francisco A. Marcos Marín, «Conferencias de Américo Castro en Chile: 1923. 165

IN MEMORIAM

Francisco A. Marcos Marín, «Samuel Armistead (1927-2013)»..... 173

Revista Iberoamericana de Lingüística (R.I.L.) tiene como principal objetivo ofrecer las contribuciones y los diferentes estudios científicos en torno a los diversos aspectos de la lingüística en el mundo iberoamericano o desde la perspectiva de los autores que trabajan en ese entorno científico. R.I.L. se publica anualmente. Toda la correspondencia debe dirigirse a la dirección postal *Universitas Castellae* (R.I.L.), Apartado 37, 47080 Valladolid, España. En todo caso es preferible el envío por correo electrónico al editor de la publicación en España (Ricardo de la Fuente) ricardodelaf@hotmail.com. En el correo electrónico, así como en el archivo que se adjunte, deben constar todos los datos del autor: afiliación universitaria y dirección postal y electrónica. Los trabajos que se envíen no deben superar los 30 ff.a doble espacio (incluidas notas y bibliografía), o 7000 palabras. El sistema de referencias debe guiarse por las normas del *MLA Handbook* o seguirá las siguientes instrucciones:

- Libros: Apellido (s), Nombre, *Título* (cursiva) (Lugar: editorial, año) páginas (abreviatura: p., pp.)
- Artículo en revista: Apellido (s), Nombre, "Título" (entrecomillado), *Nombre* de la revista (cursiva), volumen/número (año) páginas (abreviatura: p., pp.)
- Artículo en obra colectiva: Apellido (s), Nombre, "Título" (entrecomillado), Editor (es), *Título* (cursiva) (Lugar: editorial, año) páginas (abreviatura: p., pp.)
- en las notas a pie de página el orden de cualquier referencia debe ser Nombre, Apellido, etc.

No se acepta el uso de subrayados o negritas, sólo las cursivas, salvo que formen parte de algún sistema de notación objeto del trabajo. Tampoco (misma excepción) los signos: <>, «» , { }.

Abreviaturas más frecuentes: *cfr., o.c.* (obra citada), *idem, ibidem, v.* (ver).

Se debe colocar un espacio después de cada signo de puntuación, también antes de guiones (para homogeneizar se recomienda el uso del guión corto).

Para las elipsis debe utilizarse el corchete y tres puntos dentro del mismo […].

Las reseñas darán toda la información del libro, además del número de páginas, el traductor si lo hay, etc. el nombre del reseñador se debe incluir al final del texto.

Editado el número se hará llegar a los autores un ejemplar de la revista. Para publicar en esta revista es necesario ser suscriptor de la misma. Una vez aceptado el artículo, el autor está obligado a suscribirse a dos años de esta publicación.

Para cualquier duda en relación a la publicación, abreviaturas aceptadas, etc. consúltese anticipadamente con el editor de la revista.Los autores interesados en separatas las pueden solicitar anticipadamente (25) y deberán abonar 100 euros.

Los artículos enviados, que deberán ser originales, serán revisados por dos lectores externos a las entidades editoras, que realizarán las oportunas observaciones; en caso de disparidad se enviará el trabajo a un tercer lector externo, y sólo se aceptará el trabajo cuando dos de los informes sean positivos.

Las lenguas en las que los trabajos se pueden presentar son: español, portugués e inglés.

R.I.L. se envía para su indización a DIALNET.

La suscripción anual es de 20 euros para españa y 30 euros para el extranjero. Las personas interesadas en una suscripción pónganse en contacto a través del siguiente correo electrónico:

cuc@universitascastellae.es