

# ÍNDICE

PRÓLOGO.....	11
INTRODUCCIÓN .....	15
CAPÍTULO 1. LOS ESTUDIOS DE DISPONIBILIDAD LÉXICA.	
ORÍGENES Y DESARROLLO ACTUAL .....	19
1.1. La disponibilidad léxica en el mundo hispánico .....	22
1.2. Aplicaciones de los estudios de disponibilidad.....	23
1.3. Disponibilidad léxica y enseñanza de español como lengua extranjera.....	27
CAPÍTULO 2. METODOLOGÍA DE LA INVESTIGACIÓN.....	33
2.1. La muestra.....	33
2.1.1. Selección de los informantes .....	33
2.1.2. Variables seleccionadas.....	34
2.1.3. Distribución de los informantes por variables.....	37
2.2. La encuesta.....	39
2.3. Los centros de interés .....	40
2.3.1. La selección de centros de interés en los estudios de disponibilidad léxica .....	43
2.4. La edición de los datos .....	47
2.4.1. Criterios comunes a todos los trabajos de disponibilidad léxica .....	48
2.4.2. Criterios propios de la investigación en español como lengua extranjera.....	54
2.4.3. Características específicas de cada centro de interés .....	61
2.5. Tratamiento informático y estadístico de los datos .....	68
CAPÍTULO 3. ANÁLISIS CUANTITATIVO .....	73
3.1. Resultados generales.....	74
3.1.1. Número de palabras y media de palabras por informante.....	74
3.1.2. Número de vocablos.....	77
3.1.3. Densidad léxica e índice de cohesión.....	79
3.2. Resultados según variables .....	83
3.2.1. Pruebas de contingencia .....	84
3.2.2. Variable «nivel de español» .....	95
3.2.3. Variable «lengua materna» .....	106
3.2.4. Variable «sexo».....	122
3.2.5. Variable «uso del español».....	128
3.2.6. Variable «conocimiento de otras lenguas» .....	136

CAPÍTULO 4. ANÁLISIS CUALITATIVO. RELACIONES LÉXICAS Y RELACIONES «SIGNIFICATIVAS» EN LOS CENTROS DE INTERÉS.....	145
4.1. Tipos de relaciones en semántica.....	146
4.1.1. Relaciones semánticas y relaciones léxicas.....	149
4.1.2. Las relaciones «significativas».....	156
4.2. Las relaciones léxicas y «significativas» en el léxico disponible .....	160
4.2.1. Metodología empleada para el estudio de las relaciones en los centros de interés .....	161
4.2.2. Análisis de las relaciones en los centros de interés.....	162
CONCLUSIONES.....	193
REFERENCIAS BIBLIOGRÁFICAS.....	199

## PRÓLOGO

Prologar un libro científico constituye siempre para todo universitario que se precie un motivo de gran alegría y emoción, y si, como en este caso, se trata del primer fruto académico redactado por mi querida amiga y compañera de área, Marta Sánchez-Saus Laserna, tal circunstancia me colma de orgullo y enorme satisfacción. De esta profesora universitaria quisiera resaltar su gran formación como lingüista y filóloga, pues cuenta con dos titulaciones: Filología Hispánica y Lingüística, en las que, tanto en la Universidad de Sevilla como en la de Cádiz, ha sido una alumna excelente y modélica, como lo prueban sus magníficos expedientes académicos, que le sirvieron para optar a una beca del personal investigador del MEC y la obtención de varios premios, el más importante el Segundo Premio Nacional de titulaciones de 2º ciclo de la rama de Humanidades.

Corría un sábado de otoño, concretamente el 22 de octubre de 2011, cuando Marta Sánchez-Saus Laserna defendía, como un hito importante de su trayectoria académica, docente e investigadora, su tesis doctoral con mención europea titulada *Bases semánticas para el estudio de los centros de interés del léxico disponible. Disponibilidad léxica de informantes extranjeros en las universidades andaluzas*, leída, bajo mi dirección, en la Facultad de Filosofía y Letras de la Universidad de Cádiz, ante un tribunal compuesto por los profesores doctores D. Francisco Moreno Fernández (presidente), D. Óscar Loureda Lamas, D. Martin Hummel, D. Juan Andrés Villena Ponsoda y D. Luis Escoriza Morera, que mereció la máxima calificación de sobresaliente *cum laude* por unanimidad. A todos ellos quiero agradecerles profundamente que aceptaran formar parte de la comisión evaluadora.

Este libro, cuya valía científica la avala el ser un trabajo derivado de la que fuera su tesis doctoral, recoge el cuerpo central de aquella investigación, concretamente el análisis del léxico disponible de estudiantes de español como lengua extranjera en las universidades andaluzas, ya que el estudio completo incluía un valioso capítulo historiográfico sobre el concepto «centro de interés» y sus antecedentes en la semántica histórica, «tradicional» y pre-estructural, que aparecerá publicado próximamente como monografía por el Instituto Universitario de Investigación en Lingüística Aplicada de la Universidad de Cádiz.

Desde hace veinte años, concretamente desde que el doctor Humberto López Morales fuera, durante el curso 1995-96, profesor visitante del Área de Lingüística General de la Universidad de Cádiz para la impartición de la asignatura «Sociolingüística» de la antigua licenciatura en Lingüística y la formación de docentes en este campo del saber lingüístico, hemos de subrayar que su paso tuvo asimismo repercusiones relevantes en el ámbito de la investigación, como queda constatado por el hecho de que el grupo de investigación de excelencia «Semaínein» (PAI-HUM147) haya desarrollado desde entonces nuevas líneas y proyectos en el dominio de esta disciplina. Sólo así se explica 1) la participación conjunta en el «Proyecto para el estudio sociolingüístico del español de España y América» (PRESEEA); 2) la dirección de tres «Acciones Coordinadas de Investigación y Desarrollo Tecnológico» (2001, 2002 y 2003), financiadas por la Secretaría General de Universidades e Investigación de la Consejería de Educación y Ciencia de la Junta de Andalucía, con diferentes grupos de investigación andaluces, para el desarrollo de las siguientes líneas de actuación en el campo de la sociolingüística: disponibilidad léxica, variación lingüística, contacto lingüístico e interculturalidad (integración lingüística de inmigrantes) y convergencia y divergencia dialectales; 3) la inclusión en el proyecto, coordinado por el doctor López Morales, sobre «Disponibilidad léxica en escolares del mundo hispánico»; 4) la realización del proyecto «Lenguas en contacto y disponibilidad léxica: la situación lingüística e intercultural de Ceuta y Gibraltar» (BFF2000-0511), financiado por la DGICYT del Programa General del Conocimiento del MEC y los Fondos Feder en su convocatoria de 1999-2000, programa de investigación relacionado naturalmente con esta tesis y que ha llevado consigo la publicación de numerosos artículos científicos en torno a la comunidad de habla de Gibraltar y de dos monografías posteriores, como las de los profesores Carmen Ayora Esteban y Gérard Fernández Smith sobre los léxicos disponibles de Ceuta y Melilla, respectivamente, y 5) que el campo de la sociolingüística y, en concreto, sus aspectos más relevantes como el variacionismo, el contacto lingüístico, la política y planificación lingüísticas y, por supuesto, la ya mencionada disponibilidad léxica se hayan convertido en líneas de trabajo principales del citado grupo de investigación.

Por lo que respecta a la monografía que ahora presentamos, supone, en líneas generales, una aportación relevante, no sólo para el campo de los estudios de disponibilidad léxica, sino para el de sus aplicaciones, particularmente en un ámbito tan actual como el de la enseñanza del español como lengua extranjera. Pero es que, en cuanto a su contenido, este trabajo enriquece cuantitativa y cualitativamente la indagación en el terreno de la disponibilidad léxica, además de que aporta a la metodología habitual utilizada en este campo ciertas innovaciones y un tipo de análisis cualitativo muy novedoso, al tiempo que proporciona a los investigadores de E/LE un valioso material.

De este modo, este *Léxico disponible de los estudiantes de español como lengua extranjera en las universidades andaluzas* constituye el primer trabajo sobre disponibilidad léxica desarrollado sobre una amplia muestra de estudiantes y con un detallado estudio estadístico sobre las variables que afectan a esta población. Además, resultan muy interesantes las innovaciones metodológicas introducidas, como la selección de los centros de interés y de las variables, adaptados al tipo de informantes sobre el que se ha efectuado la investigación: estudiantes extranjeros que están aprendiendo español como lengua extranjera en el momento en el que se realiza el estudio. Por otro lado, y es uno de los aportes más significativos de esta obra, se lleva a cabo un análisis de las relaciones léxicas y las relaciones «significativas» presentes dentro de los centros de interés, observando su frecuencia y la naturaleza de estas conexiones semánticas. Se trata, por tanto, de una investigación que se inserta dentro de una tradición ya amplia de trabajos sobre léxico disponible, pero muy original desde el momento en que, por un lado, no se centra en la lengua materna, sino en el español como lengua extranjera, objeto de estudio al que se adapta la metodología empleada, y, por otro, se materializa gracias al análisis de tipo semántico realizado sobre el concepto de centro de interés.

Todo trabajo de estas características tiene detrás una dedicación de años, caracterizados por el rigor, la seriedad, el esfuerzo y el entusiasmo para afrontar los problemas y retos que la investigación plantea día a día. Pero, además, en este caso, estoy francamente convencido de que este estudio va a marcar un antes y un después en el marco de los trabajos de disponibilidad léxica. Y con ello no me refiero exclusivamente a su contribución aplicada en el terreno de la enseñanza de lenguas, sino a otras aportaciones relevantes y novedosas, como su análisis empírico y estadístico, la descripción tipológica de las distintas clases de relaciones léxicas y significativas del léxico disponible recogido como un complemento sustancial al análisis cualitativo de los datos y su gran soporte teórico con las bases semánticas de los centros de interés, cuestión esta que indudablemente faltaba en este tipo de trabajos y que demuestra cómo no se debe hacer lingüística aplicada sin lingüística teórica y cómo debe guardarse en todo momento un equilibrio entre lo teórico-epistemológico y la praxis de los datos empíricos.

Por todo ello, me siento muy satisfecho y orgulloso del trabajo llevado a cabo por la doctora Marta Sánchez-Saus Laserna, que se ve brillantemente plasmado en esta monografía, y la felicito por ello y por el interés y entusiasmo con el que afronta los problemas teóricos, metodológicos y prácticos de la semántica y la sociolingüística; en definitiva, estoy muy feliz por la culminación de esta publicación, que, sin duda, va a repercutir muy positivamente en sus futuros retos universitarios.

Cádiz, 13 de enero de 2016

MIGUEL CASAS GÓMEZ

## INTRODUCCIÓN

Los estudios de disponibilidad léxica cuentan ya con una amplia trayectoria tras de sí desde su nacimiento en Francia a mediados del siglo pasado, cuando los autores del *Français Fondamental* idearon una nueva metodología que perfeccionaba la selección del léxico básico de una lengua mediante la unión del vocabulario frecuente con su complementario: el léxico disponible, aquel que primero viene a la mente de un hablante ante un tema de conversación determinado.

El éxito de esta metodología de investigación del léxico lo demuestran las numerosas publicaciones sobre disponibilidad léxica que, a nivel internacional y sobre diferentes lenguas, han ido apareciendo desde entonces. No obstante, es el español la lengua que mejor ha sabido aprovechar las ventajas de estos estudios, gracias, fundamentalmente, a la existencia de un proyecto común a España e Hispanoamérica para la confección de un diccionario del léxico disponible de la gran mayoría de hablantes de nuestra lengua, el Proyecto Panhispánico de Disponibilidad Léxica, al cual se han adherido numerosos investigadores.

Sin embargo, para entender esta difusión es necesario tener en cuenta el amplio abanico de posibilidades de aplicación de los resultados obtenidos en estos estudios a un gran número de disciplinas lingüísticas. Estas abarcan el objetivo primordial que se plantearon los trabajos pioneros, que es también la finalidad última de este trabajo: el perfeccionamiento de la selección del léxico que han de estudiar los estudiantes no nativos de una lengua, a partir de la determinación de las unidades que pertenecen al léxico básico y que, por tanto, deben ser enseñadas en los niveles inferiores.

Aunque la disponibilidad léxica cuenta ya, en su relativamente corta historia, con numerosos estudios, creemos que aún es posible presentar trabajos novedosos, que abran puertas a aspectos todavía no tratados con profundidad. En este sentido, la investigación que aquí presentamos<sup>1</sup> estuvo planteada,

---

1. Este trabajo ha podido llevarse a término gracias a una beca del Programa de Formación del Profesorado Universitario del Ministerio de Educación español de la que fui beneficiaria entre 2007 y 2011.

desde el primer momento, hacia la finalidad primitiva de la disponibilidad léxica, la que tuvieron en mente los primeros investigadores franceses: el perfeccionamiento de la enseñanza del vocabulario a aquellos estudiantes que quieren aprender una lengua extranjera.

Dentro de este ámbito, este trabajo se presenta como un estudio de la disponibilidad léxica de estudiantes de español como lengua extranjera, los cuales son, en el momento de la recogida de datos, alumnos de alguna universidad andaluza. Desde el comienzo hemos tenido muy en cuenta la existencia de otros, excelentes aunque ciertamente escasos, trabajos dedicados al léxico disponible de estudiantes de español como lengua extranjera (Carcedo 2000a y Samper Hernández 2002, fundamentalmente), pero no existía ninguno que combinara una muestra lo suficientemente amplia para extraer conclusiones relevantes con el hecho de contar con informantes de lenguas maternas diferentes, nivelados según los estándares del Marco común europeo de referencia para las lenguas (MCER)<sup>2</sup>. Además, creímos que era necesario adaptar algunos aspectos de la metodología que han empleado tradicionalmente los estudios de disponibilidad léxica en español (incluido en el Proyecto Panhispánico de Disponibilidad Léxica) al tipo de informantes con los que tratamos, sobre todo en lo relacionado con la selección de los centros de interés. Por ello, en esta investigación trabajamos con una lista propia, que, por un lado, recoge la mayor parte de los contenidos semánticos que recomienda el MCER para los niveles iniciales de español como lengua extranjera y que, por otro, guarda las semejanzas suficientes como para poder realizar futuras comparaciones entre los datos aquí presentes y los de otros trabajos (cf. capítulo 2). Así pues, nuestro objetivo principal ha sido proporcionar un estudio con una base de datos amplia y diversa, de manera que sea posible alcanzar el objetivo, al que tantas veces se ha hecho referencia, de aplicar la disponibilidad léxica al perfeccionamiento de los métodos de enseñanza del léxico a alumnos extranjeros y de los materiales de aprendizaje que estos emplean.

Con este planteamiento general, el presente trabajo se organiza en cuatro partes. La primera de ellas, el capítulo 1, sirve de marco para la investigación. En él repasamos los hitos fundamentales en la historia de la disciplina, para centrarnos a continuación en las aplicaciones que ha tenido en la enseñanza de español a extranjeros, rama esta en la que se ha centrado el interés de los investigadores en los últimos años, con notables resultados.

El segundo capítulo está dedicado a los aspectos metodológicos de nuestra investigación en disponibilidad léxica, concretamente a la elección de la muestra, las variables empleadas para su estratificación, la selección de los

---

2. Publicado en su versión española por el Instituto Cervantes y el Ministerio de Educación, Cultura y Deporte (2002). Puede consultarse online aquí: <[http://cvc.cervantes.es/ensenanza/biblioteca\\_ele/marco/](http://cvc.cervantes.es/ensenanza/biblioteca_ele/marco/)>.

centros de interés y los criterios de edición de los datos. Puesto que la población estudiada es diferente de la de la mayoría de estudios –se trata de alumnos para los que el español es una segunda lengua, no de hablantes nativos–, tanto la selección de centros de interés como las variables contempladas se han adaptado a los fines concretos que perseguimos. No obstante, hemos procurado mantener todos los paralelismos que fuera posible conservar, de manera que se puedan llevar a cabo comparaciones entre los resultados que aquí aportamos y los de otros trabajos.

Estas singulares características hacen del análisis cuantitativo de los resultados, que se desarrolla en el capítulo 3, un estudio también particular. La primera parte de este análisis se refiere a datos generales de toda la muestra: número total de palabras recogidas en el estudio, número de palabras diferentes o *vocablos*, media de palabras por persona y centro de interés e índice de cohesión de cada centro. La segunda parte está dedicada al estudio de los resultados cuantitativos por variables, las cuales codifican factores que pueden influir en el nivel de conocimiento del español, como el nivel que tienen los informantes en esta lengua, el uso que hacen de ella, el conocimiento que tengan de otras lenguas románicas y su propia lengua materna, así como el sexo. Estos factores son los que pueden, además, influir de manera más práctica en la adaptación efectiva de los materiales de enseñanza de vocabulario a alumnos extranjeros, por lo que el análisis de su significatividad resulta de gran interés a la hora de aplicar los resultados de este análisis a la mejora de la enseñanza del español.

Finalmente, el último capítulo está dedicado al análisis cualitativo del léxico disponible obtenido. Nos centramos en el estudio de las conexiones más frecuentes que se dan entre palabras en el interior de los centros de interés, a partir de la propuesta de Casas (2005, 2011 y 2014) sobre la diferenciación entre relaciones léxicas y relaciones «significativas». Con la ayuda del programa *DispoGrafo* extraemos las relaciones entre pares de palabras más frecuentes, de manera que podemos analizar la naturaleza de aquellos vínculos que sustentan la organización de los centros. Para ello, será de gran importancia no perder de vista el límite entre lo intralingüístico y lo extralingüístico y, más aún, distinguir claramente los vínculos que se establecen entre significados de aquellos en los que intervienen otros componentes del signo lingüístico o los signos completos y, más allá, de las relaciones que se establecen entre las cosas, que pertenecen al mundo de la designación.

Cierran el libro las conclusiones de la investigación y las referencias bibliográficas, así como los anexos, donde se encuentran el modelo de encuesta utilizado y los diccionarios de léxico disponible, ordenado por centros de interés y por orden alfabético.

El trabajo que aquí presentamos procede fundamentalmente de nuestra tesis doctoral (Sánchez-Saus 2011), dirigida por el profesor Miguel Casas Gómez, que fue reconocida con el Premio Extraordinario de Doctorado de la

Universidad de Cádiz. Esperamos que, aun dentro de sus limitaciones, cumpla los objetivos que nos marcamos en un principio y que nos han acompañado durante estos años: ayudar a mejorar la selección del léxico que se emplea en la enseñanza de español a extranjeros mediante el estudio del léxico disponible de estos hablantes y de las variables que realmente influyen en su conocimiento del vocabulario y profundizar en las relaciones entre palabras que sostienen los procesos de producción léxica a partir de estímulos asociativos.

Nada de lo aquí escrito habría sido posible sin la dirección, apoyo y confianza del profesor Miguel Casas Gómez. Igualmente debo agradecer a toda el área de Lingüística General de la Universidad de Cádiz su afecto y tiempo dedicado a resolver todas mis dudas. Aunque la investigación se realizó en Andalucía, fue en Alemania donde tomó forma, gracias a la hospitalidad y ayuda del profesor Óscar Loureda, de la Universidad de Heidelberg. No puedo olvidar a Andrés Jiménez, sin cuya desinteresada ayuda el análisis estadístico habría sido un escollo insalvable, ni a los directores, coordinadores y profesores de español de los centros de idiomas de las universidades andaluzas. A ellos y, sobre todo, a todos los alumnos cuyas palabras son la razón de este trabajo, va mi más sincero agradecimiento. Y, desde luego, a mi familia, especialmente a Manuel y, ahora también, a Julia, por el tiempo que dejé de estar con ellos y por acompañarme en todo este camino. Gracias.