

DISPONIBILIDAD LÉXICA

---

TEORÍA, MÉTODO  
Y ANÁLISIS



WYDAWNICTWO  
UNIWERSYTETU  
ŁÓDZKIEGO

<http://dx.doi.org/10.18778/7969-437-2>

Antonio María López González

# DISPONIBILIDAD LÉXICA

---

## TEORÍA, MÉTODO Y ANÁLISIS



WYDAWNICTWO  
UNIWERSYTETU  
ŁÓDZKIEGO

ŁÓDŹ 2014

Antonio María López González – Uniwersytet Łódzki, Wydział Filologiczny  
Katedra Filologii Hiszpańskiej, 90-236 Łódź, ul. Pomorska 171/173

RECENZENT

*Jerzy Szalek*

REDAKTOR WYDAWNICTWA UŁ

*Bogusława Kwiatkowska*

SKŁAD KOMPUTEROWY

*Oficyna Wydawnicza Edytor.org*

*Lidia Ciecierska*

PROJEKT OKŁADKI

*Lukasz Orzechowski*

Zdjęcie wykorzystane na okładce: © Depositphotos.com/zagandesign

© Copyright by Uniwersytet Łódzki, Łódź 2014

Wydane przez Wydawnictwo Uniwersytetu Łódzkiego

Wydanie I. W.06638.14.0.M

ISBN 978-83-7969-437-2

Wydawnictwo Uniwersytetu Łódzkiego

90-131 Łódź, ul. Lindleya 8

www.wydawnictwo.uni.lodz.pl

e-mail: ksiegarnia@uni.lodz.pl

tel. (42) 665 58 63, faks (42) 665 58 62

Druk i oprawa: Quick Druk

# ÍNDICE

<b>Introducción</b> .....	9
i. <b>Ámbito de la investigación</b> .....	9
ii. <b>Esta investigación: Presentación y propósitos</b> .....	12
<b>I. Los estudios de disponibilidad léxica</b> .....	17
1.1. <b>La medición estadística del vocabulario</b> .....	17
1.1.1. <b>Los diccionarios normativos</b> .....	17
1.1.2. <b>La frecuencia léxica</b> .....	19
1.1.3. <b>El léxico básico</b> .....	22
1.1.4. <b>La disponibilidad léxica</b> .....	24
1.2. <b>El estudio de la disponibilidad léxica en español</b> .....	26
1.2.1. <b>Los estudios anteriores al Proyecto Panhispánico</b> .....	26
1.2.2. <b>El Proyecto Panhispánico de Disponibilidad Léxica</b> .....	30
<b>II. Bases teóricas y metodológicas de la disponibilidad léxica</b> .....	35
2.1. <b>Unidades de análisis del corpus léxico</b> .....	35
2.2. <b>El procesamiento del corpus léxico</b> .....	36
2.3. <b>Criterios de edición del Proyecto Panhispánico</b> .....	40
2.4. <b>El índice de disponibilidad léxica</b> .....	43
2.5. <b>El tamaño de la muestra</b> .....	45
<b>III. Los centros de interés</b> .....	49
3.1. <b>Concepto y selección</b> .....	49
3.2. <b>Clasificación de los centros de interés</b> .....	54
3.3. <b>Regularidades en los centros de interés</b> .....	57
3.3.1. <b>Promedio de respuestas</b> .....	58
3.3.2. <b>Número total de palabras</b> .....	65
3.3.3. <b>Número de vocablos</b> .....	68
3.3.4. <b>Índice de cohesión</b> .....	76
<b>IV. Relaciones entre las palabras en el léxico disponible</b> .....	83
4.1. <b>La fluencia semántica y la teoría de los prototipos</b> .....	84
4.1.1. <b>La elaboración de las listas</b> .....	84
4.1.2. <b>La categorización</b> .....	86
4.2. <b>La asociación de palabras y las redes semánticas</b> .....	90
4.2.1. <b>La asociación de palabras en las listas de léxico disponible</b> .....	90
4.2.2. <b>La representación gráfica de la red semántica</b> .....	94
4.2.3. <b>Las relaciones significativas</b> .....	97

<b>V. Los condicionantes extralingüísticos de la disponibilidad léxica</b> .....	103
5.1. El sexo .....	104
5.2. El nivel sociocultural .....	108
5.3. El factor espacial: urbano/rural .....	112
5.4. El tipo de centro escolar .....	115
5.5. La edad .....	119
5.6. Los condicionantes extralingüísticos en las comunidades bilingües .....	124
5.6.1. ‘Lengua materna’ .....	125
5.6.2. ‘Lengua de uso familiar’/‘Lengua habitual’ .....	129
5.6.3. ‘Modelo de enseñanza’ (monolingüe/bilingüe) .....	133
5.6.4. Las transferencias léxicas en las comunidades bilingües .....	136
<b>VI. El análisis cualitativo del léxico disponible</b> .....	143
6.1. El análisis cualitativo: fines y posibilidades de estudio .....	143
6.2. La composición de los listados de léxico disponible .....	145
6.3. Clasificación del léxico disponible .....	156
6.3.1. Análisis según el índice de disponibilidad léxica .....	157
6.3.2. Análisis descriptivo lexicológico .....	161
6.3.3. Análisis descriptivo según niveles de lengua (registros) .....	166
6.4. Los extranjerismos en el léxico disponible del español .....	169
6.5. El léxico peculiar: los dialectalismos .....	176
<b>VII. La comparación de los léxicos disponibles</b> .....	183
7.1. La norma léxica del español peninsular .....	183
7.1.1. La delimitación de la norma léxica según el léxico disponible .....	183
7.1.2. Las comparaciones interdialectales y el léxico compartido .....	185
7.1.3. Cotejo de cinco léxicos disponibles peninsulares .....	186
7.2. Análisis descriptivo de los centros de interés y su norma léxica .....	190
7.2.1. ‘Partes del cuerpo humano’ .....	190
7.2.2. ‘La ropa’ .....	194
7.2.3. ‘Partes de la casa (sin muebles)’ .....	197
7.2.4. ‘Los muebles de la casa’ .....	200
7.2.5. ‘Comidas y bebidas’ .....	203
7.2.6. ‘Objetos sobre la mesa para la comida’ .....	207
7.2.7. ‘La cocina y sus utensilios’ .....	210
7.2.8. ‘La escuela (muebles y materiales)’ .....	214
7.2.9. ‘Iluminación, calefacción y medios de airear un recinto’ .....	217
7.2.10. ‘La ciudad’ .....	221
7.2.11. ‘El campo’ .....	224
7.2.12. ‘Medios de transporte’ .....	227
7.2.13. ‘Trabajos del campo y del jardín’ .....	231
7.2.14. ‘Los animales’ .....	234
7.2.15. ‘Juegos y distracciones’ .....	237
7.2.16. ‘Profesiones y oficios’ .....	242
<b>Conclusiones</b> .....	247
<b>Bibliografía</b> .....	259
<b>Dostępność leksykalna. Teoria, metoda i analiza. Streszczenie</b> .....	279
<b>Lexical availability: theory, method and analysis. Summary</b> .....	283
<b>Od Redakcji</b> .....	287

## AGRADECIMIENTOS

Un trabajo como el que aquí se presenta no sería posible sin la ayuda y el apoyo de muchas personas, a las que quiero dedicar estas líneas, dándoles nombre propio.

Así, quiero dar las gracias a Francisco García Marcos, mi maestro en la sociolingüística, por iniciarme en el camino de la disponibilidad léxica. A él y a M.<sup>a</sup> Victoria Mateo García, mi agradecimiento por compartir tantas ideas y materiales. Igualmente quiero dar las gracias a José A. Samper Padilla, verdadero biógrafo de la disponibilidad léxica en español, por responder siempre a mis consultas y facilitarme sus publicaciones. Gracias también a José A. Bartol Hernández y Natividad Hernández Muñoz, por permitirme el acceso al banco de datos y a las herramientas informáticas de *Dispolex.com*. Asimismo, en el campo informático, quiero agradecer a Marta Sánchez-Saus Laserna y Paula Urzua sus explicaciones sobre el funcionamiento del programa *Dispografo*.

Quiero dar las gracias, en fin, a los estudiosos de la disponibilidad léxica que han compartido conmigo sus materiales y que han respondido a mis consultas. Por este motivo, además de a los arriba mencionados, les estoy agradecido: Antonio M. Ávila Muñoz, Carmen Fernández Juncal, M.<sup>a</sup> Victoria Galloso Camacho, Antonio García Megía, Adolfo González Martínez, Belén López Meirama, Francisco Moreno Fernández, Josefina Prado Aragonés, Ventura Salazar García, Marta Samper Hernández, Javier de Santiago Guervós, Cristina Tabernero Sala, Antonia Torres González, Ester Trigo Ibáñez.

Quiero tener un recuerdo aquí para dos personas, que ya no están con nosotros, y que han influido mucho con sus investigaciones en el devenir de la disponibilidad léxica y, por consiguiente, en esta monografía. Se trata de Alberto Carcedo González, lectura obligada para mis trabajos sobre la disponibilidad léxica en español como lengua extranjera, y de Max. S. Echeverría, autor del programa *Dispografo*, a quien tuve la enorme suerte de conocer como alumno del taller de Didáctica que impartió en el XIII Congreso de ALFAL.

Gracias, finalmente, a Wiaczeslaw Nowikow, director de la Cátedra de Filología Española de la Universidad de Lódz, por su ánimo e infinita paciencia para conmigo.

Y siempre, un lugar especial en mi corazón, para Janina, mi mujer, y para mis hijos Ana y Antonio, por todas las horas en común que este trabajo les ha robado, y los que, sin embargo, siempre, siempre, me han acompañado en los momentos de fatiga. Os quiero.

## INTRODUCCIÓN

### i. **Ámbito de la investigación**

El estudio de la disponibilidad léxica ha demostrado ser un campo fecundo en las investigaciones lingüísticas hispánicas desde finales de los años 90, tanto desde el punto de vista cuantitativo –número de trabajos– como cualitativo –reflexión teórica y desarrollo metodológico. De ello han levantado acta distintos estudios que, sucesivamente y a partir de una exhaustiva revisión bibliográfica, han dado cuenta del progreso de estas investigaciones.

Así, si en 2003 Samper Padilla, Bellón Fernández y Samper Hernández (2003: 27) se referían a que “varios grupos de investigación de ambos lados del Atlántico [trabajaban] en este aspecto de la léxico-estadística”, lo que había supuesto un cambio radical con respecto a las pocas investigaciones realizadas en ese campo hasta aquel año, sólo tres años más tarde, los mismos Samper Padilla y Samper Hernández (2006: 5) certificaban que “el estudio de la disponibilidad léxica [había] experimentado un desarrollo muy destacado en el mundo hispánico”. Desde entonces, la investigación en este campo no ha hecho más que crecer, lo que hace que ya en 2012 Paredes García (2012: 78–79) pueda presumir del “estado efervescente en que se encuentra la investigación de la disponibilidad léxica” y de “la eclosión de la disponibilidad léxica en España y América”, habida cuenta del amplio censo de monografías, tesis doctorales, memorias de licenciatura o máster, artículos en revistas u obras colectivas, etc., siempre en constante crecimiento.

El extraordinario desarrollo de este campo de investigación en el mundo hispánico se debe sobre todo al *Proyecto Panhispánico de Estudio sobre la Disponibilidad Léxica*, un macroyecto que coordina y dirige Humberto López Morales, con equipos inscritos en las diferentes zonas dialectales del mundo hispánico. No obstante, “hay que destacar especialmente por sus aportes teóricos y/o metodológicos las investigaciones llevadas a cabo en Puerto Rico, la República Dominicana-

na, Chile y México”, así como España, donde “el territorio estudiado cubre más de cincuenta provincias españolas [...] lo que posibilitará en un futuro próximo la publicación del *Diccionario del léxico disponible de España*” (Paredes García, 2012: 79). El fin último de este macroproyecto es la elaboración de un diccionario del léxico disponible del español.

A la coordinación de los estudiosos de la disponibilidad léxica en el ámbito hispanohablante ha contribuido también desde octubre de 2000 la página electrónica [www.disporex.com](http://www.disporex.com), creada y mantenida por José Antonio Bartol Hernández y Natividad Hernández Muñoz desde la Universidad de Salamanca. Esta página sirve como medio de contacto e información a todas las personas que investigan o se interesan por la disponibilidad léxica, proporcionando detalles sobre la forma más habitual de proceder en este campo, y acceso gratuito a las herramientas necesarias para realizar los cálculos más habituales en disponibilidad léxica y a un banco de datos de las investigaciones realizadas.

Hablar de *disponibilidad léxica* significa referirse al “caudal léxico *utilizable* en una situación comunicativa dada” (López Morales, 1996: 245), a aquellas palabras que acuden antes a la mente del hablante como reacción a un estímulo temático concreto (López Morales, 1986, 1999a). Estas palabras constituyen el *vocabulario disponible* de ese hablante o de esa comunidad de hablantes. Se entiende, entonces, “que existe en el lexicón mental una serie de términos que no se actualizan a menos que sea necesario para comunicar una información muy específica” (López Morales, 2009: 26).

Los estudios de disponibilidad léxica se iniciaron en Francia a mediados del siglo XX, cuando un grupo de lingüistas recibió el encargo de elaborar *Le Français Élémentaire*, un instrumento de planificación didáctica para la enseñanza del francés como L2, en las excolonias de l’Union Française, para los inmigrantes llegados a Francia, y para los extranjeros en general (Gougenheim *et al.*, 1956: 8–9). Para seleccionar los vocablos que debían integrar el léxico “elemental”, recurrieron a los listados léxicos de frecuencias, pero se dieron cuenta de que en esos recuentos faltaban términos cotidianos y familiares, “palabras que se referían a realidades concretas y que no aparecían en los listados basados en la frecuencia porque su uso estaba muy condicionado por el tema de la conversación o del escrito” (Bartol Hernández, 2006: 379).

Enseguida, los investigadores franceses se percataron de que el léxico fundamental de una lengua incluye dos subconjuntos léxicos fácilmente distinguibles (Michéa, 1953):

a) el *léxico básico*, frecuente y atemático, principalmente palabras gramaticales (adjetivos, verbos y sustantivos de carácter general), formado por palabras que aparecen continuamente en cualquier conversación o escrito, independientemente del tema de que se trate (ej. – *a, los, no, mucho, hay, dar, persona, poner*, etc.);

b) el *léxico disponible*, temático, de contenido semántico muy concreto, integrado por palabras que, aun siendo de uso común, sólo se emplean ligadas a un tema (ej. – *sartén* y *tenedor* ligadas a la ‘cocina’, o *carta* y *sello* relacionadas al ‘correo’).

Para recabar ese léxico disponible, los estudiosos franceses prepararon en 1953 unas pruebas asociativas a las que respondieron 416 niñas y 488 niños de 9 a 12 años en cuatro departamentos franceses. Con este estudio, y los realizados posteriormente por Mackey (1971a, 1971b) en áreas francoparlantes de Canadá, se cimentó la base metodológica de la disponibilidad léxica: la recogida de materiales mediante encuestas de listas abiertas en las que durante dos minutos los informantes anotan todas las palabras que les vienen a la mente en relación con una serie de estímulos temáticos.

Como se aprecia, la disponibilidad léxica nació –y se mantiene– ligada a una clara finalidad pedagógica: la necesidad de seleccionar los vocablos que deben ser enseñados en la clase de lengua (extranjera o materna). Al mismo tiempo, se caracteriza por indagar en el mundo conceptual-real significativo de una determinada comunidad (Gómez Devís, 2004: 8). La utilidad de estas investigaciones es, por tanto, innegable para la lingüística aplicada dada la importancia excepcional que revisten las aplicaciones pedagógicas. Junto a estas, los estudios sobre disponibilidad léxica realizan distintas tareas que, en su día, Azurmendi Ayerbe (1983: 187–195) clasificó en los siguientes tres grupos:

- El establecimiento de léxicos básicos, para los que la disponibilidad léxica refleja mejor el uso oral y es más representativa de las realidades sociales.
- La reflexión teórica sobre las regularidades en el comportamiento de la disponibilidad y la validez del método de asociación controlada, sobre las relaciones entre el comportamiento de la disponibilidad y el tamaño de la muestra, así como sobre el tamaño de respuestas óptimas.
- La comparación inter e intralingüística, sea buscando la comparación y la distancia interlingüística, sea estudiando la comparación intralingüística al analizar las variedades geolingüísticas de la misma lengua.

Además, se han multiplicado los trabajos de índole dialectológica, sociolingüística, psicolingüística y etnolingüística (cf. Samper Padilla *et al.*, 2003: 105 y ss.). Al respecto, Ávila Muñoz y Villena Ponsoda (2010: 42–44) enumeran las siguientes aplicaciones:

1) La sociolingüística “somete la disponibilidad a análisis de covariación con los factores sociales que le interesa manejar y determina la caracterización léxica de los sociolectos que integran la comunidad en cuestión” (López Morales, 1999a: 25). Trabajando principalmente a base de cuantificación y no (siempre) de variación cualitativa, diversos estudios<sup>1</sup> han demostrado que, “por encima de cualquier

---

<sup>1</sup> Entre otros, destacan: Samper Padilla (2006), Samper Padilla y Hernández Cabrera (1997), Gómez Molina (2006), Gómez Molina y Gómez Devís (2004), Mateo García (1996, 1998) o Samper Hernández (2003a).

otro condicionante sociocultural, la edad y el tipo de educación formal de los padres son los factores que determinan las diferencias encontradas en las listas de disponibilidad léxica que generan los distintos grupos de hablantes” (Ávila Muñoz y Villena Ponsoda, 2010: 43).

2) La psicolingüística se interesa por la producción de la lista de léxico durante la prueba de disponibilidad léxica. Esta prueba conlleva una tarea cognitiva compleja en la que intervienen determinados procesos psicológicos que, en gran medida, condicionan la mencionada lista, y que han sido investigados, entre otros, por Hernández Muñoz (2006) y Ávila Muñoz y Villena Ponsoda (2010).

3) La etnolingüística estudia las diferencias culturales que se observan entre las comunidades de habla estudiadas a través de los listados de vocablos disponibles, en los que el léxico disponible de cada sociedad aparece como un reflejo de su cultura. Las relaciones entre vocabulario y cultura fueron señaladas ya por Michéa (1950) e investigadas por Mackey (1971a).

4) La dialectología comparada realiza comparaciones diatópicas de los materiales proporcionados por la disponibilidad léxica. El punto de inflexión en estos estudios comparados lo marcó el desarrollo de los índices matemáticos de López Chávez (1992, 1993), que permiten la comparación de las listas de disponibilidad obtenidas en diferentes sintopías. “Todos los estudios que las han empleado nos muestran que junto a los vocablos comunes existen otros exclusivos, lo que nos permite obtener conclusiones sobre el grado de aproximación o compatibilidad entre los diferentes dialectos”<sup>2</sup> (Ávila Muñoz y Villena Ponsoda, 2010: 44).

5) La lexicología analiza la composición léxica de las listas, especificando el léxico patrimonial, los grupos léxicos, los acrónimos, las marcas, los extranjerismos, etc. En especial, se han desarrollado estudios sobre la penetración y el alcance de los extranjerismos<sup>3</sup> o la presencia de faltas de ortografía en las listas, como en el trabajo de Galloso Camacho (2003a: 185–286), en el que esta autora advierte del “panorama desolador” en materia ortográfica y hace una llamada de atención a todos los “organismos competentes” para empezar a tomar medidas sin demora.

## ii. Esta investigación: Presentación y propósitos

El volumen que tiene el lector entre sus manos tiene como objetivo analizar de una manera crítica las bases teóricas y metodológicas en las que se apoyan los estudios de léxico disponible. A partir de una amplia revisión bibliográfica, se pasa revista a las aportaciones más destacadas del Proyecto Panhispánico, presentando los diferentes ámbitos lingüísticos en los que se desarrollan estos estudios.

---

<sup>2</sup> Vid. López Morales (1999: 20–21).

<sup>3</sup> Entre otros, destacan los trabajos de Paredes García (2001b), Carrera y Bradley (2004), Bartol Hernández (2005) y Castañer Martín (2008).

En este sentido, pretende completar y actualizar los trabajos anteriores de Samper Padilla *et al.* (2003), y de Samper Padilla y Samper Hernández (2006), en los que se trazaba un amplio panorama de los logros de la disponibilidad léxica en lengua española<sup>4</sup>.

Como hemos anticipado, la disponibilidad léxica nació con fines pedagógicos en Francia (Gougenheim *et al.*, 1956). Estos mismos objetivos inspiraron a Humberto López Morales (1973) a llevar a cabo la primera investigación de disponibilidad léxica en el mundo hispánico, en la que analiza la adquisición del vocabulario y las variables que en él inciden, a partir de una muestra de escolares de la zona metropolitana de San Juan de Puerto Rico. Naturalmente, a esta finalidad didáctica hacemos alusión en este libro.

Sin embargo, con ser la perspectiva didáctica de importancia máxima para el desarrollo de estos estudios, en esta ocasión nos ocupamos solo de manera secundaria de este tipo de aplicaciones, ya que el objetivo de esta monografía es el análisis de la fundamentación teórica y metodológica de estos trabajos. Se trata, pues, de una investigación de corte epistemológico, en el sentido en que este estudio resulta en una aportación sobre el origen, la estructura, los métodos y la validez misma de la investigación sobre el léxico disponible. Para ello, este trabajo ha buceado en los estudios introductorios de los trabajos que aquí se mencionan, comparando los planteamientos teóricos de los que parten, la metodología empleada que de ellos se deriva, y los resultados que con su aplicación se obtienen.

Estructuralmente, este libro se compone de siete capítulos, aparte de la Introducción y el capítulo de Conclusiones. El capítulo I (*Los estudios de disponibilidad léxica*), está dedicado, por una parte, a la evolución histórica de los procedimientos principales que miden el vocabulario existente en una comunidad de habla. Entre estos procedimientos, el léxico disponible estudia un léxico de contenido semántico muy concreto, existente en el lexicón mental, pero de carácter potencial. Por otra parte, se da testimonio de cómo el español ha extraído el mejor provecho de la investigación cuantitativa del léxico por medio de la disponibilidad léxica. Después del trabajo inicial de 1973, Humberto López Morales siguió investigando en torno a diversos aspectos descriptivos y metodológicos de la disponibilidad léxica, investigaciones que alcanzaron su punto culminante con la publicación del *Léxico disponible de Puerto Rico* (López Morales, 1999a). Entre una y otra fecha, los estudios sobre el léxico disponible se extendieron por distintos países de Hispanoamérica (Puerto Rico, República Dominicana, Costa Rica, México, Chile), llegando a España en 1992 con los primeros estudios de Benítez Pérez (1992) en Madrid, y Samper Padilla y Hernández Cabrera (1992) en Las Palmas de Gran Canaria. Desde entonces, los estudios de léxico disponible han proliferado por toda la geografía española, sobre todo, gracias a la existencia de

---

<sup>4</sup> Con anterioridad, López Morales (1995, 1999a) y Carcedo González (1998b) ya habían descrito el devenir histórico de la disponibilidad léxica.

un ambicioso proyecto panhispánico, el Proyecto de Estudio de la Disponibilidad Léxica en Español, dirigido por el profesor López Morales.

Los capítulos II (*Bases teóricas y metodológicas de la disponibilidad léxica*) y III (*Los centros de interés*) abordan la renovación epistemológica y metodológica que han experimentado los estudios de léxico disponible en el marco del Proyecto Panhispánico, dedicados a indagar y definir la norma comunitaria adulta de las distintas comunidades de lengua española. No obstante, la discusión metodológica no ha estado exenta de controversias, como la que afectó al diseño de la fórmula matemática utilizada en los estudios estadísticos de disponibilidad léxica (Lorán y López Morales, 1983; López Chávez y Strassburger, 1987), o la que se ha manifestado a través de las múltiples propuestas sobre la cantidad y naturaleza de los campos léxicos investigados<sup>5</sup>. Incluso los criterios de edición del Proyecto Panhispánico (Samper Padilla, 1998a: 311–313), que otorgan homogeneidad al proyecto, se han visto sometidos a frecuentes matizaciones y adaptaciones, como veremos más adelante.

Una parte inherente y constitutiva de los estudios de léxico disponible es el uso de herramientas y procedimientos de análisis estadístico, que incluyen tanto la estadística descriptiva como el uso de fórmulas especializadas (índice de disponibilidad, índice de cohesión, análisis multivariante, etc.). Como afirma Mateo García (1998: 14), “mueve a cierta sorpresa un análisis que coloca entre números, índices y magnitudes sociolingüísticas una realidad tan familiar y cotidiana como la abordada por la disponibilidad léxica”. En cualquier caso, este tratamiento científico sirve para descubrir las características estadísticas más o menos constantes entre los estudios realizados en diferentes comunidades que afectan tanto a la producción léxica de los centros de interés como al comportamiento de las variables sociolingüísticas consideradas. Estos aspectos son analizados en el mencionado capítulo III y en el V (*Los condicionantes extralingüísticos de la disponibilidad léxica*), a partir de la comparación de los resultados de las investigaciones españolas del Proyecto Panhispánico.

El capítulo IV (*Relaciones entre las palabras en el léxico disponible*) se inspira en los estudios de Hernández Muñoz (2006), y Ávila Muñoz y Villena Ponsoda (2010), e indaga en las líneas de investigación abiertas por estos estudiosos, encaminadas a descubrir la estructura del lexicón mental partiendo de los datos proporcionados por los estudios de disponibilidad, y mediante la aplicación de los modelos psicolingüísticos más relevantes en la producción del léxico. El capítulo muestra que el análisis de las asociaciones que se producen entre las palabras, sea

---

<sup>5</sup> Al respecto, entre otras, podemos adelantar la propuesta de Ayora Esteban (2006), quien manifiesta la necesidad, por un lado, de recabar información contextual y, por otro, de enunciar centros de interés que ayuden al investigador a compilar listas, no solo de sustantivos, sino también de otras categorías lingüísticas como adjetivos o verbos. A estas características corresponderían centros como ‘Los colores’, ‘Defectos y cualidades físicos y morales’ y ‘Acciones’, que la citada autora incorpora a su investigación.

en el interior de una lista léxica producida por un informante, sea dentro del conjunto léxico perteneciente a un colectivo, es una potente herramienta para deducir cuál es la realidad psicológica del léxico disponible. Estudios como el de Hernández Muñoz (2006), que consideran el léxico disponible como un fenómeno cognitivo complejo constituyen la aportación más seria a la consecución de la teoría integrada de la disponibilidad léxica preconizada por López Morales (1999a: 26), en la que se conjuguen factores como la memoria semántica, la familiaridad, etc.

El capítulo V (*Los condicionantes extralingüísticos de la disponibilidad léxica*) analiza la incidencia de los factores sociales en la producción léxica. Una de las características principales del Proyecto Panhispánico de Disponibilidad Léxica es la adopción de una perspectiva sociolingüística de corte variacionista, que intenta mostrar la estratificación sociolectal de la disponibilidad léxica de la comunidad estudiada. Según las normas comunes del Proyecto Panhispánico, las variables sociales obligatorias en cada investigación son ‘sexo’, (hombre/mujer), ‘centro escolar’ (público/privado), ‘zona geográfica’ (urbano/rural) y ‘nivel sociocultural’ (alto/medio/bajo). Este capítulo se corresponde con el análisis cuantitativo que de estas variables se hace en las diferentes investigaciones. En el presente estudio, mediante la comparación de los datos procedentes de estas investigaciones, se señalan las tendencias que con respecto a cada variable marcan los distintos estudios. Se analizan, además, otras variables, como ‘la edad’ y los condicionantes extralingüísticos que se tienen en cuenta en las situaciones de contacto de lenguas.

En el capítulo VI (*El análisis cualitativo del léxico disponible*) estudiamos una serie de factores lingüísticos que determinan la composición de los listados de léxico disponible que conforman los diccionarios que resultan de las distintas investigaciones. Como afirma García Marcos (1997: 25), además de los factores sociales, es necesario tomar en cuenta en los estudios de léxico disponible “los hechos estrictamente lingüísticos que, no sólo son –o pueden ser– responsables de la configuración que adquiera un determinado nivel de léxico disponible, sino que están en la base misma de la dinámica lingüística y, sociolingüística, por supuesto, de una determinada comunidad de habla”. La naturaleza de los hechos lingüísticos que analizamos permite diferentes clasificaciones de los vocablos que integran las listas, y que reflejan la variación (diatópica, diastrática, diafásica), la importación léxica, la formación y creación de palabras, etc.

Por último, antes de las conclusiones finales, el capítulo VII (*La comparación de los léxicos disponibles*) analiza la metodología empleada en las comparaciones interdialectales. El capítulo presenta, asimismo, el cotejo de cinco léxicos disponibles de España –fin último del Proyecto Panhispánico de Disponibilidad Léxica–, en lo que representa un ensayo para establecer la norma léxica del español de España. Este ejercicio nos sirve para mostrar lo que de común y diverso tienen estos léxicos disponibles, y para ilustrar lo que será la fase final del mencionado macroproyecto: la elaboración de diccionarios de léxico disponible que incluyan los resultados de las investigaciones precedentes.



# I. LOS ESTUDIOS DE DISPONIBILIDAD LÉXICA

## 1.1. La medición estadística del vocabulario

La aparición y desarrollo de los estudios de disponibilidad léxica se relacionan con la necesidad de optimizar el proceso de adquisición de la lengua materna, de una parte, y el aprendizaje como segunda lengua por hablantes extranjeros, de otra. Esta necesidad, en lo referente a la programación de la enseñanza del léxico, se ha traducido en la evolución y desarrollo de los procedimientos encaminados a hacer efectiva una adecuada selección de los vocablos que se deben enseñar tanto en las clases de lengua materna como en las de lengua extranjera.

¿Qué palabras hay que enseñar? ¿Cuáles no? ¿Por qué unas sí y otras no? A estas preguntas ha tratado de responder la léxico-estadística, y en la evolución de los cuatro procedimientos principales para la medición estadística del vocabulario existente en una comunidad de habla –los diccionarios normativos, la frecuencia léxica, el léxico básico y la disponibilidad léxica– (López Morales, 1986: 53–66) se pueden seguir los intentos de plasmar de forma precisa el uso real que los hablantes hacen de la lengua.

### 1.1.1. Los diccionarios normativos

De los diferentes tipos de diccionarios existentes (bilingües, de sinónimos, etimológicos, de colocaciones, de rimas, de dudas, etc.), son los diccionarios generales o normativos los que se ocupan de recoger de manera exhaustiva el caudal léxico de una lengua. En el caso del español, es el *Diccionario de la Real Academia Española*, conocido por sus siglas *DRAE*, la referencia obligatoria del patrimonio léxico para todos los hablantes de español. Su carácter de “depósito” léxico de la lengua española hace que la vigésima segunda, de 2003, cuente con unas 88.500 entradas o lemas. Con respecto a la vigésima primera

edición, de 1992, de estos lemas, unos 11.500 son novedades léxicas, entre ellas *zapear*, *liposucción* o *videojuego*, y desaparecen, porque realmente han muerto, otras seis mil palabras. Se añaden, al mismo tiempo, para un mejor reflejo de la realidad lingüística universal, numerosas marcas de americanismos. Se sobreentiende, por tanto, que en los once años que van de una edición a la otra, las nuevas palabras y acepciones ahora incorporadas ya eran antes usadas, al tiempo que muchas de las que se han eliminado habían caído en el olvido desde hacía varios años.

Ni qué decir tiene que “ningún hablante conoce todos los vocablos de un diccionario, ni siquiera todas las acepciones de los considerados como más frecuentes o con mayor rendimiento funcional” (Najt y Reizábal, 1989: 220). Al respecto, tomando al azar una página del *DRAE*, Humberto López Morales (1986: 53) ha clasificado el vocabulario que integra un *diccionario normativo*, según el grado de conocimiento y uso que un hablante hace de las palabras, en las siguientes clases:

a) palabras que nos son totalmente desconocidas, ya que ni las hemos leído ni escuchado en ninguna parte, como *empulgadura*, *empuntar*, *empurrarse*, *emuntorio*, etc.;

b) palabras que nos parecen conocidas, pero que en realidad no sabemos bien lo que significan, como *emulgente*, *emulsivo*, *emulsor*, *enacerar*, etc.;

c) palabras que sí conocemos y entendemos, pero que no utilizamos –léxico pasivo– al hablar o escribir, como *emulación*, *emulador*, *enajenable*, *enanismo*, etc.;

d) palabras que en realidad conocemos, no solo porque las oímos o leemos con frecuencia, sino porque nosotros mismos las usamos –léxico activo–, como *empuñadura*, *enamorado*, *enamoradoizo*, *enamorar*, *enano*, etc.

A continuación, el mismo autor hace un listado de los vocablos que se apartan del léxico general al que, en su afán de exhaustividad, los diccionarios dan cabida a. Entre estos vocablos se hallan:

– arcaísmos, palabras anticuadas que se usaron en otra época pero que ya han caído en el olvido, como *empulgar* (‘armar la ballesta’), *enaciyar* (tratar las lanas con aceite, acije o aceite de vitriolo’);

– regionalismos, palabras que solo se conocen en alguna zona o país del mundo hispánico, como *empuntar* (‘encarrilar, encaminar, dirigir’, en Colombia, Ecuador y El Salvador), *engrudo* (‘cola de pegar’, en Canarias);

– tecnicismos, palabras científicas manejadas solo o principalmente por especialistas, como *emunción* (‘evacuación de los humores y materias superfluas y nocivas’), *enálage* (‘figura que consiste en mudar las partes de la oración o sus accidentes; como cuando se pone un tiempo del verbo por otro, etc.’);

– cultismos, palabras empleadas muy ocasionalmente por hablantes de gran cultura y de amplio dominio léxico, como *enajenar* (‘pasar o transmitir a otro el

dominio de una cosa o algún otro derecho sobre ella’), *enalbar* (‘caldear y encender el hierro en la fragua tanto, que parezca blanco de puro resplandeciente’).

El listado anterior muestra que los diccionarios normativos constituyen un buen depósito de todo –o casi todo– el caudal léxico de una lengua. Sin embargo, –concluye López Morales (1986: 54) no son adecuados pedagógicamente<sup>1</sup> para delimitar y enseñar las palabras “que tienen realmente vida común entre los hablantes de una comunidad”.

### 1.1.2. La frecuencia léxica

La exhaustividad de los diccionarios normativos, que recogen todo el caudal léxico de una lengua se vio depurada por las listas de frecuencia, compuestas por los vocablos que se repiten con más frecuencia en los textos de una comunidad. La *frecuencia léxica*, se basa en el índice ‘frecuencia’, que indica el número de apariciones de los vocablos (o palabras diferentes) en los textos computados. Estos vocablos se ordenan según rangos, es decir, estableciendo una jerarquía de acuerdo a la cantidad de veces que se repiten en un texto. Si se toman muchos textos, orales y escritos, representativos de la lengua de un país o comunidad, se llegaría a tener una lista de las palabras *realmente* usadas allí, jerarquizada por rangos de aparición, en definitiva, una lista de frecuencia.

“Las palabras más frecuentes son, principalmente, las que sirven para mantener la estructura del mensaje, mientras que las voces realmente portadoras de información se repiten relativamente poco en el discurso” (Alvar Ezquerro, 2005: 22). Tanto es así que, según Alvar Ezquerro (2003: 100) “en español, las 100 palabras más frecuentes representan más del 30% del material léxico que constituye cualquier texto; con las 1000 más frecuentes se llega a más del 50% de todo texto; con las primeras 5000 se alcanza más del 90%”. Las palabras más frecuentes son de tipo gramatical (preposiciones, determinantes, pronombres, conjunciones, adverbios), además de los nombres de los objetos inmediatos y los verbos y adjetivos elementales. Ello se puede apreciar en la lista de las 160 palabras más frecuentes (véase Cuadro 1.1) del *Diccionario de frecuencias de las unidades lingüísticas del castellano* de Alameda y Cuertos (1995), que maneja un corpus de textos de 2.000.000 de palabras del periodo 1980–1995:

---

<sup>1</sup> Un buen diccionario para la enseñanza de L2 debe contener, según Alvar Ezquerro (2003: 104–105), información sobre: a) división silábica, b) pronunciación, c) léxico usual en la lengua escrita y hablada, d) definiciones claras, breves y no sinónimas, e) abundantes ejemplos de uso, f) uso gramatical y estilístico, g) familias de palabras, h) ilustraciones contextualizadoras, i) pesos, medidas, nombres geográficos y otros elementos culturales.

**Cuadro 1.1. Las 160 palabras más frecuentes del español, según Alameda y Cuetos (1995)**

de	116480	qué	3447	estaba	1925	hacer	1292
la	78770	ni	3395	menos	1891	cosas	1274
que	67636	entre	3370	ese	1854	cómo	1271
y	60094	muy	3233	sino	1841	veces	1249
el	54027	hasta	3146	día	1798	todas	1232
en	49542	nos	3049	casa	1775	ver	1219
a	44456	porque	3019	mundo	1775	madre	1209
los	31399	ella	2974	han	1740	noche	1209
se	28254	son	2897	esa	1736	sido	1209
un	25119	sólo	2883	poco	1725	mejor	1194
no	24135	dos	2865	después	1695	he	1174
las	21768	vez	2837	donde	1692	esto	1164
con	20405	también	2789	aunque	1687	luego	1151
una	19973	él	2725	tanto	1673	toda	1150
por	19723	tan	2702	tiene	1671	aquí	1148
del	19170	desde	2683	pues	1645	hecho	1140
su	16533	así	2515	sí	1621	aquella	1133
lo	15375	te	2498	dijo	1607	momento	1110
es	14453	este	2459	(50%)	–	tres	1106
al	12290	bien	2454	hace	1601	mano	1099
como	11958	hay	2428	antes	1585	quien	1080
para	11493	siempre	2425	entonces	1564	mis	1073
más	9401	esta	2368	les	1532	ante	1072
o	8945	vida	2293	hacia	1525	misma	1053
me	8648	todos	2287	unos	1512	otras	1039
pero	8226	fue	2265	mucho	1504	tu	1039
le	7908	años	2255	ojos	1466	sea	1033
sus	7057	tiempo	2226	e	1441	forma	1031
si	5856	eso	2163	nunca	1423	padre	1027
sin	5324	ahora	2145	tenía	1417	fin	1023
era	5201	mismo	2138	decir	1416	durante	1018
había	5161	nada	2097	casi	1410	mas	1011
ya	4772	otro	2096	otros	1383	nadie	1006
cuando	4738	está	2050	ellos	1369	habían	1003
mi	4499	cada	1993	mientras	1354	allí	998
todo	4389	uno	1988	tal	1350	cuerpo	998
ha	4144	puede	1974	aquel	1336	jugar	998
sobre	4106	algo	1969	mujer	1335	contra	974
o	3697	hombre	1965	parte	1300	aún	964
ser	3467	otra	1953	gran	1292	tarde	960
						caso	953

La historia de las listas de frecuencia ha pasado por una serie de reajustes y revisiones del marco teórico y la metodología de las investigaciones (lengua oral o lengua escrita, tipo y tamaño de las muestras, volumen léxico del análisis, marco de la comunicación, ponderación de la frecuencia, organización de las listas, etc.). No obstante, el objetivo fundamental de estos recuentos ha sido aplicarlos a la enseñanza estructural y sistemática del vocabulario –bien de una lengua extranjera (en la mayoría de los casos), bien de la lengua materna.

Carcedo González (2000a: 11–15) presenta una amplia visión cronológica de diversos diccionarios de frecuencias en diferentes idiomas. La lista de diccionarios de frecuencia con intención estrictamente pedagógica incluye:

– Para el inglés: El *Teacher's Word Book*, de Thorndike (1921), formado por 10000 palabras inglesas con alta frecuencia de aparición en diversos textos a partir de un corpus de 4.500.000 de palabras, y el *Inglés B.A.S.I.C.*<sup>2</sup>, de Ogden (1944) y Richards (1968), con las 850 palabras más usuales y de mayor expresión que aparecían en las definiciones de los diccionarios. Este *Inglés Básico*, aunque gozó de extraordinaria popularidad para la enseñanza rápida del inglés, fue criticado por obligar a sus usuarios a expresarse en una forma ajena al inglés corriente.

– Para el francés: El diccionario de frecuencia de Henmon (1924) de las 5000 palabras más usuales del francés, partiendo de una muestra de 400.000 palabras tomadas de textos de la literatura francesa de la segunda mitad del siglo XIX. Este diccionario fue revisado por Vander Beke en 1935, quien publicó un diccionario de 6.067 palabras, en el que supeditaba la frecuencia al rango, al considerar más importante una palabra que aparece una vez en varios textos que una que aparece varias veces en un mismo texto. Por su parte, M. Aristizabal publicó en 1938 un vocabulario de frecuencia mixta de 12.038 vocablos, basado en 460.727 palabras recogidas en textos escritos por adultos y niños, con el objetivo de hacer más eficaz el aprendizaje de la lengua, especialmente de la ortografía en las escuelas.

– Para el español: El *Spanish Word Book* de Buchanan (1927) para la enseñanza del español a anglohablantes, una lista, que combinaba la frecuencia y el rango, con las 5234 palabras españolas más usadas en un corpus de 1.200.000 palabras, repartidas en 40 textos que incluían obras dramáticas, novelas, poesías, folklore, prosa miscelánea, textos técnicos y textos periodísticos. Posteriormente, con el objetivo de la enseñanza de la lengua materna, el *Recuento de vocabulario español* (Rodríguez Bou, 1952) y *Vocabulario usual, común y fundamental* (García de Hoz, 1952). En el primer caso, Rodríguez Bou se sirve de más de 7.000.000 de palabras de material muy heterogéneo, escrito y oral, cuyas fuentes son composiciones y frases de escolares, y textos extraídos de los periódicos, la radio, obras religiosas, pedagógicas y libros de texto. En el caso de García de Hoz, el corpus contaba 400.000 palabras extraídas de cuatro ámbitos: vida privada (cartas privadas), vida social indiferenciada (periódicos), vida social regulada (textos

---

<sup>2</sup> B.A.S.I.C.: British American Scientific International Commercial.

administrativos gubernamentales y eclesiásticos), vida cultural (textos literarios). Sirviéndose de la frecuencia y el rango, García de Hoz distingue 12.402 palabras usuales, 1.971 comunes y 208 fundamentales.

– Con respecto a otras lenguas, cabe mencionar el diccionario de ruso para estudiantes de lengua inglesa de H. Josselson (1953). Presenta una lista con las 204 palabras más frecuentes, a partir de 100.000 palabras tomadas de textos periodísticos, novelas y textos literarios de otros géneros desde 1830 hasta el primer tercio del siglo XX.

En la actualidad, el estudio de la frecuencia léxica continúa siendo un parámetro fundamental en los estudios léxico-estadísticos. Con el desarrollo paralelo de la lingüística del corpus y de la capacidad de los ordenadores y programas informáticos, las posibilidades de procesar corpus de millones de palabras han aumentado, al tiempo que los listados de frecuencia son cada vez más fiables. Muestra de ello son los listados de frecuencias del volumen *Frecuencias del español: diccionario y estudios léxicos y morfológicos* (Almela Pérez et al., 2005), basado en el Corpus Cumbre, de 20 millones de palabras referidas al lenguaje hablado y escrito en España y en todos los países de habla hispana, o los listados de frecuencias que ofrece la Real Academia Española en su página web, a partir del *Corpus de Referencia del Español Actual (CREA)*, el cual, en su versión (3.2, junio de 2008), cuenta con algo más de 160 millones de formas, procedentes de una amplia variedad de textos escritos y orales, producidos en todos los países de habla hispana desde 1975 hasta 2004.

Volviendo al uso que los hablantes hacen del léxico de una lengua, la frecuencia léxica, entre otras cosas, ha puesto de manifiesto las limitaciones de los hablantes a la hora de manejar el caudal léxico de su propia lengua, puesto que las listas muestran que son pocas las palabras que alcanzan una frecuencia alta. Es más, las investigaciones llevadas a cabo en Francia entre los años cincuenta y sesenta (Guiraud, 1954, 1959; Herdan, 1956, 1962; Verléé, 1963; Müller, 1964) demostraron que, dejando a un lado los términos de la especialidad de cada cual –los vocabularios de profesiones y oficios– el hombre culto, el intelectual, maneja entre 4000 y 5000 vocablos, en tanto que el hombre común no emplea más de 2000 vocablos.

### 1.1.3. El léxico básico

La identificación de esas 4000 ó 5000 palabras de uso real entre los hablantes cultos de una comunidad lingüística, y la delimitación de las 2000 palabras manejadas por el nivel sociocultural bajo nos lleva al concepto de *léxico básico* (López Morales, 1986: 59–62).

Los léxicos básicos perfeccionan la selección léxica al ponderar la frecuencia con la dispersión, ya que divide el universo léxico en diferentes mundos limitados,

bien por el contenido de los textos, bien por las condiciones de los mismos. La dispersión está en relación directa con la cantidad de mundos en que aparece un vocablo y la estabilidad de la frecuencia en ellos. Al ponderar la frecuencia con la dispersión, se obtiene un tercer índice: el uso real de cada palabra en la comunidad lingüística que se estudia.

López Morales (1986: 62) ejemplifica la diferencia entre léxico frecuente y léxico básico a partir de los listados del *Léxico básico del Español de Puerto Rico*. Así, entre las primeras 100 palabras de este *Léxico básico*, hay 13 en que no coinciden la frecuencia con el uso. Las palabras *aquí, dejar, hijo, mano, más, mirar, mujer, oír, usted, tú, vivir, volver* y *salir* son más frecuentes; por el contrario, *mucho, cada, como, durante, encontrar, esto, menos, mucho, mundo, persona, poner, seguir* y *tres* “son más usuales porque a pesar de tener menos frecuencia, poseen mayor dispersión, aparecen en más mundos”.

El primer léxico básico es el *Frequency Dictionary of Spanish Words*, de Juilland y Chang-Rodríguez (1964). Este diccionario está compuesto por 5.024 palabras seleccionadas de un universo léxico integrado por 500.000 unidades, repartidas a partes iguales en cinco mundos léxicos: 1) obras dramáticas, 2) obras de ficción, novelas e historias breves, 3) obras en prosa, ensayos, cartas y memorias, 4) textos periodísticos, y 5) textos técnicos. Para asegurar la homogeneidad dialectal, cronológica y cultural de ese universo léxico, se seleccionaron únicamente textos no marcados dialectalmente escritos exclusivamente en la Península Ibérica entre 1920 y 1940, en tanto que para controlar la variedad sociocultural y estilística sólo se admitió el vocabulario adulto estándar. Más tarde, y de manera sucesiva, Juilland en compañía de otros autores completaría una colección de léxicos básicos dedicados a diferentes lenguas románicas: rumano (Juilland, Edwards y Juilland, 1965), francés (Juilland, Brodin y Davidovitch, 1971) e italiano (Juilland y Traversa, 1973).

Aunque con algunos retoques en el procedimiento estadístico, el empleo de corpus escritos, de cinco subcategorías o mundos, el tamaño del corpus manejado, así como su base teórica convirtieron a las investigaciones de Juilland en modelos para la elaboración de posteriores léxicos básicos. Tal es el caso del léxico elaborado por Bortolini, Tagliavini y Zampolli (1971) para el italiano, el de Puerto Rico (Morales, 1986), y dos proyectos de investigación aún no concluidos sobre el léxico básico de España, *LEBAES*, y de las Islas Canarias, *LEBAICan* (Benítez Pérez, Hernández Cabrera y Samper Padilla, 1995).

En palabras de López Morales (1986: 62), “el léxico básico nos ofrece un perfil muy adecuado del vocabulario usual de una comunidad lingüística dada, lo que sin duda constituye un valiosísimo instrumento para la planificación léxica de la escuela”. De hecho, Mena Osorio (1986: 7–8) afirma que, por estar constituido por los ítems léxicos utilizados por la gran mayoría de los hablantes, el léxico básico o “vocabulario de base es el punto de partida para la planificación de la enseñanza de la lengua materna”.

Aparte de esta función planificadora, los léxicos básicos pueden tener otras utilidades. En la enseñanza de español a extranjeros los léxicos básicos “pueden ser material para el estudio en casa o instrumento de control a medio o largo plazo para comprobar el dominio del vocabulario” (Fajardo Aguirre, 1993: 221–222). Pero, dada su ordenación temática, necesitan ser acompañados de un libro de ejercicios adecuado para trabajar los diferentes campos léxicos. Fajardo sugiere, además, que se use el vocabulario básico para la elaboración de las definiciones de los diccionarios usados en la enseñanza de ELE para facilitar la comprensión de estas definiciones.

#### 1.1.4. La disponibilidad léxica

A pesar de su potencialidad pedagógica, el léxico básico demostró pronto ser un producto incompleto para la planificación de la enseñanza del léxico. En efecto, cuando a principios de los años cincuenta M. G. Gougenheim, P. Rivenc, R. Michéa y A. Sauvegeot examinaban las listas de frecuencia para preparar *Le Français Élémentaire* (1954), observaron que ciertas palabras muy conocidas y usadas por los franceses alcanzaban una frecuencia bajísima (*auto, metro, fête* o *garage*) o ni siquiera aparecían (*fourchette, coude, dent, jupe, leerte* o *timbre*), por lo que, por motivos estadísticos, no debían formar parte de la lengua básica. Esta conclusión era del todo inaceptable: quedaba de manifiesto la parcialidad de un léxico seleccionado sólo a partir del criterio de la frecuencia.

Surgió así el concepto de *disponibilidad* léxica, que “señala el vocabulario concreto que el hablante podría utilizar en un contexto comunicativo dado” (López Morales, 1986: 63). La disponibilidad está, pues, ligada al tema tratado. Así, como ejemplifica el mismo autor, “es poco probable que, a no ser que se hable de los medios de transporte público, se mencionen palabras como *autobús*, y que salgan en nuestra expresión palabras como *carta* y *sello* si no nos referimos específicamente a asuntos del correo”.

Una palabra disponible es, según Michéa (1953: 310), “una palabra que, sin ser particularmente frecuente, está lista para ser empleada y viene inmediatamente y naturalmente a la mente en el momento en que se tiene necesidad de ella”. En este sentido, “un estudio de disponibilidad léxica implica centrarse en el plano de la producción lingüística” (Mena Osorio, 1986: 11), en el vocabulario activo.

Para recaudar este vocabulario disponible, pero no frecuente, se preparan pruebas asociativas en torno a estímulos temáticos, los “centros de interés” (‘el cuerpo humano’, ‘la ropa’, ‘alimentos’, etc.). Así se recogen las palabras de las que el hablante dispone realmente si su expresión girara en torno a estos temas.

El *Français Fondamental*, de Gougenheim *et al.* (1958), fue la primera obra que aplicó el método de la disponibilidad léxica. Se trata de un vocabulario integrado, en primera instancia, sólo por 1475 palabras, de las cuales 1222 son ítems léxicos y el resto, ítems gramaticales. Para la delimitación del léxico disponible se realizaron 904 encuestas (488 chicos y 416 chicas) en cuatro departamentos representativos franceses alejados entre sí geográficamente a escolares de 9 a 12 años, ya que se consideraba que con esa edad los escolares ya poseían un vocabulario común, pero todavía libre de las influencias del vocabulario especializado de quien tiene una profesión (Mackey, 1971a: 30). Los encuestados debían escribir listas de 20 palabras por cada uno de los 16 centros de interés que luego han sido clásicos en la bibliografía. Para representar los sectores semánticos más extendidos en toda lengua, estos centros debían a un tiempo: a) ser universales, en el sentido de que pudieran interesar a cualquier hablante de cualquier estrato, situación o condición, y b) mantener cierta coherencia interna, es decir, ser fácilmente diferenciables.

De inmediato se iniciaron estudios similares en el Canadá francoparlante que, finalmente, se ajustaron fielmente a la metodología empleada en Francia para garantizar la comparabilidad de los resultados de las encuestas canadienses con los obtenidos en Francia. De esta manera, Mackey (1971b) pudo medir las diferencias culturales entre el vocabulario común de Francia y el de la región canadiense de l'Acadie, determinando diferencias y similitudes léxicas, y las palabras exclusivas de cada vocabulario.

Los resultados de la disponibilidad léxica contrastan en gran medida con el léxico recogido en las listas de frecuencias. Así, mientras que en listas de vocabulario frecuente las palabras más frecuentes son las palabras gramaticales (artículos, pronombres, preposiciones, etc.) y, en orden decreciente, los verbos, los adjetivos y por último los nombres, en las listas de vocabulario disponible aparecen nombres casi en su totalidad. De acuerdo con este contraste, Michéa (1950) estableció la distinción entre palabras “temáticas” y “atemáticas”; las primeras principalmente de contenido semántico muy concreto, sólo aparecían cuando se trataba de un tema o un cierto tipo de temas; las segundas, por el contrario, aparecían con gran regularidad, independientemente del tema tratado, y eran principalmente palabras gramaticales: adjetivos, verbos y sustantivos de carácter general.

Estas consideraciones llevaron a Gougenheim *et al.* (1956: 142) a plantear la necesidad de elaborar dos vocabularios, uno frecuente y otro disponible. Según López Morales (1986: 63), en el “vocabulario frecuente” se integrarían las clases de palabras más estables en la lengua, esto es, las que pueden aparecer en una mayor cantidad de contextos y situaciones diferentes; y en el “vocabulario disponible” las palabras más inestables, de baja frecuencia –escrita u oral–, pero

en todo caso conocidas y útiles. El léxico básico recogería para López Morales (1986, 1995) la norma culta, formalmente difundida a través de los canales de comunicación que toda sociedad tiene habilitados para la producción escrita, organizada en torno a una serie de mundos temáticos que presuponen una evidente formalidad comunicativa. Por el contrario, el léxico disponible se refiere a situaciones muy cotidianas para los hablantes, estando menos determinado por la formalidad cuando se actualiza en el discurso, y perteneciendo más bien a parcelas léxicas orales no reflejadas en la literatura. “De ahí que la convergencia de ambos [léxicos] en el *léxico fundamental* de una determinada comunidad de habla sea considerada como uno de los mejores baluartes con los que puede contar una tarea de planificación lingüística” (Mateo García, 1998: 38).

El diccionario de léxico disponible de una comunidad es, en palabras de Galloso Camacho y Prado Aragonés (2004: 372), “un auténtico diccionario de uso, no solo fiable por la metodología empleada, sino por su actualidad” [...] “con el valor añadido de que se trata de un trabajo que proporciona información dialectológica, sociolingüística y de frecuencia”. Se trata de un diccionario de «palabras temáticas», vinculadas fuertemente con determinados ámbitos conceptuales, que incluye sobre todo sustantivos, bastantes verbos, y pocos adjetivos.

## **1.2. El estudio de la disponibilidad léxica en español**

Los estudios de disponibilidad léxica nacieron en Francia, pero ha sido en el ámbito hispánico donde se han desarrollado con más fuerza, primero en tierras hispanoamericanas, después extendiéndose exitosamente por España. Este apartado analiza el devenir de estos estudios para el español, en los que el *Proyecto Panhispánico de estudio sobre la Disponibilidad Léxica* constituye un punto de inflexión, gracias a la homogeneización y unificación de los criterios de diseño metodológico que permiten el intercambio de datos y el desarrollo de estudios comparados entre los resultados de los diferentes proyectos locales.

### **1.2.1. Los estudios anteriores al Proyecto Panhispánico**

En el ámbito hispánico la disponibilidad léxica se inicia con la investigación de H. López Morales (1973) sobre la disponibilidad léxica en escolares de San Juan. Para su estudio utilizó a 63 alumnos de escuelas de San Juan de Puerto Rico y diez centros de interés. López Morales analiza la influencia de las variables ‘nivel socioeconómico’ y ‘nivel escolar’ sobre el vocabulario de los alumnos de enseñanza básica. Más tarde utilizaría los datos de este estudio

para reflexionar sobre la programación del léxico en la escuela (1978), y discutir acerca de la *teoría del déficit* de Bernstein (1971, 1975).

López Morales es, por tanto, el pionero, impulsor (y coordinador) de los estudios de disponibilidad léxica en español, habiéndose convertido su investigación sobre el *Léxico disponible de Puerto Rico* (1999a) en modelo de otros estudios en el mundo hispánico. Antes y después de esa investigación, López Morales publicó una larga lista de trabajos (1979, 1983, 1986, 1991b, 1993, 1995, 1996, 1999b, 2001, 2005, 2008, 2009), en los que –bien como objeto exclusivo de atención, bien junto al estudio de otros aspectos de la lengua– demuestra el enorme campo de aplicaciones de la disponibilidad léxica para la lingüística y sus disciplinas colindantes.

El magisterio del profesor López Morales inspiró desde mediados de los años 80 distintas investigaciones sobre disponibilidad léxica en América (México, Chile y Costa Rica) y España, si bien estas investigaciones no presentan unas normas metodológicas unitarias: la nómina de centros de interés es variada en la cantidad y en la composición, la selección de la muestra incluye diferentes grupos de edad y nivel de escolarización, y el tamaño de las muestras es muy diverso, como se muestra en el Cuadro 1.2.

De las investigaciones desarrolladas fuera del Proyecto Panhispánico destacan las siguientes aportaciones:

- En México, los trabajos de López Chávez, quien, en colaboración con Strassburger Frías (1987, 1991), desarrolló la fórmula matemática que se utiliza actualmente para el cálculo de la disponibilidad, así como también inició la senda de los cotejos dialectales con los análisis de los índices de compatibilidad de los léxicos disponibles de Puerto Rico, República Dominicana, Madrid y Gran Canaria (López Chávez, 1992, 1995b).

- En Chile, la labor investigadora de Echeverría (1991, 1993) y Valencia (1993, 1994a, 1994b, 1997, 1998a, 1998b, y posteriores) que culminó con la publicación del volumen titulado *Disponibilidad léxica en estudiantes chilenos* (Valencia y Echeverría, 1999), un vasto trabajo con fines pedagógicos, basado en una muestra de 2052 estudiantes, que aporta una radiografía precisa y detallada de la disponibilidad léxica a nivel nacional chileno, muy útil para profesores de español y planificadores pedagógicos del país.

- En el caso de Costa Rica, Murillo Rojas (1993, 1994, 1998) y Sánchez Corrales y Murillo Rojas (1993, 1997, 1999, 2002, 2004) han desarrollado diversas investigaciones sobre la disponibilidad léxica de alumnos de preescolar. Fruto de estas investigaciones ha sido la publicación del diccionario de disponibilidad léxica de niños de entre cinco años y medio y seis años y medio de todas las regiones del país (Sánchez Corrales y Murillo Rojas, 2006).

Cuadro 1.2. Estudios de disponibilidad léxica realizados fuera del Proyecto Panhispanico

Investigación	Muestra/Encuesta	VARIABLES	Centros de interés
López Morales, 1973: San Juan, Puerto Rico.	63 escolares: 1°, 3°, 5° grado.	Sexo, Nivel sociocultural y Nivel escolar.	10: Alimentos; Animales; La casa; La cocina; El cuerpo humano; Materiales de construcción; Muebles; Naturaleza; La ropa; Transportes.
Román Morales, 1985: Dorado, Puerto Rico.	293 escolares: 1°, 3°, 5° grado.	Sexo, Nivel escolar, Zona geográfica, Nivel socioeconómico.	10: Los mismos que López Morales (1973).
Butrón, 1987: Puerto Rico.	154 escolares: 5° grado.	Sexo, Nivel socioeconómico y Procedencia geográfica.	8: Alimentos; Animales; La casa; La ropa; Naturaleza; Colores; El cuerpo humano; Juegos y entretenimientos.
Justo Hernández, 1986: México, D.F.	54 alumnos: Preuniversitarios. Listas cerradas: 10 palabras.	Tipo de centro, Ubicación del centro, Sexo.	10: centros para investigar sobre los Colores; Ropa; Zapatos; Cosméticos; Automóviles; Anuncios comerciales; Mochilas, bolsas y maletas; Muebles; Tapices, alfombras; Pintura de casa; Vestido.
Cañizal Arévalo, 1987: México, D.F.	68 informantes.	Sexo, Tipo de escuela, Tiempo dedicado a ver la televisión. Agrupaciones léxicas.	6: Partes del cuerpo; Objetos que hay en el interior de una casa; Alimentos y bebidas; Medios de transporte; Profesiones y oficios; Defectos físicos y morales.
Ruiz Basto, 1987: México.	120 alumnos: Bachillerato.	Sexo, Tipo de escuela.	6: Los mismos que Cañizal Arévalo (1987), con la modificación "Defectos y cualidades físicas y morales".
López Chávez, 1993: México, D.F.	216 alumnos: Preescolar, 6° grado de Primaria.	-	16 centros de interés clásicos.
Mena Osorio, 1986: Concepción, Chile.	300 escolares: 4°, 6°, 8° grado.	Sexo, Nivel educativo, Nivel sociocultural.	10: Los mismos que López Morales (1973), pero reformula "Materiales de construcción" como "Herramientas y materiales de construcción".
Echeverría <i>et al.</i> , 1987: Concepción, Chile.	400 alumnos: 1° y 3° E. Media.	Nivel escolar, Sexo, Nivel socioeconómico	10: Los mismos que López Morales (1973), pero reformula "Materiales de construcción" como "Herramientas y materiales de construcción".

Echeverría, 1991: Chile.	300 escolares: 4°, 6°, 8° grado. 400 alumnos: 1°, 3° E. Media.	Sexo, Procedencia sociocultural, Nivel de escolaridad.	10: Los mismos que López Morales (1973), pero reformula “Materiales de construcción” como “Herramientas y materiales de construcción”.
Vargas Sandoval, 1991; González Becker y Vargas Sandoval, 1992: V región de Chile.	500 hablantes (20–50 años).	Sexo, Nivel escolar, Nivel socioeconómico.	6 (abstractos): Arte, Derecho, Economía, Política, Sentimientos, Teología.
Valencia y Echeverría, 1999a: Chile.	2052 estudiantes: 4° E. Media.	Sexo, Nivel sociocultural, Régimen educativo, Regionalidad, Procedencia poblacional.	18: Las partes del cuerpo; Ropa y calzado; La casa; Muebles; Alimentos; La cocina; Medios de transporte; Trabajo de campo y jardín; Plantas y animales; Juegos y entretenimientos; Profesiones y oficios; La política; La actividad económica; Las artes; Ciencia y tecnología; Mundo espiritual; Procesos mentales; Problemas del ambiente.
Murillo Rojas, 1993, 1994; Sánchez Corrales y Murillo Rojas, 1993, 2006: Costa Rica.	50 niños: Preescolar. 10 láminas motivadoras.	Sexo, Lugar de procedencia, Ocupación del padre.	6: Partes del cuerpo; Familia; Animales; Medios de transporte; Profesiones y oficios; Comidas y bebidas.
Ríos González, 2010a: Costa Rica.	512 alumnos: Preuniversitarios.	Sexo, Ubicación del centro educativo, Tipo de centro educativo, Nivel socioeconómico.	14: Saludos; Temas de conversación; Juegos y diversiones; Sexo; Tecnología; Ropa; Despedidas; Música; Medios de comunicación; La familia; Delincuencia y asuntos ilegales; Vida diaria; La muerte; Léxico juvenil.
Azumendi Ayerbe, 1983: País Vasco.	1224 informantes. Listas cerrada: 20 palabras.	Cantidad de producción, Categoría léxica, Distancia interlingüística, Corrección interlingüística, Divisiones semánticas. Factor geográfico, Edad, Sexo, Tipo de escolarización, Lengua más utilizada en casa, Lengua más utilizada con los amigos y lengua utilizada en la encuesta.	15: Partes del cuerpo; Prendas de vestir y de calzar; Partes y objetos que hay en una vivienda; Alimentos y bebidas; Relaciones familiares; Ciudad o pueblo; Los transportes; Enseñanza; Trabajo; Juegos, diversiones y entretenimientos; Deportes; Acciones que normalmente realizan todos los días; La naturaleza, las vacaciones; Agricultura y vegetales. Ganadería y animales. Pesca y peces; Dinero, bancos y comercio.

### 1.2.2. El Proyecto Panhispánico de Disponibilidad Léxica

El Proyecto Panhispánico de Disponibilidad Léxica, que dirige y coordina Humberto López Morales, es el marco teórico y metodológico en el que se desarrollan los estudios de disponibilidad léxica en español. En marcha desde finales de los años 90<sup>3</sup>, cuenta con equipos inscritos en las diferentes zonas dialectales del mundo hispánico: investigadores de todas las comunidades de España (por orden de incorporación al proyecto, Madrid, Canarias, Andalucía, País Vasco, Castilla y León, Asturias, Valencia, Aragón, Galicia, Cataluña, Cantabria, Castilla-La Mancha, La Rioja, Navarra, Murcia, Extremadura e Islas Baleares) y de buena parte de Hispanoamérica (Puerto Rico, República Dominicana, Chile, Costa Rica, Uruguay, Paraguay, Cuba, Argentina, Colombia, etc.).

El objetivo principal del Proyecto es elaborar Diccionarios de Disponibilidad Léxica (por centros de interés, por variables, etc.) para las diversas zonas del mundo hispánico. Las diferentes jornadas de coordinación celebradas en la Universidad del País Vasco (1999), en San Millán de la Cogolla (2003 y 2005), y en la Universidad de Salamanca (2011) han servido para que los investigadores en disponibilidad léxica llegaran a acuerdos metodológicos que garantizan dentro del proyecto la homogeneidad de criterios, en cuanto al tipo de encuesta, los campos léxicos que se toman en consideración, los informantes, las variables analizadas, los criterios de edición de los listados y el tratamiento matemático de los datos. Todo ello permite establecer comparaciones de tipo lingüístico, etnográfico y cultural entre los resultados de las comunidades de habla estudiadas, dibujar áreas de difusión y, en general, servir de punto de partida para análisis posteriores.

Las fases para la elaboración de cada diccionario de disponibilidad son:

- a) Selección de la muestra de informantes.
- b) Recogida de materiales mediante una encuesta.
- c) Transcripción de los materiales y su homogeneización mediante los protocolos de edición convenidos.
- d) Cálculo de la disponibilidad léxica global y por campos y grupos de informantes, lo que llevará a configurar los diversos apartados del diccionario.
- e) Publicación del diccionario.

Los 16 centros de interés obligatorios en el proyecto panhispánico son:

1. 'Partes del cuerpo humano'.
2. 'La ropa'.
3. 'Partes de la casa (sin los muebles)'.

---

<sup>3</sup> Samper Padilla *et al.* (2003: 28) dan cuenta del carácter fundacional de la reunión en la Universidad del País Vasco en Bilbao en 1999, que supuso el empuje definitivo para el Proyecto Panhispánico de Disponibilidad Léxica. Allí se reunieron equipos de universidades que iban a empezar sus investigaciones (Salamanca, Cádiz, Castellón), con otros con investigaciones ya en curso (Puerto Rico, Las Palmas de Gran Canaria, el País Vasco). Todos ellos, debatieron y perfilaron los principios metodológicos del proyecto, al que desde ese momento definitivamente se le dio carácter coordinado.

4. ‘Los muebles de la casa’.
5. ‘Alimentos y bebidas’.
6. ‘Objetos colocados en la mesa para la comida’.
7. ‘La cocina y sus utensilios’.
8. ‘La escuela: muebles y materiales’.
9. ‘Iluminación, calefacción y medios de airear un recinto’.
10. ‘La ciudad’.
11. ‘El campo’.
12. ‘Medios de transporte’.
13. ‘Trabajos del campo y del jardín’.
14. ‘Los animales’.
15. ‘Juegos y distracciones’.
16. ‘Profesiones y oficios’.

Las muestras analizadas están compuestas por estudiantes del nivel preuniversitario (último curso del bachillerato) a fin de que los resultados no se vean contaminados por los tecnicismos adquiridos en el nivel universitario, los cuales se alejan de la norma de la comunidad, según López Morales (1999a: 28) y Samper Padilla *et al.* (2003: 65). El patrón es mantener homogénea la edad y el nivel cultural de los informantes (estudiantes del curso inmediatamente anterior a la entrada en la universidad) y subcategorizarlos de acuerdo con una serie de variables. Estas variables obligatorias en el Proyecto Panhispánico incluyen ‘sexo’ (varón/mujer), ‘tipo de centro’ (público/privado), ‘ubicación del centro’ (urbano/rural), ‘nivel sociocultural’ (para cuyo establecimiento se tiene en cuenta los estudios y la profesión de los padres).

Las normas de edición comunes a todo el Proyecto Panhispánico han sido recogidas sistemáticamente por Samper Padilla (1998a: 311–313). De ellas trataremos en § 4.3.

Por lo que respecta a los trabajos publicados, en Hispanoamérica destaca El *Léxico disponible de Puerto Rico*, de López Morales (1999a), culminación de los estudios de disponibilidad léxica en esta isla, iniciados ya en 1973. Este trabajo se ha convertido en modelo para los siguientes estudios.

Previamente había visto la luz el *Léxico disponible de la República Dominicana*, de Alba (1995a). Junto con el estudio de López Morales (1999a), pronto se convirtió en referencia obligada para las comparaciones de tipo dialectal entre diferentes sintopías del mundo hispánico (Alba, 1998, 2000b; Bartol Hernández, 1998; González Martínez, 2001; Carcedo González, 2003, 2004). Además, en Argentina también se han publicado los léxicos disponibles de Resistencia (Wingeyer, 2007) y Córdoba (Darluzo y González, 2010).

Ha sido, no obstante, en España donde el Proyecto Panhispánico ha alcanzado más auge. Véase en el Cuadro 1.3 la nómina de investigaciones con diccionario y monografía publicada, ordenadas por comunidades autónomas.

**Cuadro 1.3. Estudios de disponibilidad léxica realizados dentro del Proyecto Panhispánico**

Investigador (Año)	Comunidad
1	2
Madrid	
Benítez Pérez (1992b, 1995)	Madrid
Paredes García (2001a)	Alcalá de Henares (Madrid)
Canarias	
Samper Padilla y Hernández Cabrera (1997); Samper Padilla (1998b); Samper Hernández (2003a, 2009)	Gran Canaria
Andalucía	
Mateo García (1998)	Almería
González Martínez (1997, 2002)	Cádiz
Cruzado Romero (2003)	Moguer (Huelva)
Bellón Fernández (2003, 2011)	Córdoba
Prado Aragonés y Galloso Camacho (2005, 2006)	Huelva
Ortolano Ríos (2005)	Ayamonte (Huelva)
Ahumada Lara (2006)	Jaén
Ávila Muñoz (2006); Ávila Muñoz y Villena Ponsoda (2010)	Málaga
Trigo Ibáñez (2007, 2011)	Sevilla
Pastor Milán y Sánchez García (2008)	Granada
Ayora Esteban (2003, 2006)	Ceuta
Fernández Smith <i>et al.</i> (2008)	Melilla
Castilla y León	
Galloso Camacho (1997, 1998a, 2002a)	Zamora
Galloso Camacho (2001, 2002b, 2003a)	Ávila, Salamanca y Zamora
Bartol Hernández (2004a)	Soria
Fernández Juncal (2008)	Burgos
Santiago Guervós (2008)	Segovia
Cruz Alonso (2012)	Castilla y León
Castilla-La Mancha	
Hernández Muñoz (2002a, 2004)	Cuenca
Hernández Muñoz (2006)	Castilla-La Mancha
Valencia	
Gómez Devís (2004); Gómez Molina y Gómez Devís (2004)	Valencia
Martínez Olmos (2007)	Alicante

1	2
Otras comunidades	
Etxebarria Arostegui (1996, 1999)	Bilbao
Carcedo González (2001)	Asturias
Serrano Zapata (2002)	Lleida (Cataluña)
Arnal Purroy (2004, 2008a)	Aragón
López Meirama (2008, 2011)	Galicia
Areta Lara (2009)	Navarra
Prado Aragonés <i>et al.</i> (2009)	Zonas fronterizas: Andalucía y Extremadura (España) - Algarve y Alentejo (Portugal)
Fernández Juncal (2013)	Cantabria

Entre estas investigaciones cabe destacar:

– Las investigaciones en la isla de Gran Canaria, dirigidas por J. A. Samper Padilla y C. E. Hernández Cabrera. Aportaron al Proyecto Panhispánico los criterios de edición (Samper Padilla, 1998a: 311–313), que ya habían sido aplicados exitosamente en los estudios previos de Gran Canaria (Samper Padilla y C. E. Hernández Cabrera, 1997).

– Andalucía es la comunidad de España donde más adelantados están los estudios de disponibilidad léxica. El proyecto de estudio de la disponibilidad léxica, dirigido por H. López Morales y F. García Marcos, ha culminado el objetivo de elaborar un diccionario de disponibilidad para cada provincia, siendo el siguiente paso la elaboración del léxico disponible de Andalucía. Desde las universidades andaluzas también se han realizado las investigaciones del léxico disponible en las ciudades autónomas de Ceuta (Ayora Esteban, 2003, 2006) y Melilla (Fernández Smith *et al.*, 2008), así como de las zonas fronterizas entre España y Portugal de Extremadura y El Algarve y Alentejo (Prado Aragonés *et al.*, 2009).

– Las extraordinaria actividad investigadora del equipo de la Universidad de Salamanca, liderado por J. A. Bartol Hernández y N. Hernández Muñoz. Este equipo, no sólo ha completado las investigaciones de su región, sino que se ha hecho cargo también del estudio de la región de Castilla-La Mancha (Hernández Muñoz, 2000a, 2004, 2006). Este equipo ha mostrado un interés especial por los aspectos relacionados con la psicolingüística (Galloso Camacho, 1997, 1998b; Hernández Muñoz, 2004, 2005, 2006). Este equipo, que también es responsable de la página electrónica *Dispolex*, ha asumido asimismo la elaboración del léxico disponible de España, lo que supone aunar los diccionarios de las investigaciones particulares que se han venido desarrollando en cada una de las ciudades del país (Samper Padilla y Samper Hernández, 2006: 16).



## II. BASES TEÓRICAS Y METODOLÓGICAS DE LA DISPONIBILIDAD LÉXICA

### 2.1. Unidades de análisis del corpus léxico

La léxico-estadística suele emplear el concepto de *palabra*, desde una perspectiva formal, como el material gráfico que se encuentra limitado por espacios en blanco. Esta unidad se puede identificar, aislar y contar fácilmente, al tiempo que favorece la explicación de las propiedades distribucionales (formales) de sus constituyentes (flexión nominal y verbal). Esa es, por ejemplo, la solución adoptada en el proyecto DIES-M, en el que Ávila (1995), para su tratamiento informático, define ‘palabra gráfica’ como “una secuencia de letras antes y después de la cual aparece un espacio en blanco”.

La segmentación y el análisis del corpus de la disponibilidad léxica, no obstante, han obligado a ampliar y adaptar las unidades de análisis del corpus léxico. El término *palabra*, considerado como “palabra gráfica”, es adecuado para estudios de densidad léxica de textos, pero no para catalogar y analizar las respuestas producidas por los informantes, que pueden constar de más de una palabra. En efecto, estas respuestas incluyen lexías simples (*mano, manzana, silla...*), lexías complejas (*columna vertebral, mesita de noche, cuarto de baño...*) e, incluso, sintagmas verbales (*regar las plantas...*).

Cuando se habla de palabras y vocablos, términos habituales en la estadística y la disponibilidad léxicas, en realidad se trata de *unidades léxicas*, las cuales, como acabamos de decir, pueden constar de más de una palabra (*plato llano, arrancar malas hierbas, cortar el césped...*). La *unidad léxica* es un concepto de suma importancia en la didáctica del léxico, ya que constituye la unidad de significado en el lexicón mental. Se trata de la unidad conceptual básica en el aprendizaje del vocabulario, al poder estar formada por más de una palabra. Por consiguiente, la unidad léxica es adecuada para la medición del vocabulario, ya que de lo que se trata en las pruebas de disponibilidad léxica es de recuperar

y medir las unidades del vocabulario almacenadas en el lexicón mental. Este vocabulario, según aclaran Nattinger y De Carrico (1992), no sólo se almacena como lexemas individuales, sino que también lo hace de forma redundante, como segmentos de frases o partes largas de enunciados.

Terminológicamente, en los estudios léxico-estadísticos se ha generalizado el empleo del término *palabra* referido a todas y cada una de las respuestas computables de los informantes, en tanto que se usa *vocablo* en referencia a cada una de las unidades que componen el vocabulario, y que resulta de la aplicación de criterios homogeneizadores o lematizadores (agrupación de todas las repeticiones de una palabra en una única entrada o vocablo). La oposición *palabra/vocablo* se basa en la distinción básica entre ‘palabra’ y ‘vocablo’, establecida por Müller (1973: 230) para los estudios de léxico-estadística. El vocablo se define como “la actualización de un lexema en el discurso” (o en un corpus léxico obtenido mediante estímulos indirectos, en este caso), en tanto que la idea de palabra hace referencia a “toda ocurrencia de un vocablo cualquiera”. En la estadística léxica (Justicia, 1995), el vocablo es cada una de las unidades que componen el vocabulario; es lo que los lingüistas llaman vocabulario. Cada vocablo, en cuanto que entrada de diccionario, engloba un conjunto de tipos léxicos o palabras, dados textualmente bien por flexión, en el caso de los verbos, bien por derivación, en el caso de los sustantivos o adjetivos. Por ejemplo, pertenecen al mismo vocablo *decir* los tipos “digo, decían, han dicho, diciendo, ser dicho, etc.”

Müller (1973: 231) relaciona la distinción “palabra/vocablo” con otra entre léxico, perteneciente a la lengua, y vocabulario, perteneciente al discurso. Para Müller, el vocabulario de un texto es sólo una muestra del léxico de la lengua. En palabras de Justicia (1995: 16), “el léxico hace referencia al sistema de palabras que componen una lengua”, mientras que “cuando esas unidades virtuales de la lengua se actualizan en el discurso del hablante reciben el nombre de vocablos y el conjunto de los mismos lo llamamos vocabulario”.

## 2.2. El procesamiento del corpus léxico

En la estadística léxica se utiliza el término ‘lema’ como sinónimo de vocablo. Los lemas son los elementos resultantes de un proceso que reduce a una forma léxica única todas las posibilidades paradigmáticas de un término. En el proceso de la lematización se reagrupan las formas heterogéneas de un mismo vocablo y se separan las formas homógrafas que responden a vocablos diferentes. Como resultado, se obtiene un inventario lematizado que contiene como unidad la unidad léxica, y las formas flexivas que la representan se consideran parte de la misma en los cómputos principales (Gómez Devís, 2004: 77).

El procesamiento del material recogido en las encuestas es, no obstante, el aspecto más controvertido de la investigación en disponibilidad léxica, como ponen

de manifiesto Gómez Devís (2004: 76) y Ayora Esteban (2006: 78). La complejidad de esta tarea ha dado pie a que reciba diferentes denominaciones por parte de los estudiosos: edición de los materiales –la formulación más generalizada– (Samper Padilla, 1998a; Carcedo González, 2001; Galloso Camacho, 2002b; Bartol Hernández, 2004a; etc.), edición y codificación de los materiales (López Morales 1999a), estandarización de los datos (Alba, 1995a; Galloso Camacho, 2002a, 2002b), pautas de unificación (Mateo García, 1998), validación (Valencia y Echeverría, 1999a), etc. Cada una de estas denominaciones refleja que la edición de los materiales depende fundamentalmente de los objetivos de la investigación: dialectales, interlingüísticos, cognitivistas, planificación de la adquisición léxica, etc.

A propósito de los cambios acaecidos en la preparación de los materiales a lo largo del desarrollo de los estudios de disponibilidad léxica, Hernández Muñoz (2006: 266) recuerda que estos dependen de los reenfoques sucesivos habidos en este tipo de investigaciones. A continuación, la autora menciona los siguientes enfoques:

– *Enfoque pedagógico*, utilizado por los autores franceses, cuyo fin es la elaboración de material didáctico, para el cual no se requiere la conservación de las variantes morfológicas, los extranjerismos, los dialectalismos, etc.

– *Enfoque descriptivo*, adoptado por los estudios hispánicos, cuyo fin es la descripción lingüística del léxico de una comunidad de habla y el cotejo de la variación dialectal y social, para el cual se requiere la conservación de todas las peculiaridades léxicas.

– *Enfoque psicolingüístico*, apoyado por investigadores como Hernández Muñoz (2006)<sup>1</sup> y Ávila Muñoz y Villena Ponsoda (2010), cuyo fin es el estudio de las implicaciones psicolingüísticas del léxico disponible, para poder pasar de las manifestaciones subjetivas a los comportamientos universales.

En cualquier caso, de todas estas opciones terminológicas, la fórmula que se ha generalizado para designar esta parte metodológica en los estudios de disponibilidad léxica es *Criterios de edición del corpus*. En palabras de Hernández Muñoz (2004: 29), este término compartido por la mayoría de los lingüistas “expresa el objetivo último de su existencia: preparar los datos ofrecidos por los informantes en las encuestas para ser procesados por los medios informáticos con el objetivo de editarlos en tablas para su análisis posterior”.

Este análisis posterior servirá para cumplir los objetivos principales del Proyecto Panhispánico: elaborar un diccionario de léxico disponible del español general, y posibilitar estudios comparativos entre los léxicos disponibles de diferentes áreas y países hispanohablantes. “De ahí la necesidad de lematizar con el fin de unificar variantes de una misma unidad léxica para proceder a los cotejos

---

<sup>1</sup> Hernández Muñoz (2006: 269) plantea hipotéticamente una perspectiva conceptual, que transforma los listados de acuerdo con las peculiaridades “más semánticas” y “menos léxicas” y que “defiende que la unidad conceptual prevalece sobre la morfológica”.

dialectales” (Gómez Devís, 2004: 78). Por la misma razón, cuando se unifican variantes de una misma unidad léxica, se deben aplicar criterios comunes al resto de los estudios de disponibilidad léxica.

En el marco del Proyecto Panhispánico, ha sido Samper Padilla (1998a)<sup>2</sup> quien ha fijado los criterios de edición para el procesamiento de los datos que los informantes actualizan en las encuestas. No faltan, empero, cuestiones discutibles, dudas y variaciones, con respecto a las sugerencias generales del Proyecto Panhispánico, en especial “los referidos al alcance de las asociaciones y a la necesidad de lematizar con el fin de unificar variantes de una misma unidad léxica”, que más adelante expondremos.

Hernández Muñoz (2004: 29–30) y Gómez Devís (2004: 78) señalan que el proceso de edición está regido por unos principios teóricos, explícitos o implícitos, que guían las decisiones personales del investigador, incluyendo las que representan divergencias<sup>3</sup> a los criterios comunes. Estos principios sobre los que se fundamenta la toma de decisiones, se presentan pares de conceptos en oposición:

a) *fidelidad al corpus/estandarización*.

La estricta fidelidad al corpus, defiende un respeto máximo a las respuestas de los informantes y responde a que cuanto menor sea la manipulación de los datos recogidos más reales serán los resultados. Este principio también asegura que en la edición del material se produce la menor pérdida de información posible. Bartol Hernández (2004: 1077) llama a este principio *mantenimiento de la información* recabada en las encuestas.

Frente a él se sitúa la *estandarización*, a la que Bartol Hernández (2004b: 1077) denomina “principio *uniformador*, imprescindible en estudios que se enmarcan en un proyecto mayor cuyo objetivo final es la publicación de un único diccionario de léxico disponible. De acuerdo con este objetivo son o serán de aplicación los criterios lexicográficos de lematización”. El criterio uniformador implica siempre algún tipo de pérdida de información (lingüística, cultural, etc.).

b) *lematización/conceptualización*.

El principio de “*lematización más conceptualización* evalúa qué tipo de unidades recogen, analizan y comparan los diccionarios de disponibilidad léxica: *conceptos* entendidos como unidades mentales o *palabras*, signos gráficos diferentes” (Hernández Muñoz, 2004: 30). A continuación esta autora aclara que los extremos de esta dualidad vendrían dados por una “*lematización rigurosa* que podría estar representada por la reducción de los sustantivos con moción al masculino singular y los verbos al infinitivo por simples criterios morfológicos”, en tanto que la *conceptualización* podría “admitir las palabras derivadas con sufijos

<sup>2</sup> Los criterios expuestos en ese artículo fueron discutidos y aprobados más tarde en la reunión de coordinación de Bilbao (marzo, 1999), organizada por la profesora Maitena Etxeberria.

<sup>3</sup> En opinión de Samper Padilla *et al.* (2003: 99), “son divergencias poco relevantes, que afectan a cuestiones de detalle”.

aspectuales que no aparecen recogidas en el diccionario de la Academia, pero que claramente son objetos distintos para los hablantes”.

Hay que tener en cuenta que las decisiones que se tomen, siguiendo uno u otro criterio pueden modificar, en algunos casos sustancialmente, el índice de disponibilidad de una palabra. Además, como advierte Ayora Esteban (2006: 78), “hay que tener cuidado y no perder información a causa de un proceso de limpieza excesiva de los datos”. Una excesiva estandarización de las producciones reales que proporcionan los informantes es un inconveniente para describir la variación léxica motivada por factores sociolingüísticos, psicolingüísticos, etc.

En opinión de Fernández Smith *et al.* (2008: 51), “parte del problema tiene que ver con el hecho de que se toma el diccionario normativo de la Academia como fuente para identificar lexicalizaciones”, cuando, por una parte, “«lexicalizado» no implica «de uso común» y, por otra parte, [...] no parece muy sensato recurrir al repertorio léxico de una norma estándar para llevar a cabo estudios sobre normas dialectales o, más aún, diastráticas, como es el caso del Proyecto Panhispánico”. La solución a estos problemas de edición es doble. En primer lugar, (Hernández Muñoz, 2004: 30) concluye que la referencia última para tomar decisiones en la edición de los textos no debe ser ni el investigador ni el diccionario, sino los propios informantes, lo que, de hecho, ya hacen en la práctica los investigadores, ya que siempre procuran recoger los términos propios de la comunidad de hablantes observada. En segundo lugar, algunos autores reclaman la adopción de criterios más laxos a la hora de unificar vocablos (Fernández Smith *et al.*, 2008: 51), y la necesidad de replantearse algunos criterios de edición y de concretar de forma precisa la lematización del corpus en aras del rigor y objetividad del Proyecto Panhispánico de léxico disponible como de la fiabilidad de los datos (Gómez Devís, 2004: 79).

Antes de pasar a exponer las normas generales de edición del Proyecto Panhispánico, nos detenemos a comentar el proceso de depuración de las listas, sin duda, el aspecto que más debate ha suscitado con posterioridad a la publicación de los “Criterios de edición del léxico disponible” (Samper Padilla, 1998a). La depuración de las listas consiste en establecer el grado de asociación necesario de las palabras que contienen las listas con el centro de interés para que sean aceptadas, el tratamiento que debemos dar a los extranjerismos, marcas comerciales o dialectalismos, etc.

Al respecto, los primeros trabajos del Proyecto Panhispánico (tomados enseguida como modelos) son parcos en explicaciones y taxativos a la hora de abordar la depuración de las listas. Así, López Morales (1999a: 40) explica que “por ejemplo, se conservaron algunas respuestas que no parecían corresponder fielmente al centro en cuestión (*sacerdocio, administración de empresas, ciencias políticas*, etc. para ‘Profesiones y oficios’, o *lámpara, abanico, cojín, tostadora*, etc. para ‘Muebles de la casa’”, en tanto que Carcedo González (2001: 36) afirma que “la

muy necesaria y comprensible labor inicial de depurado” se ocupa de “lexemas manifiestamente ajenos al área que se solicita, palabras desconocidas por los diccionarios al uso, etc.” Trabajos posteriores, como el de Ayora Esteban (2006: 79) en Ceuta, aplican idéntico criterio e indican expresamente que “hemos eliminado las palabras totalmente ajenas al centro de interés en cuestión”.

Lo cierto es que, como recuerdan Ávila Muñoz y Villena Ponsoda (2010: 95), “entre los planteamientos de partida del PPHDL no se encuentra una norma común sobre el carácter temático o asociativo de los listados”. Los ejemplos anteriores continuaban la tradición francesa (Gougenheim, Michéa, etc.) en la que las listas se diseñaban siguiendo un criterio temático y no asociativo. Sin embargo, continúan Ávila Muñoz y Villena Ponsoda, frente a esta opción, diversos autores (Benítez Pérez, 1992b; Mateo García<sup>4</sup>, 1998; Galloso Camacho, 2003a; Bartol Hernández, 2004a; Hernández Muñoz, 2006; Ávila Muñoz y Villena Ponsoda, 2010; Trigo Ibáñez, 2011) siguen una tendencia “consistente en mantener todas las respuestas dadas por los informantes, aunque no estén ligadas directamente al campo nocional, o incluso la de conservar las variantes formales de las palabras, nos acercan hacia un carácter más asociativo de la disponibilidad”.

Como vemos, la depuración de las listas plantea un dilema en la edición de los datos: considerar solo aquellas palabras que se relacionen semánticamente con el estímulo inicial (planteamiento estrictamente categorial de asociaciones directas) o bien dar cabida también a aquellas otras relacionadas asociativamente con el centro de interés (planteamiento asociativo con asociaciones directas e indirectas). Obviamente la elección de uno u otro procedimiento, o de una postura intermedia, tiene una repercusión directa nada desdeñable en los valores de la producción léxica, como Ávila Muñoz y Villena Ponsoda (2010: 49) han demostrado.

### 2.3. Criterios de edición del Proyecto Panhispánico

Sea como fuere, de manera estricta o con ligeras modificaciones, todos los trabajos inscritos en el Proyecto Panhispánico aplican las siguientes normas generales de edición, como ya se ha dicho, publicadas por Samper Padilla (1998a):

1. Eliminación de los términos repetidos.

Se eliminan los términos repetidos en un mismo campo léxico de la encuesta. Aquí también se incluye el caso en que aparezca la variante plena y la abreviada de una misma palabra (*bici/bicicleta*), su forma en singular y plural (*mano/ma-*

---

<sup>4</sup> Esta autora aclara que mantiene “prácticamente todas las palabras aparecidas en un determinado centro de interés, aunque aludan indirectamente al mismo”, por la información que ofrecen “acerca de cómo se procesa la realidad que está siendo sometida a examen verbal”, y por la posibilidad de “plantear futuras investigaciones interdisciplinarias a partir de estos materiales” (Mateo García, 1998: 86–87).

*nos*), o su forma no marcada y una variante flexiva (*pollo/polluelo*). Aquí se incluyen los términos homonímicos, como *banco* en el centro de interés ‘La ciudad’, por estar descontextualizados<sup>5</sup>.

## 2. Corrección ortográfica.

Como principio general, se adopta la norma ortográfica seguida por la RAE. Muchos de los errores ortográficos se producen en palabras que no se usan con frecuencia en el lenguaje escrito. Además, se debe tener presente que, al comienzo de la encuesta, los informantes reciben la indicación de que no deben preocuparse por este aspecto.

## 3. Unificación ortográfica.

Se adopta una sola entrada para aquellas palabras que admiten dos grafías (*mahonesa/mayonesa*; *hierba/yerba*, *quiosco/kiosco*). La variación es más frecuente con los extranjerismos (*volleyball/voleibol*; *football/fútbol*). Suele adoptarse la forma que el *DRAE* o el Diccionario Panhispánico de Dudas presenta como preferible (*béisbol*, *espagueti*, *fútbol*, *suéter*, *voleibol*). No obstante, autores como Carcedo González (2001), Arnal Purroy *et al.* (2004) y Hernández Muñoz (2006) optan por la forma elegida mayoritariamente por los informantes, la cual coincide frecuentemente con la solución recomendada por el *DRAE* (*mayonesa*, *biquini*, *póquer*, etc.), pero difiere en otros casos (*croissant*, *flash*, *kiwi*, *whisky*).

Por lo demás, la unificación no afecta a aquellos casos en que junto a la forma extranjera aparece la correspondiente voz española, sea ésta traducción de aquella o de diferente origen; así se conservan dos entradas para *voleibol* y *balonvolea*, y *fútbol* y *balompié*. El resto de extranjerismos no normalizados por el *DRAE* se transcriben según la ortografía de la lengua de procedencia (*foie-gras*, *petit-Suisse*).

## 4. Neutralización de las variantes meramente flexivas.

Se adopta la forma no marcada del paradigma: los verbos en infinitivo, y los sustantivos y adjetivos, siempre que sea posible, en masculino y singular. Así, la variación singular/plural se elimina a favor del singular (*medias* y *media* > *media*), salvo casos de ambigüedad (en “Juegos y distracciones”, por ejemplo, se mantiene *cartas*, y no *carta*). Son excepciones de esta norma:

- las denominaciones de juegos: *veo*, *veo*; *cartas*, *damas*, *chinos*;
- los diminutivos lexicalizados: *mesilla* o *mesita*, *cucharilla*, *zapatilla*;
- los sustantivos que funcionan como complemento de un verbo u otro sustantivo: *cuidar animales*, *estación de autobuses*;
- los casos en que la oposición singular/plural supone significados distintos: *pecho/pechos*<sup>6</sup>;

<sup>5</sup> López Morales (1999a: 39, n. 22) ya había advertido que, tratándose de “metodología asociativa, por lo tanto descontextualizada, los posibles problemas de homonimia que se presenten quedan sin solución; la aparición de éstos es, sin embargo, altamente escasa”.

<sup>6</sup> Sin embargo, Orlando Alba (1995a), López Morales (1999a), Carcedo González (2001), González Martínez (2002) y Bartol Hernández (2004a) optan por unificarlos en *pecho(s)*, atendiendo

- los *pluralia tantum* en su forma más corriente: *tijeras*;
- las palabras que se usan mayoritariamente en plural: *gafas, bragas*;
- las diferencias genéricas basadas en heterónimos o sufijos especiales; se mantienen ambas formas: *toro/vaca; caballo/yegua; gallo/gallina; actor/actriz; pimienta/pimienta*.

No obstante, a veces en las encuestas aparecen términos solo en su forma femenina cuando también es posible la forma masculina, como, por ejemplo en el campo *Profesiones y oficios* en aquéllas que van ligadas tradicionalmente al sexo femenino (*niñera, limpiadora, empleada de hogar, secretaria*). En ese caso, diversos investigadores (Samper Padilla y Hernández Cabrera, 1997; Mateo García, 1998; Ayora Esteban, 2006) han optado por seguir la norma de la comunidad y mantener la forma femenina. Gracias a esta decisión, se conserva “información sociolingüística relevante”, ya que “los encuestados asignan automáticamente la palabra en cuestión al entorno femenino” (Mateo García, 1998: 85).

Hay, incluso, autores como Bartol Hernández (2004b: 1078) y Hernández Muñoz (2006: 275) que, en el caso de la variación numérica y de género, respetan siempre las formas aportadas por los informantes. De este modo, si no hay variación, se consigna la forma existente en el corpus, aunque sea en plural (*células, gafas, correos*). Si hay variación se lematiza en singular y se indica entre paréntesis la terminación de plural: *arteria(s), calcetín(es), short(s)*; si hay variación genérica, Bartol Hernández la hace constar en el lema: *ternero/a, ciervo/a, médico/a, profesor/a*.

El paréntesis también es usado por Gómez Devís (2004) y Hernández Muñoz (2006) en las formas verbales que aparecen en más de una forma no personal: así *andando* y *andar* se unifican en *anda(ndo)*.

#### 5. Unificación de los derivados regulares.

Se recogen en una sola forma las variantes morfológicas que no implican alteración del significado léxico. Así se funden las formas de diminutivo o aumentativo con las formas no marcadas (*fresquito > fresco; cochazo > coche*).

Se mantienen aparte, sin embargo, los diminutivos lexicalizados, es decir, que presentan un matiz que no es simplemente la modificación minorativa de la base léxica (*mesa/mesita de noche; cocina/cocinilla*).

#### 6. Unificación de formas plenas y acortamientos.

Afecta a las palabras que en español son susceptibles de ser acortadas. Mediante paréntesis se marca la parte que en ocasiones no aparece: *bolígrafo – boli > bolí(grafo); frigorífico – frigo > frigo(rífico); televisión – tele > tele(visión)*.

#### 7. Unificación de entradas compuestas por varios elementos y sus variantes acortadas.

---

a que se trata de una única palabra con significados diferentes, como prueba su lematización en los diccionarios.

La parte que se da por consabida y que no está presente en algunas actualizaciones aparece en los listados entre paréntesis: *columna – columna vertebral > columna (vertebral)*; *goma – goma de borrar > goma (de borrar)*.

#### 8. Marcas comerciales.

Se mantienen sólo aquellas marcas que están plenamente lexicalizadas, es decir, cuando en una situación determinada puedan sustituir al nombre del producto aún en el caso de que lo designado no se corresponda con la marca concreta. Es lo que ocurre en España con *coca-cola* (que designa cualquier bebida refrescante carbonatada con sabor a cola), *cola-cacao* (cualquier cacao en polvo), *danone* (yogur), etc.

Sin embargo, pueden encontrarse verdaderas dificultades a la hora de decidir cuándo se ha lexicalizado totalmente una marca, porque hay situaciones intermedias en que un nombre comercial puede funcionar como un representante de un producto para sus hablantes (que desconocen formas alternativas de designación) y no para otros. Para superar esta dificultad, algunos autores (Benítez Pérez, 1992b; Mateo García, 1998: 86; Galloso Camacho, 2002b: 30; Gómez Devís, 2004: 80; Ávila Muñoz, 2005: 19; Hernández Muñoz, 2006: 275) incluyen en sus estudios todas las marcas comerciales que aparecen en sus encuestas, considerándolo la opción menos arriesgada y la más objetiva e imparcial.

Ante las discrepancias existentes entre los criterios generales y las modificaciones particulares de distintos autores, Hernández Muñoz (2006: 278) defiende la coexistencia de criterios de edición diferentes, en función del objetivo final de la investigación. Así, si se trata de contribuir a la elaboración de un diccionario de léxico disponible del español general, los criterios de edición son más restrictivos para alcanzar la estandarización de las variantes regionales. Por el contrario, si de lo que se trata es de realizar el estudio del léxico disponible de una comunidad de habla particular, se hace imprescindible conservar todas las variantes. Ambos objetivos son asumidos por el Proyecto Panhispánico.

## 2.4. El índice de disponibilidad léxica

Las grandes investigaciones de Francia (Gougenheim, 1967; Gougenheim *et al.*, 1956, 1964) y de Canadá (Mackey, 1971a, 1971b), y también las de menor alcance llevadas a cabo en Inglaterra (Dimitrejevic, 1969), en España (Azurmendi Ayerbe, 1983) y en Puerto Rico (López Morales, 1973, 1978, 1979; Román-Morales, 1985) basaban las tareas de recuento exclusivamente en la frecuencia. No obstante, en estas investigaciones se aprecia que “los vocablos recogidos manifiestan grados diferentes de disponibilidad, lo que significa que algunas palabras fluyen con más facilidad y rapidez que otras semánticamente relacionadas con ellas” (Mena Osorio, 1986: 13).

Fue Charles Müller (1973) [1968] quien definió los criterios para el tratamiento estadístico de la disponibilidad léxica. Este lingüista señaló que había que tener en cuenta, no sólo la frecuencia, sino también el orden de aparición de cada uno de los términos utilizados por los informantes en cada uno de los listados, dado que “varias palabras podían alcanzar la misma frecuencia y no tener el mismo índice de disponibilidad” (López Morales, 1996: 246). En consecuencia, desde 1983 la disponibilidad léxica maneja dos factores: frecuencia y grado. Cada palabra disponible tiene una frecuencia específica, que se analiza contando la cantidad de veces que la palabra es mencionada por los sujetos de la prueba. El grado de disponibilidad o espontaneidad indica el lugar que ocupa la palabra en la lista. Las primeras, con mayor grado, acuden primero a la memoria; después el grado va descendiendo.

Para ponderar el grado –lugar que la palabra ocupa en las listas– con la frecuencia, y obtener finalmente el índice real de disponibilidad de cada palabra, Lorán y López Morales (1983) idearon una fórmula que fue puesta en práctica en las investigaciones de Cañizal Arévalo (1987) en México, Mena Osorio (1986) en Concepción, y Echeverría *et al.* (1985, 1987) en San Juan. Sin embargo, estos últimos, que trabajaban con listas abiertas, advirtieron que la fórmula de Lorán y López Morales perdía su valor discriminador a partir del vigesimotercer lugar. A raíz de esto, G. Butrón (1987, 1991) mejoró bastante la fórmula de Lorán y López Morales, controlando bastante el desajuste, pero no eliminándolo.

Por ello, López Chávez y Strassburger Frías (1987) diseñaron una nueva fórmula que demostró su superioridad sobre la anterior en las investigaciones de México (Ruiz Basto, 1987; Cañizal Arévalo, 1987; Mesa Canales, 1989), de Madrid (Benítez Pérez, 1995), y de Puerto Rico (López Morales, 1999a), ya que producía una adecuación descriptiva sumamente plausible, tanto en el grupo como en el individuo (López Chávez y Strassburger, 1991). Desde entonces se usa la fórmula de López Chávez y Strassburger, que incluye como factores relevantes para hallar el índice de disponibilidad:

- a) la frecuencia absoluta de la palabra,
- b) la frecuencia absoluta de la palabra en cada posición,
- c) el número de participantes en la encuesta,
- d) el número de posiciones obtenidas en la encuesta en el centro de interés analizado,
- e) las posiciones en que fue dicha la palabra.

La fórmula de López Chávez y Strassburger (1991) implementa los criterios de frecuencia y orden, al combinar la máxima posición alcanzada en el centro de interés en la encuesta ( $n$ ), el número de la posición en cuestión ( $i$ ), y el índice de la palabra tratada ( $j$ ):

$$D(P_j) = \sum_{i=1}^n \left( \frac{f_{ij}}{I} \right) e^{-2.3 \left( \frac{i-1}{n-1} \right)}$$

El índice de disponibilidad cuantifica, de este modo, que “frente a una situación concreta, aquellas palabras que primero acuden a nuestra memoria son más disponibles en relación con dicha situación que aquellas que no hacen su aparición de inmediato” (López Morales, 1996: 248). El índice de disponibilidad léxica de cada vocablo es un valor entre 0 y 1, donde el valor 1 equivale a la máxima disponibilidad y el valor 0 a una disponibilidad léxica nula. Este valor “se interpreta como el grado de accesibilidad de un vocablo en el lexicón cuando el tema de la conversación lo requiere” (Paredes García, 2012: 85). De este modo, el índice de disponibilidad explica adecuadamente la relación que existe entre el conocimiento y la producción de léxico.

Las investigaciones del Proyecto Panhispánico utilizan la fórmula de J. López Chávez y C. Strassburger Frías. Esta se incluye en los algoritmos de los programas informáticos de tratamiento de léxico disponible: Lexidisp para el entorno Windows, desarrollado por Moreno Fernández *et al.* (1995), y Dispolex, accesible a través de Internet (Bartol Hernández y Hernández Muñoz, 2004).

## 2.5. El tamaño de la muestra

Como vimos anteriormente, las investigaciones realizadas fuera del Proyecto Panhispánico se caracterizan por una gran heterogeneidad metodológica. Estas diferencias entre estudios son también muy marcadas en lo referente al tamaño de las muestras, que iban de 50 individuos en los estudios con preescolares en Costa Rica (Murillo Rojas, 1993, 1994; Sánchez Corrales y Murillo Rojas, 1993, 2006), o 54 informantes en el estudio con alumnos preuniversitarios en México (Justo Hernández, 1986) hasta las voluminosas muestras usadas en el macroestudio de Azurmendi Ayerbe (1983) en el País Vasco con 1224 informantes, o la muestra de 2052 estudiantes de 4º curso de Enseñanza Media estudiada en Chile por Valencia y Echeverría (1999a). Entre estos extremos se halla una gran diversidad en el tamaño de las muestras empleadas por los estudiosos, con cantidades que van ascendiendo desde los 63 escolares que conforman la muestra de San Juan de Puerto Rico (López Morales, 1973) hasta los 500 encuestados por Vargas Sandoval (1991) y González Becker y Vargas Sandoval (1992) en la V región de Chile.

Amén de otras diferencias metodológicas ya comentadas, esta diversidad en el tamaño de las muestras suponía un impedimento para realizar una comparación fiable entre los resultados de las distintas investigaciones, dado que el tamaño de la muestra incide directamente en la cantidad total de unidades léxicas y de vocablos (Samper Padilla *et al.*, 2003: 69). Por ello, para facilitar y hacer más fiables las comparaciones de los datos entre las distintas zonas dialectales, en la reunión de Bilbao del año 1999 los equipos que estudian la disponibilidad léxica dentro del Proyecto Panhispánico acordaron que fueran 400 los encuestados por cada área de estudio. En opinión de (Samper Padilla *et al.*, 2003: 69), “la cifra acordada (400) parece adecuada al tamaño de las poblaciones que se estudian”.

Un repaso a las muestras estudiadas en las investigaciones ya concluidas dentro del Proyecto Panhispánico (ver Cuadro 2.1) deja a la vista que, a pesar de los acuerdos adoptados, aún se está lejos de conseguir la homogeneidad muestral en las investigaciones del PPD. Únicamente en el área andaluza –con la excepción de Huelva– se ha cumplido casi de manera escrupulosa con la directriz de contar con 400 informantes por comunidad investigada (el estudio de Málaga sobrepasa esta cifra en 58 individuos). Hay, además, otro grupo de trabajos que, de manera aproximativa, manejan muestras cercanas a los 400 sujetos. Tal es el caso de los estudios de Melilla, la comunidad valenciana (Valencia, Alicante), Aragón, Asturias, Cantabria, Navarra, Tenerife y la República Dominicana.

**Cuadro 2.1. Tamaño muestral en diferentes investigaciones dentro del Proyecto Panhispánico**

Comunidad	Muestra	Comunidad	Muestra
Puerto Rico	558	Castilla-La Mancha	680
Rep. Dominicana	347	Madrid	257
Gran Canaria	539	Ávila	80
Resistencia, Argentina	501	Burgos	175
Córdoba, Argentina	214	Salamanca	100
Tenerife	449	Segovia	100
Almería	400	Soria	100
Cádiz	400	Zamora	100
Córdoba	400	Asturias	470
Granada	400	Cantabria	333
Huelva	220	Aragón	417
Jaén	400	Alicante	460
Málaga	458	Valencia	465
Sevilla	400	Navarra	430
Ceuta	266	Bilbao	245
Melilla	315	Lérida	253
Cuenca	117	Galicia	800

Por el contrario, las cifras son muy variadas en el resto de las investigaciones. En la banda superior de la distribución, se hallan muestras por encima de los 500 encuestados en Resistencia (Argentina), Gran Canaria y Puerto Rico, 680 informantes en Castilla-La Mancha, y hasta 800 informantes en Galicia; en la banda inferior, se cuentan muestras en el intervalo 275 a 175 informantes en Ceuta, Madrid, Lérida, Bilbao, Huelva, Córdoba (Argentina) y Burgos. Los estudios de Castilla-León ha asignado a cada provincia 100 informantes, a excepción de Burgos con 175 informantes y Ávila con 80 encuestados.

Múltiples han sido los factores que han incidido en la toma de decisiones en cada investigación, a la hora de establecer la configuración final de la muestra. En los estudios realizados dentro del PPDL esta configuración ha dependido de los siguientes factores:

**a) La representatividad estadística de la muestra.** Dada la influencia del enfoque sociolingüístico en los estudios de disponibilidad léxica, con frecuencia los investigadores han hecho constar que la relación entre la muestra con respecto del universo estadístico del que procede satisface ampliamente la *ratio* de representatividad habitualmente aceptada en los estudios sociolingüísticos de corte variacionista. Esta *ratio* fue establecida en 0,025% por Labov (1972). Así, por ejemplo, Hernández Muñoz (2006: 247) a partir del padrón municipal de 2002 de la Comunidad de Castilla-La Mancha establece que la muestra con la que trabaja representa el 0,61% de la población adolescente, y el 0,038% de la población total de la comunidad, porcentajes ambos que superan el 0,025% establecido por W. Labov. A continuación, la autora apostilla que “este número de informantes también satisface ampliamente las pautas fijadas en la reunión sobre léxico disponible realizada en Bilbao en 1999, que fijaba un número de 400 estudiantes”, de lo que se deduce que Hernández Muñoz lo interpreta como un valor mínimo que puede ser sobrepasado, en este caso con creces. En cualquier caso, la *ratio* de representatividad de 0,025% de Labov (1972) se cumple en los demás trabajos aquí contemplados, incluido los que manejan una muestra de 400 informantes, como es el caso de Almería (Mateo García, 1998: 67), donde la muestra supone el 15,56% del conjunto de la población escolar que cursaba estudios de C.O.U. en Almería en esa fecha.

Aparte de la *ratio* establecida por Labov, algunos autores utilizan instrumentos estadísticos para asegurar que el tamaño de la muestra garantiza la representatividad. Este es el caso del estudio de Gómez Devís (2004: 66) sobre Valencia, en el que es el criterio de representatividad estadística el que delimita el valor final de la muestra. Según Gómez Devís, “el tamaño de la muestra, 465 informantes obtenido tras aplicar la fórmula de Garret (1980: 270), algo superior a la magnitud estadísticamente exigida [inicialmente de 417 informantes], nos garantiza un nivel de confianza del 95% en los resultados obtenidos, [...] garantizando un nivel de precisión entre el 92,0% y el 98% en la exactitud de los datos con respecto al total de la población acotada”. Más adelante, Gómez Devís (2004: 68) aclara que “de acuerdo con las pautas de la sociolingüística se ha utilizado la técnica del muestreo por cuotas con afijación proporcional, atendiendo a las variables: sexo, tipo de enseñanza, lengua habitual, nivel sociocultural y zona espacial de residencia”.

**b) La estructura diatópica de la comunidad estudiada.** En su estudio del léxico disponible de Asturias, Carcedo González (2001: 25) señala que “seguir, en lo metodológico, las pautas comunes al proyecto [panhispánico] [...] no impide que se realicen, en aspectos puntuales, las modificaciones necesarias para adaptarlo a las circunstancias concretas de la comunidad cuyo vocabulario se rastrea”.

Una de ellas es la forma en la que se selecciona a los informantes. De esta forma, el autor procura que la distribución geográfica de los 470 sujetos que integran la muestra de Asturias corra pareja al reparto proporcional de la población en las diferentes zonas de la región, así como que el reparto según el factor social sexo responda muy de cerca, en sus proporciones, a la distribución que en razón del sexo observa la población asturiana. Con una distribución análoga, trabaja Martínez Olmos (2007) al posestratificar el muestreo según las cuatro comunidades de habla alicantinas definidas por estudios sociolingüísticos anteriores.

**c) El tamaño de la población estudiada.** Ayora Esteban (2006: 64, n. 2) reduce el número de informantes en su estudio sobre Ceuta, de los 400 recomendados a los 266 finalmente analizados, “por considerarlo muy elevado en proporción a la extensión y al tamaño de la población que estudiamos, dado que es el área con menor número de habitantes, junto con Gibraltar y Melilla, de todas las que se estudian en el Proyecto Panhispánico”. La muestra ceutí suponía más del 45% de la población total, ya que la población a la que correspondía la muestra osciló en el periodo estudiado (1999–2001) entre 597 y 589 alumnos. El tamaño de la población estudiada también influye en las submuestras provinciales de la investigación de Hernández Muñoz (2006: 248) en Castilla-La Mancha. La distribución por provincias de los 680 sujetos de la muestra castellano-manchega intenta en la medida de lo posible respetar el reparto proporcional de los centros educativos según la población estudiada. Por su parte, en Almería la muestra queda dividida según la representatividad demográfica de cada una de las zonas consideradas en el conjunto provincial (Mateo García, 1998: 69).

**d) La división en grupos con un valor prefijado.** Metodológicamente, se ha extendido un reparto por cuotas prefijadas por cada centro educativo que compone la muestra. El valor más habitual es el de 20 alumnos por centro. Muchos son los autores que se inclinan por esta asignación. Así, por ejemplo, es utilizado en las investigaciones de González Martínez (2002) en Cádiz, de Pastor Milán y Sánchez García (2008) en Granada, o de Bellón Fernández (2011) en Córdoba, en las que la muestra de la provincia está compuesta por un total de 400 estudiantes de 2º curso de Bachillerato, procedentes de 20 centros educativos distintos, seleccionándose en cada uno de los mismos a un grupo de 20 informantes escogidos aleatoriamente. Esta selección de 20 alumnos por centro también estaba presente en muestras con tamaños diferentes, como el estudio de Asturias de Carcedo González (2001), el estudio de Castilla-La Mancha (Hernández Muñoz, 2004), y los estudios que el grupo de investigación de la Universidad de Salamanca realiza en Castilla-León. Otros trabajos, sin embargo, no aplican esta asignación de cuotas, como es el caso de Martínez Olmos (2007: 86), quien señala que su trabajo sobre Alicante cuenta con una media de 34 alumnos por grupo.

## III. LOS CENTROS DE INTERÉS

### 3.1. Concepto y selección

“Para la elaboración de los léxicos disponibles se emplean los datos obtenidos en pruebas asociativas controladas, que constituyen un modo artificial de conseguir que afloren a la superficie [...] las palabras de las que dispone, para uso inmediato, un determinado hablante o un grupo específico de ellos” (Samper Padilla *et al.*, 2003: 46–47). Se trata de encuestas de listas abiertas en las que durante dos minutos los informantes deben anotar todas las palabras que saben en torno a unos determinados centros de interés. Los centros de interés son unos estímulos temáticos relativamente amplios que persiguen provocar las respuestas de los sujetos investigados.

La noción de los “centros de interés” procede de la pedagogía, y fue recomendada por Michéa (1950: 189) como base de la enseñanza de las palabras concretas. Su origen se halla en *El método de los centros de interés*, también conocido como *método Decroly*, en honor a su fundador, el médico y psicopedagogo Ovide Decroly (1871–1932), representante del movimiento de la Escuela Nueva. El centro de interés, junto con la globalización y el juego, constituye el tercer pilar básico de esta metodología. Para Decroly, los centros de interés de los niños tienen cuatro ejes que atienden a las necesidades de los niños:

- alimentarse para conservar y desarrollar la vida: alimento, respiración...
- protegerse de la intemperie: calor, frío, humedad, viento...
- defenderse contra los peligros y enemigos: limpieza, enfermedad, accidentes...
- actuar y trabajar solidariamente, recrearse y mejorar: acción, alegría, vida en sociedad, diversión y descanso...

Desde muy pronto el niño relaciona estos centros con otros, como son los animales, las plantas, la naturaleza, las sociedades, las civilizaciones, las culturas, etc. Estos centros de interés se vinculan con aquello que motiva o estimula al niño. Como afirma Sánchez-Saus, la metodología de los centros de interés

pone en relación directa la enseñanza de idiomas y la disponibilidad léxica gracias a las *palabras-clave*:

Esta metodología de los centros de interés, en efecto, agrupa los distintos contenidos que quieren enseñarse en torno a diferentes temas, que podrían funcionar como *palabras-clave*, en los que se concentran los intereses del niño. Este método se aplica con mucha facilidad a la enseñanza de lenguas extranjeras, pues permite concentrar la enseñanza, sobre todo la del léxico, alrededor de temas, tal y como hace el método de la disponibilidad (Sánchez-Saus, 2011: 109).

En los estudios de disponibilidad léxica, la selección de los centros de interés ha ido variando en naturaleza y cantidad a lo largo de la historia, en función del marco teórico base del estudio, o de la necesidad de su adecuación al entorno concreto de las comunidades de habla, los objetivos específicos de las investigaciones, la edad de los informantes, etc. Un breve repaso a las investigaciones realizadas en el ámbito hispánico (*supra* § 1.2.1; y Samper Padilla *et al.*, 2003: 46–54) nos muestra la diversidad de modelos temáticos empleados, que por lo que se refiere al mundo hispánico, se inicia con los diez centros de interés que empleó López Morales (1973) en San Juan de Puerto Rico, y que también utilizaron Román-Morales (1985) en Dorado (EE.UU.), y Mena Osorio (1986) y Echeverría (1987, 1991) en Chile<sup>1</sup>:

01. ‘Alimentos’.
02. ‘Animales’.
03. ‘La casa’.
04. ‘La cocina’.
05. ‘El cuerpo humano’.
06. ‘Materiales de construcción’.
07. ‘Muebles’.
08. ‘Naturaleza’.
09. ‘La ropa’.
10. ‘Transportes’.

Hasta llegar a la configuración definitiva de los 16 centros de interés del Proyecto Panhispánico, el número y selección de las áreas temáticas sondeadas ha variado: seis centros emplea Cañizal Arévalo (1987) en México, otros seis son los investigados por Vargas Sandoval (1991) en Chile, 8 centros sondea Butrón (1987) en Puerto Rico, 10 centros forman la nómina de las investigaciones de López Morales (1973) en San Juan de Puerto Rico, Román-Morales (1985) en Dorado (EE.UU.), y Mena Osorio (1986) y Echeverría (1987, 1991) en Chile, también 10 es el número de centros investigado por Justo Hernández (1986) en México, a 15 asciende el número de centros de interés en la investigación de Azurmendi Ayerbe (1983) en San Sebastián, España, y hasta un total de 18 centros

---

<sup>1</sup> Los autores chilenos añaden ‘Herramientas’ al centro ‘Materiales de construcción’.

incorporan Valencia y Echeverría (1998, 1999a) en sus trabajos sobre disponibilidad léxica en Chile.

La frecuencia de aparición de los centros de interés más asiduos en los estudios para el español anteriores al Proyecto Panhispánico<sup>2</sup> queda como sigue:

- ‘Alimentos’ (7), ‘Alimentos y bebidas’ (4), ‘Comidas y bebidas’ (2).
- ‘Animales’ (9), ‘Plantas y animales’ (1), ‘Agricultura y vegetales. Ganadería y animales. Pesca y peces’ (1).
- ‘La casa’ (8), ‘Partes y objetos que hay en una vivienda’ (1).
- ‘La cocina’ (7).
- ‘El cuerpo humano’ (6), ‘Partes del cuerpo’ (7).
- ‘Materiales de construcción’ (2), ‘Herramientas y materiales de construcción’ (3).
- ‘Muebles’ (7), ‘Muebles de la casa’ (1), ‘Objetos que hay en el interior de una casa’ (2), ‘Partes y objetos que hay en una vivienda’ (1).
- ‘Naturaleza’ (6), ‘La naturaleza, las vacaciones’ (1).
- ‘La ropa’ (8), ‘Ropa y calzado’ (1), ‘Prendas de vestir y de calzar’ (1).
- ‘Transportes’ (6), ‘Medios de transporte’ (6).
- ‘Profesiones y oficios’ (5), ‘Trabajo’ (1).
- ‘Juegos y entretenimientos’ (2), ‘Juegos, diversiones y entretenimientos’ (1), ‘Juegos y distracciones’ (1).
- ‘Familia’ (2), ‘Relaciones familiares’ (1).

Como se aprecia, la nómina de centros es amplia y se utilizan diferentes criterios a la hora de parcelar los temas de la realidad. Véase una amplia revisión de las pruebas asociativas empleadas en los diferentes trabajos en López Morales (1999a: 32–37), Carcedo González (2001: 32–34) y Samper Padilla *et al.* (2003: 46–55). No obstante, por su especificidad y originalidad, destacan las propuestas de:

- Justo Hernández (1986) en México, quien analiza el léxico de los colores.
01. ‘Ropa’.
  02. ‘Zapatos’.
  03. ‘Cosméticos’.
  04. ‘Automóviles’.
  05. ‘Anuncios comerciales’.
  06. ‘Mochilas, maletas y bolsos’.
  07. ‘Muebles’.
  08. ‘Tapices y alfombras’.
  09. ‘Pintura de casa (exterior – interior)’.
  10. ‘Vestido’.

---

<sup>2</sup> Se consideran pertenecientes a una misma investigación los trabajos de Vargas Sandoval (1991), y González Becker y Vargas Sandoval (1992) sobre la V región de Chile; Murillo Rojas (1993, 1994) y Sánchez Corrales y Murillo Rojas (1993) para las investigaciones de principios de los años 90 sobre los niños preescolares en Costa Rica, considerándose el trabajo de ambas autoras en 2006–12 años después– una investigación nueva.

– Vargas Sandoval (1991) en Chile, con centros de contenido abstracto<sup>3</sup>.

01. ‘Arte’.
02. ‘Derecho’.
03. ‘Economía’.
04. ‘Política’.
05. ‘Sentimiento’.
06. ‘Teología’.

– Valencia y Echeverría (1998, 1999a), quienes, a partir del número 12, incluyen algunos centros poco usuales.

12. ‘La política’.
13. ‘La actividad económica’.
14. ‘Las artes’.
15. ‘Ciencia y tecnología’.
16. ‘Mundo espiritual’.
17. ‘Procesos mentales’.
18. ‘Problemas del ambiente’.

– Azurmendi Ayerbe (1983) en San Sebastián, España, por la diversidad de los centros de interés seleccionados:

01. ‘Partes del cuerpo’.
02. ‘Prendas de vestir’.
03. ‘Partes y objetos que hay en una vivienda (casa o piso)’.
04. ‘Alimentos y bebidas’.
05. ‘Relaciones familiares y no familiares’.
06. ‘Ciudad o pueblo: sus partes – lo que se ve y se hace en ellos’.
07. ‘Los transportes: medios y vehículos que se utilizan’.
08. ‘Enseñanza: clases de centros y sus partes; objetos y material utilizados, asignaturas, etc.’

09. ‘Trabajo: clases de trabajos, de oficios, de profesiones’.
10. ‘Diversiones. Entretenimientos’.
11. ‘Deportes: prendas de vestir y objetos que se utilizan’.
12. ‘Acciones que normalmente se utilizan todos los días’.
13. ‘La naturaleza – las vacaciones’.
14. ‘Agricultura y vegetales – ganadería y animales – pesca y peces’.
15. ‘Dinero – bancos – comercio’.

Sin embargo, a partir de los 90, todos las investigaciones realizadas dentro del Proyecto Panhispánico, tienen en común los 16 centros de interés tradicionales:

01. ‘Partes del cuerpo humano’.
02. ‘La ropa’.
03. ‘Partes de la casa (sin los muebles)’.

---

<sup>3</sup> Anteriormente en México, Cañizal Arévalo (1987) había trabajado con ‘Defectos físicos y morales’ y Ruiz Basto (1987) con ‘Cualidades y defectos físicos y morales’.

04. ‘Los muebles de la casa’.
05. ‘Alimentos y bebidas’.
06. ‘Objetos colocados en la mesa para la comida’.
07. ‘La cocina y sus utensilios’.
08. ‘La escuela: muebles y materiales’.
09. ‘Iluminación, calefacción y medios de airear un recinto’.
10. ‘La ciudad’.
11. ‘El campo’.
12. ‘Medios de transporte’.
13. ‘Trabajos del campo y del jardín’.
14. ‘Los animales’.
15. ‘Juegos y distracciones’.
16. ‘Profesiones y oficios’.

Para Gougenheim *et al.* (1967) estos campos constituían los sectores semánticos más representativos, ya que cumplían la condición de la universalidad, por cuanto que “se creían los más comúnmente compartidos por los hablantes, independientemente de su procedencia social, geográfica, profesión, etc. Y además se consideraban suficientemente diferenciados semánticamente” (Carcedo González, 2000a: 57).

Algunos autores, no obstante, han criticado esta selección. Así, Mackey (1971a) pone objeciones a la imperfecta diferenciación semántica de los centros, ya que algunos pueden compartir vocabulario. Galisson (1991: 163), por otra parte, los critica precisamente por ser demasiado universalistas y amplios. Gómez Devís (2004: 56) cuestiona la utilidad de algunos de estos centros (‘Medios de Iluminación, calefacción y ventilación’; ‘Tareas del campo y del jardín’, etc.) y propone introducir otros nuevos, acordes con la realidad y el desarrollo social (‘Colores’, ‘Naturaleza o cuestiones medioambientales’, ‘Organización social y política’, ‘La vida religiosa’, ‘Ciencia y tecnología’, ‘Sentimientos y emociones’, ‘Cualidades humanas’, etc.), en tanto que Izquierdo (2003: 424) propone incluir ‘La familia’, ‘La enfermedad’, ‘El tiempo atmosférico’, ‘Los sentimientos’, ‘El carácter’, etc.

López Morales (1999a: 33), sin embargo, aun reconociendo los problemas teóricos que presenta clasificar exhaustivamente en centros de interés los intereses humanos, defiende el mantenimiento en el Proyecto Panhispánico de los mismos centros de interés originarios de las investigaciones francesas. La posibilidad de establecer comparaciones entre los resultados hispánicos y los obtenidos entre los francohablantes, y la falta de una solución teórica definitiva aconsejan mantener la selección de centros original (López Morales, 1999a).

En cualquier caso, las características propias de la comunidad de habla estudiada, así como los objetivos que las investigaciones han perseguido, han hecho que distintos investigadores dentro del Proyecto Panhispánico hayan añadido otros centros de interés a la lista de 16 centros clásicos. Así, Benítez Pérez (1992b), Samper Padilla y Hernández Cabrera (1995b), Etxebarria Arostegui (1996), Mateo

García (1998), Ayora Esteban (2003), Arnal Purroy *et al.* (2004), Gómez Devís (2004), Prado Aragonés y Galloso Camacho (2005), Ahumada Lara (2006), Casanova Ávalos (2006), entre otros, añaden el centro ‘Los colores’. Mateo García (1998) y Ayora Esteban (2003) incluyen además el centro ‘La mar’. Hernández Muñoz (2006) introduce ‘Inteligencia’, un campo abstracto, difuso en sus límites, y que evoca palabras sin referente en la realidad. Ahumada Lara (2006) introduce en Jaén el centro ‘El olivo y el aceite’, un campo muy relacionado cultural y económicamente con esta comarca.

### 3.2. Clasificación de los centros de interés

Según Hernández Muñoz (2006: 134), los centros de interés clásicos “son categorías semánticas dispares: ni su ontogenia, ni su organización interna, ni su alcance son los mismos”. Basa esta afirmación en que las características de los centros de interés (número de unidades que los componen, distribución de las mismas, relaciones entre ellas, número de subgrupos, concomitancia con otras categorías, límites más o menos difusos, correspondencia con categorías naturales, etc.) son distintas en cada caso.

Esta disparidad semántica se muestra ya en la misma relación que vincula las unidades léxicas con el estímulo generador. Como acertadamente ha notado Borrego Nieto:

Existen relaciones heterogéneas entre el enunciado del campo y el léxico resultante: en *Partes del cuerpo* o en *Animales* se obtienen listas de partes del cuerpo o de animales, es decir, los enunciados son hiperónimos de las palabras que constituyen las respuestas. En cambio en *El campo* y *La ciudad*, se obtienen los vocablos correspondientes a las realidades que evocan, y no hipónimos de los enunciados (Borrego Nieto, 2004: 67).

Es decir, entre los centros de interés de la prueba de disponibilidad léxica se definen centros abiertos –como ‘El campo’ o ‘La ciudad’–, y centros cerrados –como ‘Partes del cuerpo’ o ‘Los animales’. Estos últimos, los centros más concretos, como ‘Partes del cuerpo’, “están constituidos por un número menor de clases lógicas, lo que reduce el tipo de evocaciones que puede despertar en el hablante” (Carcedo González, 2003: 209).

Las diferencias semánticas existentes entre los centros de interés de la disponibilidad léxica han dado pie a Hernández Muñoz (2006: 130–156) a realizar una taxonomía de los mismos, usando los criterios de inclusividad y estructura interna (ver Cuadro 3.1). De acuerdo al primer criterio, el nivel de inclusividad, distingue:

- centros *inclusivos*, en los que “los elementos que los componen se relacionan con la etiqueta categorial con una relación de pertenencia estricta”

(2006: 150), según la proposición lógica del tipo “X es Y”: ‘Partes del cuerpo’, ‘La ropa’, ‘Partes de la casa’, ‘Los muebles’, ‘Comidas y bebidas’, ‘Objetos colocados sobre la mesa para la comida’, ‘Medios de transporte’, ‘Trabajos del campo y del jardín’, ‘Los animales’, ‘Juegos y distracciones’ y ‘Profesiones y oficios’;

- centros *relacionales*, en los que los elementos no se relacionan tan fácilmente con la etiqueta categorial y la proposición que los relaciona es “X es parte de Y” o “X tiene algún tipo de relación con Y”: ‘La escuela’, ‘Calefacción e iluminación’, ‘La ciudad’, ‘El campo’. En cuanto a ‘La cocina y sus utensilios’, se compone de dos categorías distintas, una inclusiva –‘sus utensilios’– y otra relacional –‘la cocina’.

En cuanto a la organización de sus elementos, los centros pueden ser:

a) *categorías naturales*, que responden a las características clásicas de una categoría semántica: límites más o menos definidos relacionados con características físicas o abstractas que nacen de la experiencia del mundo, y con identificación sencilla de los elementos que la constituyen (‘La ropa’, ‘Partes de la casa’, ‘Los muebles de la casa’, ‘Comidas y bebidas’, ‘La cocina y sus utensilios’, ‘Calefacción e iluminación’, ‘Medios de transporte’, ‘Trabajos del campo y del jardín’, ‘Animales’ y ‘Profesiones y oficios’);

b) *categorías bien definidas*, que poseen límites concretos y tienen una estructura de lista cerrada, y en las que el grado de pertenencia es igual para todos los miembros (‘Partes del cuerpo humano’);

c) *categorías ad hoc*, creadas con un objetivo concreto, y que no pueden explicarse ni mediante prototipos ni por semejanza de familia, puesto que su organización depende en exclusiva del conocimiento del mundo que tenga el hablante (‘Objetos sobre la mesa para la comida’);

d) *categorías radiales*, que consisten “en una subcategoría central con extensiones o ramificaciones sustentadas en diversos procedimientos entre los que destacan la metonimia o la metáfora”, y en las “que una palabra activa múltiples dominios cognitivos (redes conceptuales) que a su vez activan otros conceptos que no necesitan estar relacionados con los primeros” (2006: 147). Este es el caso de ‘La cocina y sus utensilios’ –con características también naturales–, ‘La escuela’, ‘La ciudad’, ‘El campo’ y ‘Juegos y distracciones’.

A estas categorías Sánchez-Saus (2011: 199) añade, siguiendo a Tomé Cornejo (2011), el modelo de organización de los *esquemas*. Los *esquemas –marcos (frames)* según Minsky (1975), o *esquemas visuales* según de Vega (1984)– proceden de la ciencia cognitiva y hacen referencia a “estructuras de conocimiento que contienen información estructural fija a partir del conocimiento genérico [...] acerca de cómo son los objetos de la realidad (del tipo *cómo es un aula de un colegio, cómo es una biblioteca* y, por tanto, qué esperamos encontrar cuando estamos en ellas)” (Sánchez-Saus, 2011: 205).

**Cuadro 3.1. Clasificación de los centros de interés de la disponibilidad léxica**

Centro Interés	Estructura interna					Niveles de inclusividad	
	natural	bien definida	ad hoc	radial	esquema	inclusiva	relacional
01. El Cuerpo Humano	x	x				x	
02. La ropa	x					x	
03. Partes de la casa	x					x	
04. Los muebles de la casa	x					x	
05. Comidas y bebidas	x					x	
06. Objetos en mesa para comida			x			x	
07. La cocina y sus utensilios	x			x	x	x	x
08. La escuela				x	x		x
09. Calefacción, iluminación	x						x
10. La ciudad				x	x		x
11. El campo				x	x		x
12. Medios de transporte	x					x	
13. Trabajos del campo/jardín	x					x	
14. Los animales	x					x	
15. Juegos y distracciones	x			x	x	x	
16. Profesiones y oficios	x					x	

La diferente estructura interna de los centros de interés afecta a la cohesión de los conjuntos léxicos. En disponibilidad léxica, ésta es medida por el índice de cohesión, concepto introducido por Max Echeverría *et al.* (1987: 68), que “mide el grado de coincidencia en las respuestas de los sujetos para un mismo centro”. El índice de cohesión relaciona los valores obtenidos en palabras y en vocablos, dividiendo el promedio de respuestas en cada centro de interés por el número de palabras diferentes (vocablos). Una mayor variedad en las respuestas implica un menor índice de cohesión y, por tanto, un campo abierto o difuso; por el contrario, un mayor número de coincidencias resulta en un mayor índice de cohesión, típico de un campo compacto o cerrado. Este último caso indica una relación estrecha entre los vocablos y su lugar de aparición en las listas.

Como comprobaremos en el próximo capítulo, la aplicación de este índice demuestra hay campos esencialmente compactos, como ‘El cuerpo humano’, ‘La ropa’ o ‘Los animales’, con un alto índice de vocabulario común, frente a campos difusos, como ‘La ciudad’, ‘El campo’, ‘Juegos y distracciones’ y ‘Trabajos del campo y del jardín’, estos últimos caracterizados por la dispersión de los vocablos que se recogen en las encuestas.

### 3.3. Regularidades en los centros de interés

En dos estudios diferentes, Samper Padilla *et al.* (2003), y Samper Padilla y Samper Hernández (2006) han demostrado que los centros de interés del Proyecto Panhispánico no ofrecen el mismo grado de productividad. Es más, por centros de interés se presentan ciertas regularidades en los datos cuantitativos que se repiten de unos estudios a otros, independientemente de la comunidad de habla de la que se trate. Estas regularidades se muestran tanto en el número total de unidades léxicas y el promedio de respuestas por sujeto, como en el número de vocablos, amén de otros índices cuantitativos como la densidad léxica y el índice de cohesión.

No obstante, la disparidad en el número de sujetos encuestados representa una dificultad añadida a la hora de efectuar comparaciones fiables entre los resultados de diferentes estudios, ya que esta incide directamente en la cantidad total de unidades léxicas y de vocablos en las investigaciones de disponibilidad léxica. Para solucionar este inconveniente, se suele utilizar el promedio de respuestas por sujeto en los estudios comparativos con otras comunidades de habla hispanas. Sin embargo, los datos cuantitativos relativos al número de vocablos están condicionados por el número de informantes que constituyen las muestras de las distintas comunidades, por lo que una comparación fiable en este índice sólo puede hacerse a partir de muestras semejantes.

Por ello, para minimizar la influencia del tamaño de las muestras en la comparación, limitaremos la base comparativa a aquellas investigaciones que manejan una muestra de 400 informantes o cuyo valor se halle en un intervalo de  $\pm 100$  unidades con respecto al valor recomendado en el Proyecto Panhispánico. Excluimos de la comparación la investigación de O. Alba (1995a) sobre la República Dominicana, concentrándonos en las investigaciones realizadas en el ámbito del estado español. Las investigaciones que cumplen con esta condición son<sup>4</sup>: Tenerife (Torres González y García Falero, 2009), Almería (Mateo García, 1998), Cádiz (González Martínez, 2002), Córdoba (Bellón Fernández, 2003), Granada (Pastor Milán y Sánchez García, 2008), Jaén (Ahumada Lara, 2006), Málaga (Ávila Muñoz, 2006), Sevilla (Trigo Ibáñez, 2011), Melilla (Fernández Smith, 2008), Asturias (Carcedo González, 2001), Cantabria (Fernández Juncal, 2013), Aragón (Arnal Purroy *et al.*, 2004), Alicante (Martínez Olmos, 2007), Valencia (Gómez Devís, 2004).

---

<sup>4</sup> Aunque la investigación de Navarra (Areta Lara, 2009) se hizo sobre una muestra de 430 informantes, no se incluye aquí debido a que no tuvimos acceso a esta en el momento de hacer la comparación.

### 3.3.1. Promedio de respuestas

Este valor, que mide la riqueza léxica de un centro de interés o de los informantes en particular, es uno de los datos más fiables en un análisis cuantitativo de disponibilidad, ya que no se ve afectado ni por el tamaño de la muestra ni por la presencia de informantes atípicos. Se calcula dividiendo el número total de unidades léxicas de un centro de interés (repetidas o no) entre el número total de informantes.

Un estudio comparativo de los promedios de respuestas por sujeto en las sintopías antes mencionadas, ayuda a catalogar la productividad de los ámbitos nocionales. En el Cuadro 3.2 se han calculado los valores medios para cada centro de interés y para cada sintopía. Adicionalmente aparecen sombreados en gris oscuro los valores máximos para cada centro de interés, y en un tono más claro los valores mínimos. Además se subrayan los valores coincidentes en centros de interés diferentes dentro de una misma sintopía.

**Cuadro 3.2. Promedios de palabras por centro de interés en investigaciones del Proyecto Panhispánico con muestras en torno a 400 informantes**

C. I.	Ten.	Alm.	Cád.	Cór.	Gra.	Jaén	Mál.	Sev.	Mel.	Ast.	Cant.	Arag.	Ali.	Val.	Me- dia
01. CUE	27,3	25,3	24,4	25,5	22,8	24,0	22,4	25,9	24,6	24,9	22,4	26,4	24,5	27,0	24,8
02. ROP	23,9	20,4	21,0	22,1	19,3	21,4	19,3	22,0	19,8	21,8	21,1	23,6	20,7	23,6	<u>21,4</u>
03. CAS	15,8	18,3	14,8	16,4	16,3	<u>17,5</u>	15,0	16,4	15,7	14,8	14,2	17,0	<u>15,7</u>	15,7	16,0
04. MUE	17,0	16,3	13,2	15,1	14,1	15,4	13,8	15,6	15,6	14,2	14,3	15,8	15,8	14,8	15,1
05. ALI	26,8	28,8	24,8	26,5	24,4	25,0	23,1	25,6	23,3	25,3	25,5	28,3	24,6	27,8	25,7
06. MES	16,6	19,8	14,1	15,0	13,7	15,3	13,3	14,9	13,6	15,0	15,0	17,7	<u>15,7</u>	16,2	15,4
07. COC	17,6	20,0	18,1	18,8	19,43	19,0	17,3	18,8	16,6	18,9	17,9	20,5	17,1	20,5	18,6
08. ESC	24,4	20,1	22,4	22,8	19,5	19,6	20,4	23,0	21,2	21,3	23,0	24,1	22,0	22,5	21,9
09. ILU	11,9	9,3	10,1	11,8	11,7	11,3	11,1	11,5	10,8	12,4	12,1	14,0	12,9	12,5	11,7
10. CIU	23,6	21,1	21,6	23,8	21,2	22,2	22,1	21,7	20,4	21,6	21,0	23,3	21,4	23,8	22,1
11. CAM	18,2	22,5	20,0	21,5	18,8	19,8	18,2	20,0	16,3	19,0	19,2	22,2	20,2	20,8	19,8
12. TRA	18,5	23,0	16,3	17,9	15,7	16,9	15,2	16,8	14,4	16,1	17,1	19,7	17,7	18,6	17,4
13. TRC	15,4	16,6	10,4	11,2	10,0	11,2	10,5	10,7	8,8	10,6	10,4	13,2	11,6	11,7	11,6
14. ANI	29,1	27,6	26,7	28,7	25,1	26,6	24,7	27,8	26,2	28,5	26,5	29,9	26,1	28,7	27,3
15. JUE	19,2	18,0	17,7	19,2	14,5	<u>17,5</u>	16,1	17,8	15,4	18,3	17,2	20,3	18,9	19,1	17,8
16. PRO	22,0	20,7	21,4	22,8	20,6	21,3	19,3	22,2	18,7	22,0	20,6	23,4	22,3	21,9	<u>21,4</u>
Media	20,5	20,5	18,6	19,9	17,9	19,0	17,6	19,4	17,6	19,0	18,6	21,2	19,2	20,3	19,2

Una primera ojeada a los datos ilustra sobre la reincidencia en los valores máximos en Almería (6), Aragón (6) y Tenerife (4), frente a una repetición de valores mínimos en Melilla (7), Málaga (4), Cantabria (2) y Granada (2). En prin-

cipio, habría que descartar la influencia del diferente tamaño de las muestras al tratarse del parámetro “palabras por informante”, aunque aparecen máximos en Tenerife, Aragón y Valencia, con muestras superiores a 400 informantes, y mínimos en Melilla y Cantabria, con 315 y 333 informantes, respectivamente.

Por ello, volvemos nuestra atención a la posible influencia de la composición sociológica de las muestras. De este modo, en el caso de Almería, los máximos se pueden achacar a que el 30% de la muestra (120 alumnos) pertenece a colegios privados, cuyas medias superan en más de 20 palabras como mínimo a las producciones de los colegios públicos –el 70% restante de la muestra– (Mateo García, 1998: 226–227). La diferencia es, en cualquier caso, muy apreciable, y su impacto se ha de hacer notar en los promedios de cada centro de interés. A ello hay que sumar la polarización sociocultural del tipo alto/no alto, que se da en esta provincia andaluza (Mateo García, 1998: 222), en los valores medios de palabras por informante en los niveles bajo (19,3), medio (20,5) y alto (25,4), respectivamente. Por lo que se refiere a Tenerife, hay que destacar que el porcentaje de informantes pertenecientes a la clase alta (16,3%) –la que suele mostrar una mayor productividad– es de los más elevados entre las distintas investigaciones; en Aragón es del 12,9%. Algunos de estos máximos se deben, además, con toda probabilidad a la confluencia de varios factores en la selección de la muestra. Ello se aprecia en Aragón (Arnal Purroy *et al.*, 2004), donde el grupo femenino es muy numeroso (285 mujeres, el 68,3% de la muestra) y obtiene medias superiores en once campos léxicos, entre ellos ‘La cocina’, al igual que en Valencia.

En cuanto a los valores mínimos, marcados en tono claro se nota un mayor tamaño muestral de los estratos sociales más bajos. Destaca el caso de Melilla, con una muestra conjunta de nivel bajo y medio-bajo de 78,09% frente a 21,91% de informantes de nivel medio-alto y alto. Esta asimetría muestral a favor de los niveles socioculturales bajos seguramente minimiza el impacto de los mejores promedios de los informantes de nivel medio-alto y alto en los resultados globales de Melilla (Fernández Smith *et al.*, 2008: 78). En Málaga (4 mínimos) y Granada (2 mínimos), también hay mayor porcentaje muestral de los estratos sociales más bajos. Así, en Málaga un 61,3% de informantes es de nivel bajo o medio-bajo, con diferencias significativas que muestran un condicionamiento de la productividad léxica por la clase social (Ávila Muñoz, 2006); por su parte, en Granada, donde se maneja el ‘nivel de instrucción de los padres’ como variable social, tan solo 86 informantes (21,5%) tienen padres con estudios superiores concluidos. Los porcentajes granadinos para los distintos grupos de la variable ‘nivel de instrucción de los padres’ son: Sin estudios: 21 informantes, 15 palabras; Estudios primarios: 166 informantes, 17 palabras; Estudios secundarios: 127 informantes, 18 palabras; Estudios superiores: 86 informantes, 19 palabras (Pastor Milán y Sánchez García, 2008: 38).

El Cuadro 3.3 ofrece los campos léxicos ordenados según el rango que ocupan en cada comunidad. Se ha sombreado los rangos que coinciden con el rango de la media general del Cuadro 3.2. Además, se añade el rango medio de cada centro de interés.

**Cuadro 3.3. Centros de interés, ordenados por rangos, según el promedio de palabras en investigaciones del Proyecto Panhispánico con muestras en torno a 400 informantes**

C. I.	Ten.	Alm.	Cád.	Cór.	Gra.	Jaén	Mál.	Sev	Mel.	Ast.	Cant.	Arag.	Ali.	Val.	Rango Medio	Me-dia
14. ANI	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1,07	1
05. ALI	3	1	2	2	2	2	2	3	3	2	2	2	2	2	2,14	2
01. CUE	2	3	3	3	3	3	3	2	2	3	4	3	3	3	2,86	3
10. CIU	6	6	5	4	4	4	4	7	5	6	6	7	6	4	5,29	4
08. ESC	4	9	4	5	6	8	5	4	4	7	3	4	5	6	5,29	5
16. PRO	7	7	6	5	5	6	7	5	7	4	7	6	4	7	5,93	6
02. ROP	5	8	7	7	8	5	6	6	6	5	5	5	7	5	6,07	6
11. CAM	10	5	8	8	9	7	8	8	9	8	8	8	8	8	8	8
07. COC	11	10	9	10	7	9	9	9	8	9	9	9	11	9	9,21	9
15. JUE	8	13	10	9	12	10	10	10	12	10	10	10	9	10	10,21	10
12. TRA	9	4	11	11	11	12	11	11	13	11	11	11	10	11	10,5	11
03. CAS	14	12	12	12	10	10	12	12	10	13	14	13	13	13	12,14	12
06. MES	13	11	13	14	14	14	14	14	14	12	12	12	13	12	13	13
04. MUE	12	15	14	13	13	13	13	13	11	14	13	14	12	14	13,14	14
09. ILU	16	16	16	15	15	15	15	15	15	15	15	15	15	15	15,21	15
13. TRC	15	14	15	16	16	16	16	16	16	16	16	16	16	16	15,71	16

El comportamiento de los rangos es muy similar en las distintas comunidades comparadas, y especialmente constante en los extremos de la distribución: valores superiores (1. ‘Animales’, 2. ‘Alimentos y bebidas’, 3. ‘El cuerpo humano’) y valores inferiores (15. ‘Iluminación, calefacción y medios de airear un recinto’, 16. ‘Trabajos del campo y del jardín’). En la ordenación, también tienen una distribución bastante constante los rangos medios (8. ‘El campo’, 9. ‘La cocina’, 10. ‘Juegos y distracciones’, 11. ‘Medios de transporte’). Coinciden el rango medio de ‘La ciudad’ y ‘La escuela’ (5,29), y el rango del promedio de palabras por informante en los centros ‘Profesiones y oficios’ y ‘La ropa’ (6).

En el resto de rangos, si bien aparecen, tanto en las diferentes comunidades como en el rango medio, discordancias con respecto al rango del valor medio, estas diferencias por lo general no son significativas (a lo sumo de 1, 2 ó 3 puntos). Tan sólo en Almería se muestran diferencias mayores: rango 4 en ‘Medios de transporte’ para un rango en la media de 11, rango 9 en ‘La escuela’ para un rango 5 en la media.

Los datos anteriores se resumen en el Cuadro 3.4, ordenando el promedio general de respuestas de las sintopías comparadas por intervalos de productividad. El valor de la desviación estándar informa sobre la dispersión de la distribución.

**Cuadro 3.4. Centros de interés, según la media general de respuestas por informante en investigaciones del Proyecto Panhispánico con muestras en torno a 400 informantes**

Rango	Rango Medio	Centros de interés	Promedio	Intervalo	Desviación
1	1,07	14. Los animales	27,3	24,0–28,0	1,555
2	2,14	05. Alimentos y bebidas	25,7		1,744
3	2,86	01. El cuerpo humano	24,8		1,570
4	<u>5,29</u>	10. La ciudad	22,1	19,0–24,0	1,127
5	<u>5,29</u>	08. La escuela	21,9		1,577
6	5,93	16. Profesiones y oficios	<u>21,4</u>		1,296
6	6,07	02. La ropa	<u>21,4</u>		1,520
8	8,00	11. El campo	19,8		1,676
MEDIA	–	–	19,2		1,134
9	9,21	07. La cocina	18,6	15,0–19,0	1,236
10	10,21	15. Juegos y distracciones	17,8		1,601
11	10,50	12. Medios de transporte	17,4		2,148
12	12,14	03. Partes de la casa	16,0		1,122
13	13,00	06. Objetos sobre la mesa para...	15,4		1,749
14	13,14	04. Los muebles de la casa	15,1		1,054
15	15,21	09. Iluminación, calefacción	11,7	11,0–15,0	1,168
16	15,71	13. Trabajos del campo y del jardín	11,6		2,124

Con esta información, es posible delimitar los siguientes cuatro grupos en cuanto al promedio de respuestas por individuo:

– los centros más productivos, entre 24 y 28 respuestas, son ‘Los animales’ (27,3), ‘Alimentos y bebidas’ (25,7) y ‘El cuerpo humano’ (24,8);

– en un segundo escalón, con una producción menor, aunque siempre superior a la media, hay un grupo de centros con valores muy próximos en el entorno de 20–22 respuestas: ‘La ciudad’ (22,1), ‘La escuela’ (21,9), ‘Profesiones y oficios’ (21,4), ‘La ropa’ (21,4) y, algo más distanciado, ‘El campo’ (19,8);

– hay un tercer grupo numeroso, con seis centros, en el intervalo 15–19 respuestas, con valores ya por debajo de la media: ‘La cocina’ (18,6), ‘Juegos y distracciones’ (17,8), ‘Medios de transporte’ (17,4), ‘Partes de la casa’ (16,0), ‘Objetos sobre la mesa’ (15,4), y ‘Los muebles de la casa’ (15,1);

– la menor evocación queda asociada con las áreas temáticas ‘Iluminación, calefacción y medios de airear un recinto’ (11,7) y ‘Trabajos del campo y del jardín’ (11,6), con valores casi idénticos. Como indica Samper Padilla *et al.* (2003: 59), ello “implica que estamos ante áreas léxicas más cerradas y poco familiares para los informantes de este tipo de estudios”.

Esta clasificación coincide, *grosso modo*, con la de Samper Padilla *et al.* (2003: 58) y, en especial, con Samper Padilla y Samper Hernández (2006: 30–31). Las únicas

**Cuadro 3.5. Promedios de palabras por centro de interés en las investigaciones del Proyecto Panhispánico**

	01. CUE	02. ROP	03. CAS	04. MUE	05. ALI	06. MES	07. COC	08. ESC	09. ILU	10. CIU	11. CAM	12. TRA	13. TRC	14. ANI	15. JUE	16. PRO	Media
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
P.R.	18,6	12,7	9,9	11,0	17,5	8,7	12,3	10,6	4,9	13,1	9,1	10,9	6,1	18,6	12,7	15,6	12,0
R.D.	23,0	17,9	15,7	14,0	23,9	11,4	13,9	15,7	8,4	16,5	14,6	14,5	7,5	24,4	14,7	17,6	15,9
G.Ca.	27,0	22,9	16,6	17,6	28,7	16,0	20,3	23,4	9,6	22,9	22,1	18,8	12,6	29,6	19,7	22,9	20,7
Ten.	27,3	23,9	15,8	17,0	26,8	16,6	17,6	24,4	11,9	23,6	18,2	18,5	15,4	29,1	19,2	22,0	20,5
Alm.	25,3	20,4	18,3	16,3	28,8	19,8	20,0	20,1	9,3	21,1	22,5	23,0	16,6	27,6	18,0	20,7	20,5
Cád.	24,4	21,0	14,8	13,2	24,8	14,1	18,1	22,4	10,1	21,6	20,0	16,3	10,4	26,7	17,7	21,4	18,6
Cór.	25,5	22,1	16,4	15,1	26,5	15,0	18,8	22,8	11,8	23,8	21,5	17,9	11,2	28,7	19,2	22,8	19,9
Gra.	22,8	19,3	16,3	14,1	24,4	13,7	19,43	19,5	11,7	21,2	18,8	15,7	10,0	25,1	14,5	20,6	17,9
Hue.	24,4	20,0	15,2	16,0	27,4	15,9	20,3	26,0	13,5	22,1	21,6	19,0	11,5	27,7	18,6	21,8	20,1
Jaén	24,0	21,4	17,5	15,4	25,0	15,3	19,0	19,6	11,3	22,2	19,8	16,9	11,2	26,6	17,5	21,3	19,0
Mál.	22,4	19,3	15,0	13,8	23,1	13,3	17,3	20,4	11,1	22,1	18,2	15,2	10,5	24,7	16,1	19,3	17,6
Sev.	25,9	22,0	16,4	15,6	25,6	14,9	18,8	23,0	11,5	21,7	20,0	16,8	10,7	27,8	17,8	22,2	19,4
Ceu.	25,7	21,1	16,3	16,3	26,4	17,1	19,8	22,4	11,3	24,3	20,6	17,8	10,8	28,3	19,0	22,8	20,0
Mel.	24,6	19,8	15,7	15,6	23,3	13,6	16,6	21,2	10,8	20,4	16,3	14,4	8,8	26,2	15,4	18,7	17,6
Cue.	16,6	22,9	18,7	17,5	26,9	20,6	19,6	26,5	16,3	24,0	21,2	18,2	17,4	27,1	24,1	23,8	21,3
CLM.	22,2	20,7	16,5	13,9	24,1	15,4	18,9	23,6	13,9	21,6	18,8	16,6	11,8	26,1	17,4	20,9	18,9

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Mad.	25,7	21,9	16,3	17,0	27,4	16,7	20,9	22,6	12,0	19,1	13,2	9,1	12,0	30,0	20,3	23,0	19,2
Ávi.	22,9	21,1	15,8	15,4	24,6	16,4	18,6	21,9	14,3	21,2	19,9	17,3	11,9	28,6	18,8	22,3	19,4
Bur.	22,3	21,8	15,3	14,9	25,0	15,9	18,1	23,7	14,5	21,9	20,0	18,7	12,4	27,2	19,2	23,7	19,7
Sala.	23,3	22,2	16,5	16,0	25,7	16,7	19,5	23,0	15,6	22,3	21,9	18,9	12,3	29,1	20,8	23,1	20,4
Seg.	24,7	22,9	16,7	14,1	25,8	16,6	19,8	24,2	15,8	22,4	21,7	18,2	12,2	27,5	18,8	22,5	20,2
Sor.	24,8	23,2	17,0	16,4	27,7	17,7	20,5	26,1	15,6	24,0	22,5	19,8	13,1	29,2	21,0	22,9	21,3
Zam.	22,8	21,8	15,7	15,1	25,1	16,4	18,1	22,1	13,8	21,9	21,4	17,3	12,4	26,2	19,6	22,3	19,5
Ast.	24,9	21,8	14,8	14,2	25,3	15,0	18,9	21,3	12,4	21,6	19,0	16,1	10,6	28,5	18,3	22,0	19,0
Cant.	22,4	21,1	14,2	14,3	25,5	15,0	17,9	23,0	12,1	21,0	19,2	17,1	10,4	26,5	17,2	20,6	18,6
Ara.	26,4	23,6	17,0	15,8	28,3	17,7	20,5	24,1	14,0	23,3	22,2	19,7	13,2	29,9	20,3	23,4	21,2
Ali.	24,5	20,7	15,7	15,8	24,6	15,7	17,1	22,0	12,9	21,4	20,2	17,7	11,6	26,1	18,9	22,3	19,2
Val.	27,0	23,6	15,7	14,8	27,8	16,2	20,5	22,5	12,5	23,8	20,8	18,6	11,7	28,7	19,1	21,9	20,3
Gal.	29,3	24,7	18,1	16,5	28,5	17,3	18,8	24,7	14,7	23,6	21,2	19,6	12,9	29,8	18,0	22,5	21,3
Lér.	27,1	23,7	13,9	12,4	25,8	14,4	16,3	22,5	11,6	22,8	19,3	17,5	9,2	26,8	16,8	20,2	18,8
Billb.	21,9	19,6	13,0	13,6	24,6	13,9	16,5	20,7	11,1	20,1	18,0	16,3	9,8	24,4	17,1	19,5	17,5
x	24,2	21,3	15,8	15,1	25,6	15,6	18,5	22,1	12,3	21,7	19,5	17,2	11,5	27,2	18,2	21,5	19,21

diferencias significativas entre los promedios de este estudio y los que presenta la clasificación de Samper Padilla y Samper Hernández (2006) se dan en ‘La escuela’ (21,9 por 23,0), ‘La ropa’ (21,4 por 22,2) e ‘Iluminación, calefacción y medios de airear un recinto’ (11,7 por 12,8). Las diferencias de rango afectan a las parejas de centros de interés con valores similares ‘La ciudad’ (4) y ‘La escuela’ (5), y ‘Profesiones y oficios’ (6) y ‘La ropa’ (6), en los que el orden en la pareja cambia. Como ya notaban los estudiosos canarios en 2006, los ámbitos mencionados habían mejorado su posición con respecto al promedio que habían obtenido en 2003. Sin duda, el aumento del promedio de palabras por informante que han ido presentando las investigaciones con el paso de los años, y en esos campos léxicos en particular, no se debe a una mejora real de la producción léxica de los encuestados, sino a un cambio en los criterios de edición. De este modo, cada vez son más los investigadores (Mateo García, 1998; Galloso Camacho, 2003; Bartol Hernández, 2004a; Hernández Muñoz, 2004, 2006; Ávila Muñoz, 2006; Ávila Muñoz y Villena Ponsoda, 2010; Trigo Ibáñez, 2011, entre otros) que adoptan un criterio laxo en la edición de los materiales, optando por mantener todas las formas en los listados, ya pertenezcan directamente al campo léxico estudiado o ya sean fruto de asociaciones secundarias (véase § 2.2).

En lo que se refiere a la dispersión de la distribución, hay que destacar la amplitud del intervalo entre el valor máximo y el valor mínimo en los centros ‘Medios de transporte’ (23 palabras en Almería por 14,4 palabras en Melilla) y ‘Trabajos del campo y del jardín’ (16,6 palabras en Almería por 8,8 palabras en Melilla), coincidiendo el máximo y mínimo en ambas comunidades en las dos ocasiones. Por el contrario, la distribución más agrupada es la que se da en el centro ‘Muebles de la casa’, lo que delata un conjunto de respuestas más estable.

Para concluir, el Cuadro 3.5 muestra los promedios por informante de los trabajos publicados hasta 2014, sea cual sea el tamaño de la muestra. Estas investigaciones son: Puerto Rico (López Morales, 1999a), República Dominicana (Alba, 1995a), Gran Canaria (Samper Padilla y Hernández Cabrera, 1997), Tenerife (Torres González y García Falero, 2009), Almería (Mateo García, 1998), Cádiz (González Martínez, 1997, 2002), Córdoba (Bellón Fernández, 2003, 2011), Granada (Pastor Milán y Sánchez García, 2008), Huelva (Prado Aragonés y Galloso Camacho, 2005, 2006), Jaén (Ahumada Lara, 2006), Málaga (Ávila Muñoz, 2006), Sevilla (Trigo Ibáñez, 2011), Ceuta (Ayora Esteban, 2003, 2006), Melilla (Fernández Smith *et al.*, 2008), Cuenca (Hernández Muñoz, 2004), Castilla-La Mancha (Hernández Muñoz, 2006), Madrid (Benítez Pérez, 1992b, 1995), Ávila (Galloso Camacho, 2003a), Burgos (Fernández Juncal, 2008), Salamanca (Galloso Camacho, 2003a), Segovia (Santiago Guervós, 2008), Soria (Bartol Hernández, 2004a), Zamora (Galloso Camacho, 2003a), Asturias (Carcedo González, 2001), Cantabria (Fernández Juncal, 2013), Ara-

gón (Arnal Purroy, 2004, 2008a), Alicante (Martínez Olmos, 2007), Valencia (Gómez Devís, 2004), Galicia (López Meirama, 2008, 2011), Lérída (Serrano Zapata, 2002) y Bilbao (Etxebarría, 1996, 1999).

**Cuadro 3.6. Comparación del promedio de palabras por centro de interés:  
Muestras diferentes vs. Muestras aprox. 400 informantes**

C. I.	$X \rightarrow \sum$	Rango	$X \approx 400$	Rango
14. ANI	27,2	1	27,3	1
05. ALI	25,6	2	25,7	2
01. CUE	24,2	3	24,8	3
08. ESC	22,1	4	21,9	5
10. CIU	21,7	5	22,1	4
16. PRO	21,5	6	21,4	6
02. ROP	21,3	7	21,4	6
11. CAM	19,5	8	19,8	8
07. COC	18,5	9	18,6	9
15. JUE	18,2	10	17,8	10
12. TRA	17,2	11	17,4	11
03. CAS	15,8	12	16,0	12
06. MES	15,6	13	15,4	13
04. MUE	15,1	14	15,1	14
09. ILU	12,3	15	11,7	15
13. TRC	11,5	16	11,6	16
Media	19,2		19,2	

Como muestra el Cuadro 3.6, pasar de los 14 estudios con muestras en torno a 400 informantes a los 31 con diferentes tamaños muestrales no implica, en este caso, variación significativa alguna en los promedios calculados. Las diferencias son en cada centro de interés de unas pocas décimas, y la jerarquía de ámbitos léxicos es idéntica. Incluso el valor medio de la totalidad de los 16 centros de interés clásicos es coincidente: 19,2 palabras por informante. Por lo tanto, se demuestra que, para los tamaños de muestras manejados en las investigaciones del Proyecto Panhispánico, con un mínimo de 80 informantes en Ávila (Galloso Camacho, 2003a) y un máximo de 800 informantes en Galicia (López Meirama, 2008, 2011), el parámetro que mide el promedio de palabras por informante es el valor más fiable a la hora de establecer comparaciones dentro de este macroproyecto.

### 3.3.2. Número total de palabras

Volviendo a las investigaciones con muestras en torno a 400 informantes, comparamos ahora el número total de palabras producidas en cada centro de interés.

**Cuadro 3.7. Total de palabras por centros de interés en investigaciones del Proyecto Panhispanico con muestras en torno a 400 informantes**

C. I.	Ten.	Alm.	Cád.	Cór.	Gra.	Jaén	Mál.	Sev.	Mel.	Ast.	Cant.	Arag.	Ali.	Val.	Media
01. CUE	12245	10138	9773	10206	9111	9602	10300	10348	7760	11683	7453	11029	11288	12560	10250
02. ROP	10740	8178	8402	8829	7715	8559	8860	8789	6253	10233	7020	9853	9504	10987	8852
03. CAS	7104	7323	5909	6548	6534	6984	6894	6578	4941	6943	4731	7093	7238	7287	6579
04. MUE	7614	6531	5302	6054	5644	6173	6347	6245	4926	6682	4747	6584	7261	6896	6215
05. ALI	12032	11523	9931	10582	9749	10020	10593	10224	7336	11987	8171	11787	11306	12952	10585
06. MES	7447	7934	5624	6016	5467	6103	6113	5952	4272	7054	4994	7385	7219	7516	6364
07. COC	7887	8001	7252	7520	7773	7605	7925	7508	5241	8875	5957	8537	7872	9551	7679
08. ESC	10936	8037	8948	9116	7818	7840	9378	9210	6673	10029	7650	10061	10103	10483	9020
09. ILU	5357	3713	4026	4735	4700	4513	5085	4583	3402	5805	4043	5827	5941	5802	4824
10. CIU	10610	8425	8657	9532	8491	8876	10153	8689	6435	10164	6987	9711	9864	11053	9118
11. CAM	8180	8989	7987	8598	7524	7922	8371	8003	5129	8940	6380	9260	9284	9679	8160
12. TRA	8285	9200	6504	7151	6266	6750	7001	6728	4532	7588	5696	8225	8138	8631	7192
13. TRC	6918	6661	4153	4488	3990	4467	4820	4273	2777	4989	3460	5490	5335	5453	4805
14. ANI	13060	11052	10683	11464	10035	10657	11343	11128	8244	13386	8840	12466	12000	13368	11266
15. JUE	8604	7191	7064	7691	5798	7010	7392	7107	4856	8618	5720	8485	8714	8882	7367
16. PRO	9893	8273	8562	9116	8232	8531	8851	8873	5898	10332	6872	9769	10269	10204	8834
Media	9182	8198	7424	7978	7178	7601	8089	7765	5542	8957	6170	8848	8833	9456	7944

En el Cuadro 3.7, adjunto, en buena lógica, no se registran diferencias apreciables con respecto a la ordenación de los campos léxicos según el promedio de palabras por informante del apartado anterior. Lllaman, eso sí, la atención las variaciones que se presentan con respecto a los valores máximos y mínimos. En los valores máximos, estas variaciones se producen a favor de las muestras de mayor tamaño: Asturias (470), Valencia (465), Alicante (460), Tenerife (449). Por su parte, los valores mínimos se trasladan a las muestras de menor tamaño: Melilla (315) y Cantabria (333). Ello demuestra, por tanto, la dificultad de establecer comparaciones fiables a partir de los valores absolutos, y la importancia, por consiguiente, de ceñirse a un tamaño de muestra común que facilite este tipo de comparaciones, o, en todo caso, la conveniencia de manejar el promedio de palabras por informante como modo de cuantificar la disponibilidad léxica para las comparaciones intermuestrales.

**Cuadro 3.8. Centros de interés, ordenados por rangos, según el total de palabras en investigaciones del Proyecto Panhispánico con muestras en torno a 400 informantes**

C. I.	Ten.	Alm.	Cád.	Cór.	Gra.	Jaén	Mál.	Sev.	Mel.	Ast.	Cant.	Arag.	Ali.	Val.	Rango Medio	Media
14. ANI	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1,07	1
05. ALI	3	1	2	2	2	2	2	3	3	2	2	2	2	2	2,14	2
01. CUE	2	3	3	3	3	3	3	2	2	3	4	3	3	3	2,86	3
10. CIU	6	6	5	4	4	4	4	7	5	6	6	7	6	4	5,29	4
08. ESC	4	9	4	5	6	8	5	4	4	7	3	4	5	6	5,29	5
02. ROP	5	8	7	7	8	5	6	6	6	5	5	5	7	5	6,07	6
16. PRO	7	7	6	5	5	6	7	5	7	4	7	6	4	7	5,93	7
11. CAM	10	5	8	8	9	7	8	8	9	8	8	8	8	8	8,00	8
07. COC	11	10	9	10	7	9	9	9	8	9	9	9	11	9	9,21	9
15. JUE	8	13	10	9	12	10	10	10	12	10	10	10	9	10	10,21	10
12. TRA	9	4	11	11	11	12	11	11	13	11	11	11	10	11	10,50	11
03. CAS	14	12	12	12	10	11	12	12	10	13	14	13	13	13	12,21	12
06. MES	13	11	13	14	14	14	14	14	14	12	12	12	14	12	13,07	13
04. MUE	12	15	14	13	13	13	13	13	11	14	13	14	12	14	13,14	14
09. ILU	16	16	16	15	15	15	15	15	15	15	15	15	15	15	15,21	15
13. TRC	15	14	15	16	16	16	16	16	16	16	16	16	16	16	15,71	16

En el Cuadro 3.8, los rangos muestran también un comportamiento similar al promedio de palabras por informante. Las variaciones que aparecen son insignificantes, y son producto, en cualquier caso, de lo próximo de los valores absolutos en cada caso: por ejemplo, ‘La ropa’ (8852), ‘Profesiones y oficios’ (8834).

El Cuadro 3.9 muestra que la diferencia más apreciable se refiere a la dispersión de los valores absolutos en cada centro de interés. La dispersión mayor se presenta en los tres centros más productivos (‘Los animales’, ‘Alimentos y bebidas’ y

‘El cuerpo humano’), donde se potencia el efecto del tamaño de la muestra. De manera análoga, las distribuciones más agrupadas se dan en ‘Partes de la casa’, ‘Los muebles de la casa’ e ‘Iluminación y calefacción’, donde los campos léxicos se hallan más claramente definidos y el número de respuestas suele ser más estable. La distribución de los campos léxicos en intervalos de 2000 palabras de amplitud define los mismos cuatro grupos del apartado anterior en cuanto a la producción léxica de los mismos.

**Cuadro 3.9. Centros de interés, según la media general del total de palabras en investigaciones del Proyecto Panhispanico con muestras en torno a 400 informantes**

Rango	Rango Medio	Centros de interés	Promedio	Intervalo	Desviación
1	1,07	14. Los animales	11266	10000–12000	1555,735
2	2,14	05. Alimentos y bebidas	10585		1530,867
3	2,86	01. El cuerpo humano	10250		1494,826
4	<u>5,29</u>	10. La ciudad	9118	8000–10000	1309,238
5	<u>5,29</u>	08. La escuela	9020		1254,332
6	6,07	02. La ropa	8852		1347,591
7	5,93	16. Profesiones y oficios	8834		1288,100
8	8	11. El campo	8160		1214,788
MEDIA	–	–	7944		1125,328
9	9,21	07. La cocina	7679	6000–8000	1076,889
10	10,21	15. Juegos y distracciones	7367		1251,389
11	10,5	12. Medios de transporte	7192		1256,047
12	12,21	03. Partes de la casa	6579		830,879
13	13,07	06. Objetos sobre la mesa para...	6364		1082,802
14	13,14	04. Los muebles de la casa	6215		831,650
15	15,21	09. Iluminación, calefacción	4824		4000–6000
16	15,71	13. Trabajos del campo y del jardín	4805	1130,777	

### 3.3.3. Número de vocablos

Otro criterio fundamental para estudiar las divergencias entre los centros de interés es la cantidad de vocablos (o palabras diferentes) que éstos generan. La cantidad de vocablos de que se compone cada centro informa de la homogeneidad o heterogeneidad de las respuestas actualizadas por los informantes, es decir, de la mayor o menor amplitud de asociaciones que genera cada área temática (Samper Padilla *et al.*, 2003: 60).

**Cuadro 3.10. Total de vocablos por centros de interés en investigaciones del Proyecto Panhispánico con muestras en torno a 400 informantes**

C. I.	Ten.	Alm.	Cád.	Cór.	Gra.	Jaén	Mál.	Sev.	Mel.	Ast.	Cant.	Arag.	Ali.	Val.	Me- dia
01. CUE	434	385	296	339	266	335	380	354	283	336	278	356	328	339	336
02. ROP	354	312	251	365	205	361	515	381	282	252	388	332	323	330	332
03. CAS	349	347	186	433	386	577	373	303	282	263	322	346	350	263	341
04. MUE	494	494	152	452	245	345	425	361	371	252	362	363	442	308	362
05. ALI	822	704	495	636	447	565	726	653	536	589	636	649	626	561	617
06. MES	408	507	222	455	226	383	395	362	315	297	353	324	397	256	350
07. COC	551	642	373	612	353	518	527	551	465	383	480	520	519	372	490
08. ESC	852	643	452	669	445	558	582	653	510	475	716	669	552	473	589
09. ILU	346	227	161	445	559	538	405	434	292	455	551	442	540	256	404
10. CIU	959	1151	736	1073	925	890	936	1031	707	896	901	989	1011	755	926
11. CAM	1103	1231	949	1225	1041	991	1054	1235	778	966	982	1224	1196	967	1067
12. TRA	355	329	216	426	315	383	401	420	258	234	458	403	316	280	342
13. TRC	856	1053	380	856	593	759	623	843	547	469	700	547	755	491	677
14. ANI	447	447	503	706	365	440	662	621	442	429	524	542	419	467	501
15. JUE	789	1227	692	927	762	975	734	1043	621	718	936	894	864	546	838
16. PRO	762	1000	589	845	597	692	769	825	675	634	708	761	824	715	743
Media	618	669	416	654	483	582	594	629	460	478	581	585	591	461	557

En la producción de vocablos (ver Cuadro 3.10), solo parece haber incidido el mayor tamaño de las muestras en los valores máximos de Tenerife (4) y Málaga (1). No obstante, no parece ser este el único factor influyente en un mayor aporte de palabras diferentes, por cuanto que Asturias, Alicante y Valencia, con una muestra que supera los 400 informantes en unos 60 a 70 individuos más, no obtienen ningún valor máximo entre las comunidades de habla comparadas. Además, el mayor número de máximos se halla en Almería (7), al tiempo que también aparece un valor máximo en distintos campos léxicos de Córdoba, Granada, Jaén, Sevilla –con 400 informantes– e incluso en Cantabria, comunidad esta última que contaba con una muestra más pequeña (333 informantes). Los valores mínimos se reparten entre Cádiz (8), Granada (6), Melilla (2) y Valencia (1).

Nuevamente los valores obtenidos en la distribución general de vocablos son un reflejo tanto de las características sociológicas de la muestra como de la manera en la que se han editado los materiales. Un ejemplo del efecto de la caracterización sociológica de la muestra, como ya apuntamos en § 3.3.1, se percibe en la menor producción léxica de Granada, la cual es un reflejo de que la mayoría de los informantes procedan de familias con un nivel de instrucción inferior a otras comunidades estudiadas.

En cuanto al efecto de la actitud del investigador ante el corpus obtenido, la aceptación de una mayor extensión y amplitud en las asociaciones posibles en cada ámbito conlleva un repertorio de vocablos más amplio. Así ocurre en Almería, donde los centros que tienen un mayor número de vocablos están todos ellos en o por encima de 1000: ‘La ciudad’ (1151), ‘El campo’ (1231) –solo superado en cuatro unidades por Sevilla–, ‘Trabajos del campo y del jardín’ (1053), ‘Juegos y distracciones’ (1227), ‘Profesiones y oficios’ (1000). Señala Mateo García (1998: 160) que “ese mayor volumen de vocablos [...] no hace más que reflejar en gran medida la realidad cotidiana y lo que podríamos considerar parte de las «vivencias colectivas», puesto que a nuestros encuestados se les pidió que actualizaran todo aquel léxico que le sugiriese el epígrafe señalado”. Mateo García añade, además, que “en estos centros, sobre todo en ‘La ciudad’, ‘El campo’ y ‘La mar’ –[centro que ella suma a esta investigación]– los adjetivos abstractos tienen una presencia muy elevada, debida a las diversas sensaciones que crean en los hablantes (ansiedad, admiración, pena, paz, tranquilidad, etc.)”.

Por otra parte, la manera de lematizar los grupos léxicos, sea regidos por un nombre o por un verbo, con mayor o menor fidelidad al corpus, mayor o menor manipulación, puede llegar a dar visiones significativamente diferentes de la realidad. Precisamente estos grupos léxicos tienen un gran protagonismo en los centros de interés con valores máximos en Almería; con respecto a la muestra general los grupos léxicos representan los siguientes porcentajes: ‘Objetos sobre la mesa para la comida’ (50,8%), ‘La cocina y sus utensilios’ (41,9%), ‘Trabajos del campo y del jardín’ (69,4%), ‘Juegos y distracciones’ (43,4%). Ahora bien, compárense los dos valores extremos (máximo y mínimo) en el Cuadro 3.10 en el centro de interés ‘Juegos y distracciones’: en Almería (1227 vocablos), en Valencia (546 vocablos). La diferencia es tal que se llega a duplicar el repertorio de una a otra comunidad; una situación similar se da también con respecto a ‘Trabajos del campo y del jardín’: en Almería (1053) y Sevilla (1043) frente a Valencia (491), Asturias (469 vocablos) y Cádiz (380). En este caso, los valores más bajos coinciden con las investigaciones en la que se fusionan las combinaciones “verbo +complemento nominal” (*montar a caballo, ir al cine, ir/salir con los amigos, jugar a las cartas, tocar la guitarra, escuchar música*) por el núcleo nominal: *caballo, cine, amigos, cartas, guitarra, música*; de otra parte, los valores más altos corresponden a las investigaciones donde se mantienen todas las combinaciones como entradas diferentes, sea una asociación sintagmática esperable o no (*tocar un instrumento; componer música*).

En cuanto al orden de los centros de interés, según las cifras totales de vocablos (ver Cuadro 3.11), la primera posición es ocupada de manera constante, en todas las comunidades, por el centro de interés ‘El campo’, seguido de manera casi unánime por ‘La ciudad’ en la segunda posición. Ambos son campos léxicos con una gran amplitud de evocaciones.

A partir de la tercera posición se inicia una paulatina diversificación en los rangos asignados en las distintas comunidades comparadas. En cuanto a los valores inferiores, el rango medio muestra que no existen grandes diferencias en la producción de vocablos hasta en cuatro centros, todos ellos agrupados en torno al valor 13. Estos centros de interés son: ‘Medios de transporte’, ‘Partes de la casa’, ‘El cuerpo humano’ y ‘La ropa’. Todos estos campos léxicos se componen de un vocabulario fácil de acotar.

**Cuadro 3.11. Centros de interés, ordenados por rangos, según el total de vocablos en investigaciones del Proyecto Panhispánico con muestras en torno a 400 informantes**

C. I.	Ten.	Alm.	Cád.	Cór.	Gra.	Jaén	Mál.	Sev.	Mel.	Ast.	Cant.	Arag.	Ali.	Val.	Rango Medio	Media
11. CAM	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
10. CIU	2	3	2	2	2	3	2	3	2	2	3	2	2	2	2,29	2
15. JUE	6	2	3	3	3	2	4	2	4	3	2	3	3	5	3,21	3
16. PRO	7	5	4	5	4	5	3	5	3	4	5	4	4	3	4,36	4
13. TRC	3	4	8	4	5	4	7	4	5	7	6	7	5	6	5,36	5
05. ALI	5	6	6	8	7	7	5	<u>6</u>	6	5	7	6	6	4	6,00	6
08. ESC	4	7	7	7	8	8	8	<u>6</u>	7	6	4	5	7	7	6,5	7
14. ANI	10	11	5	6	10	11	6	8	9	9	9	8	11	8	8,64	8
07. COC	8	8	9	9	11	10	9	9	8	10	10	9	9	9	9,14	9
09. ILU	16	16	15	12	6	9	12	10	12	8	8	10	8	16	11,29	10
04. MUE	9	10	16	11	14	15	11	14	10	15	13	12	10	12	12,29	11
06. MES	12	9	12	10	15	13	14	13	11	12	14	16	12	16	12,79	12
12. TRA	13	14	13	14	12	12	13	11	16	16	11	11	16	13	13,21	13
03. CAS	15	13	14	13	9	6	16	16	<u>14</u>	13	15	14	13	15	13,29	14
01. CUE	11	12	10	16	13	16	15	15	13	11	16	13	14	10	13,21	15
02. ROP	14	15	11	15	16	14	10	12	<u>14</u>	14	12	15	15	11	13,43	16

Como se ve, la distribución de los valores jerarquizados presenta menos uniformidad que la que reflejaba la productividad de palabras. Ello, más allá de posibles diferencias en los criterios de edición de los listados léxicos, puede ser explicado también por la influencia de diferentes realidades psicosociales y medioambientales del entorno.

La ordenación de los centros de interés según la media de vocablos (ver Cuadro 3.12) muestra claramente la muy diferente riqueza léxica de los 16 centros de interés. No es ésta una distribución uniforme en su ordenamiento, apreciándose fácilmente un agrupamiento asimétrico de los cuatro grupos en intervalos de 200 unidades:

**Cuadro 3.12. Centros de interés, según la media general del total de vocablos en investigaciones del Proyecto Panhispánico con muestras en torno a 400 informantes**

Rango	Rango Medio	Centros de interés	Promedio	Intervalo	Desviación
1	1	11. El campo	1067	900–1100	134,738
2	2,29	10. La ciudad	926		122,988
3	3,21	15. Juegos y distracciones	838	700–900	173,311
4	4,36	16. Profesiones y oficios	743		105,980
5	5,36	13. Trabajos del campo y del jardín	677	500–700	180,513
6	6	05. Alimentos y bebidas	617		93,183
7	6,5	08. La escuela	589		113,290
MEDIA	–	–	557		78,189
8	8,64	14. Los animales	501		96,048
9	9,14	07. La cocina	490	300–500	87,781
10	11,29	09. Iluminación, calefacción	404		124,471
11	12,29	04. Los muebles de la casa	362		95,054
12	12,79	06. Objetos sobre la mesa para...	350		80,098
13	<u>13,21</u>	12. Medios de transporte	342		73,514
14	13,29	03. Partes de la casa	341		88,193
15	<u>13,21</u>	01. El cuerpo humano	336		44,255
16	13,43	02. La ropa	332		72,464

– dos centros con una variedad léxica extraordinariamente alta: ‘El campo’ (1067 vocablos) y ‘La ciudad’ (926) vocablos;

– en un segundo escalón se sitúa ‘Juegos y distracciones’ (838 vocablos) y ‘Profesiones y oficios’ (743 vocablos);

– un tercer grupo entre 500 y 700 vocablos, en el que se halla la media, donde se cuentan ‘Trabajos del campo y del jardín’ (677), ‘Alimentos y bebidas’ (617), ‘La escuela’ (589) y ‘Los animales’ (501);

– y un último y numeroso grupo, con la mitad de los centros de interés clásicos, en el que, salvo ‘La cocina’ (490 vocablos), las demás categorías presentan unos valores muy próximos con una amplitud de 72 unidades. Estos centros son: ‘Iluminación y calefacción’ (404), ‘Muebles de la casa’ (362), ‘Objetos sobre la mesa para la comida’ (350), ‘Medios de transporte’ (342), ‘Partes de la casa’ (341), ‘El cuerpo humano’ (336) y ‘La ropa’ (332).

En sus respectivos estudios, Samper Padilla *et al.* (2003: 62–63), y Samper Padilla y Samper Hernández (2006: 33–34) también ofrecen una ordenación de los centros de interés a partir de las cifras de vocablos, aunque reconocen en ambos casos que algunas de estas cifras son muy dispares, ya que “estos datos

cuantitativos están condicionados por el número de informantes que constituyen las muestras de las distintas comunidades” (Samper Padilla y Samper Hernández, 2006: 31, n. 16). Recuérdese que, para evitar esta influencia, aquí hemos excluido las investigaciones con muestras con valores alejados a cardinal 400.

En el caso de las dos ordenaciones citadas, los estudiosos canarios ya notaban un notable aumento del número de vocablos que conseguían los centros de interés 08, 12 y 2, y un descenso en los rangos en los campos 04 y 01. En cuanto a la ordenación que aquí presentamos a partir de muestras más similares cuantitativamente, según el orden de rangos se asemeja más a la primera de ellas (Samper Padilla *et al.*, 2003). Hay cinco categorías que aparecen con un rango diferente, si bien en el caso ‘Los animales’, ‘La cocina’ y ‘La ropa’ la variación es de solo un rango. ‘El cuerpo humano’ varía en dos puntos y ‘Medios de transporte’ en tres posiciones. En el caso de la ordenación de Samper Padilla y Samper Hernández (2006), la clasificación es mucho más dispar, mostrando hasta 10 variaciones a partir del rango 5.

No obstante, tanto en las clasificaciones del equipo canario como en la que aquí manejamos, a lo que hay que prestar atención es al contraste de la distribución del promedio de vocablos con la tradicional –y a veces engañosa– ordenación por rangos. En efecto, la distribución del promedio de vocablos no sigue la progresión aritmética que sugiere la ordenación por rangos, sino que presenta una ordenación discontinua y asimétrica. En nuestra clasificación, cuatro campos léxicos se reparten en los dos intervalos superiores, y hasta ocho centros de interés ocupan el último intervalo. Por ello, manejar el rango medio da una imagen más fidedigna de la realidad, cuando no se parte de muestras similares en su tamaño.

En lo que sí coinciden las tres clasificaciones es en señalar la existencia de cuatro centros de interés (‘El campo’, ‘La ciudad’, ‘Juegos y distracciones’ y ‘Profesiones y oficios’) con un vocabulario muy amplio, pero al mismo tiempo con unas posiciones muy estables en todas las distribuciones. También los dos intervalos inferiores determinan un comportamiento muy aproximado en el ordenamiento de la producción léxica, aunque con variaciones en los rangos. Estas variaciones se deben, no solo a la caracterización sociolingüística y a matices en la edición del corpus, sino, sobre todo, a lo aproximado de las cantidades de vocabulario que integran esos centros de interés. Ello provoca que una pequeña variación en alguno de los factores anteriormente mencionados tenga un efecto inmediato en la ordenación del campo léxico con respecto a las restantes categorías. Nótese así lo aproximado de muchos de los valores en todas las clasificaciones a partir de media tabla.

A la vista de los cuadros anteriores (promedio de palabras por informante, palabras totales y total de vocablos), se puede apreciar que la mayor o menor riqueza de vocablos no se corresponde con la productividad de unidades léxicas en las diferentes áreas temáticas; es decir, no se cumple la máxima: a mayor número de respuestas, mayor número de palabras diferentes.

La comparación del promedio de palabras por informante y el total de vocablos que presenta el Cuadro 3.13 muestra el diferente comportamiento de los centros de interés con respecto a estas magnitudes. Dado que se trata de magnitudes diferentes, efectuamos la comparación a partir del rango medio de productividad de unidades léxicas por sujeto y de vocablos en las sintopías analizadas. Así se obtiene una visión más aproximada de la realidad.

El contraste es evidente. Los centros con mayor número de respuestas son ‘Los animales’ (1), ‘Alimentos y bebidas’ (2), ‘El cuerpo humano’ (3). Estos descienden en la producción de vocablos a las posiciones (9), (6) y (13) respectivamente. Por su parte los tres primeros rangos en la producción de vocablos están ocupados por ‘El campo’ (1), ‘La ciudad’ (2) y ‘Juegos y distracciones’ (3), que en el número de palabras por informante ocupaba los rangos (8), (5) y (10) respectivamente.

Por la parte baja de la clasificación, los centros menos productivos en palabras son ‘Iluminación y calefacción’ (15) y ‘Trabajos del campo y del jardín’ (16), con rangos más altos en el número de vocablos, (11) y (5) respectivamente. Del mismo modo, la clasificación de los centros de interés, según el número de vocablos se cierra en la posición (13) compartida, con ligeras diferencias por ‘Objetos sobre la mesa para la comida’, ‘El cuerpo humano’, ‘Medios de transporte’, ‘Partes de la casa’ y ‘La ropa’, centros de interés todos ellos con una mayor producción de palabras, especialmente ‘El cuerpo humano’, que llega al rango (3) en el número de respuestas.

**Cuadro 3.13. Comparación de los rangos del promedio de palabras por informante y de la media general de vocablos de los 16 centros de interés tradicionales**

Palabras por informante	Rango	Vocablos
1	2	3
14. Los animales (1,07)	1	11. El campo (1)
05. Alimentos y bebidas (2,14)	2	10. La ciudad (2,29)
01. El cuerpo humano (2,86)	3	15. Juegos y distracciones (3,21)
	4	16. Profesiones y oficios (4,36)
10. La ciudad (5,29) 08. La escuela (5,29)	5	13. Trabajos del campo y del jardín (5,36)
16. Profesiones y oficios (5,93) 02. La ropa (6,07)	6	05. Alimentos y bebidas (6) 08. La escuela (6,5)
	7	
11. El campo (8)	8	
07. La cocina (9,21)	9	14. Los animales (8,64) 07. La cocina (9,14)

1	2	3
15. Juegos y distracciones (10,21) 12. Medios de transporte (10,5)	10	
	11	09. Iluminación, calefacción (11,29)
03. Partes de la casa (12,14)	12	04. Los muebles de la casa (12,29)
06. Objetos sobre la mesa para la comida (13) 04. Los muebles de la casa (13,14)	13	06. Objetos sobre la mesa para la comida (12,79) 01. El cuerpo humano (13,21) 12. Medios de transporte (13,21) 03. Partes de la casa (13,29) 02. La ropa (13,43)
	14	
09. Iluminación, calefacción (15,21)	15	
13. Trabajos del campo y del jardín (15,71)	16	

Por último, en el Cuadro 3.14 se muestra el diferencial entre los rangos medios del promedio de unidades léxicas por sujeto y del número total de vocablos. Un diferencial positivo indica un rango superior en vocablos, en tanto que un diferencial negativo es característico de un campo donde se ha producido un elevado número de respuestas, pero en las que se repiten con más frecuencia los vocablos.

**Cuadro 3.14. Diferencial de los rangos del promedio de palabras por informante y de la media general de vocablos de los 16 centros de interés tradicionales**

+10	‘Trabajos del campo y del jardín’ (+10,35)
+9	
+8	
+7	‘El campo’ (+7) ‘Juegos y distracciones’ (+6,92)
+6	
+5	
+4	‘Iluminación y calefacción’ (+3,92)
+3	‘La ciudad’ (+3)
+2	‘Profesiones y oficios’ (+1,57)
+1	‘Muebles de la casa’ (+0,85)
0	‘Objetos sobre la mesa para la comida’ (+0,21) ‘La cocina’ (+0,07)

Cuadro 3.14 (cont.)

-1	‘Partes de la casa’ (-1,15) ‘La escuela’ (-1,21)
-2	
-3	‘Medios de transporte’ (-2,71)
-4	‘Alimentos y bebidas’ (-3,86)
-5	
-6	
-7	‘La ropa’ (-7,36)
-8	‘Los animales’ (-7,57)
-9	
-10	‘El cuerpo humano’ (-10,35)

Si se establecen entornos simétricos de centro 0 y radio 2, las sucesivas ampliaciones del radio en  $\pm 2$  unidades permiten delimitar los siguientes grupos de campos léxicos:

a) un grupo más estable con un diferencial de  $\pm 2$ , compuesto por ‘La cocina’, ‘La escuela’ y ‘Profesiones y oficios’ con rangos medios entre 4 y 9, y ‘Objetos sobre la mesa para la comida’, ‘Muebles de la casa’ y ‘Partes de la casa’, con rangos bajos entre 12 y 13;

b) un grupo con un diferencial de  $\pm 4$ , compuesto por ‘Alimentos y bebidas’ y ‘La ciudad’ con rangos superiores entre 2 y 6, y ‘Medios de transporte’ e ‘Iluminación y calefacción’ con rangos inferiores entre 10 y 15;

c) dado el conjunto vacío del diferencial  $\pm 6$ , el siguiente grupo queda definido por un diferencial  $\pm(7-7,5)$  con los centros ‘Juegos y distracciones’, ‘El campo’, ‘La ropa’ y ‘Los animales’;

d) finalmente con un diferencial  $\pm 10$  aparecen los dos centros de interés con la relación “palabras por informante – palabras diferentes” más desigual: ‘El cuerpo humano’ y ‘Trabajos del campo y del jardín’.

### 3.3.4. Índice de cohesión

Como acabamos de ver, entre los rangos del promedio de palabras y del número de vocablos se dan divergencias, a veces muy apreciables: centros con una alta productividad en vocablos tienen un promedio de palabras bajo (‘Juegos y distracciones’, ‘El campo’, ‘Trabajos del campo y del jardín’), y viceversa, centros con una baja productividad en vocablos tienen un promedio de palabras alto (‘El cuerpo humano’, ‘Los animales’, ‘La ropa’). Al mismo tiempo, aparecen centros con un alto grado de coincidencia entre los rangos de ambos índices (‘La cocina’, ‘Objetos colocados sobre la mesa para la comida’, ‘Muebles de la casa’, ‘Partes de la casa’, ‘La escuela’).

La explicación a este diferente comportamiento de ambos índices en cada centro de interés radica en el grado de coincidencia de las respuestas entre los informantes. Si aumenta la disparidad en cuanto a las palabras con que los individuos de la muestra responden, aumenta la variedad o cantidad de palabras diferentes en ese determinado centro.

La constatación de este hecho llevó a Max Echeverría (1991; *et al.*, 1987) a introducir el índice de cohesión, que muestra el mayor o menor grado de coincidencia en las respuestas de los informantes. Este índice se obtiene dividiendo el promedio de respuestas por sujeto en cada centro de interés entre el número de palabras diferentes tabuladas en ese centro. El valor que se obtiene oscila entre 0 y 1. Si el índice se acerca a 1 las respuestas se repiten con más frecuencia y el centro es, por tanto, más cerrado o compacto, y si se aproxima a 0 hay una mayor dispersión en las respuestas y, en consecuencia, el centro es más abierto o difuso. De este modo, el índice de cohesión refleja el grado de cohesión o coherencia semántica de los centros, que permite que se les pueda limitar a un número reducido de clases lógicas, cohesión alta, o que hace casi imposible la delimitación de un número de clases lógicas, cohesión baja (Echeverría *et al.*, 1987: 57).

En cuanto a las sintopías que comparamos, el Cuadro 3.15 muestra los valores del índice de cohesión de cada centro de interés. La distribución de valores máximos y mínimos, muy repartida entre todas las comunidades estudiadas, parece descartar una influencia significativa del diferente tamaño de algunas muestras.

En cuanto a los valores máximos destacan Cádiz (9) y Granada (4), donde, siempre en términos relativos, se da una mayor repetición del vocabulario y una menor diversidad del mismo. Los valores mínimos están mucho más repartidos, sobresaliendo solo Málaga (6), donde se tiende a una mayor dispersión en las respuestas.

La ordenación por rangos (ver Cuadro 3.16) permite constatar mejor la gradación compacto/difuso de los diferentes centros de interés. Varios son los hechos dignos de comentario aquí:

- Una presencia muy apreciable en las distintas sintopías de rangos repetidos (aquí subrayados), lo que indica que la distancia entre rangos a veces es muy pequeña e incluso inexistente, según qué centros de interés y en según qué comunidades.

- Es en los extremos de la distribución –rangos de la banda alta (1 a 3) y de la banda baja (14 a 16)– donde aparece más definida la ordenación, con un comportamiento algo más constante de los rangos correspondientes. En cualquier caso, en comparación con los cuadros de rangos del promedio de palabras por informante o del total de vocablos, hay mucha más diversidad en los rangos de cada centro de interés.

**Cuadro 3.15. Índice de cohesión por centros de interés en investigaciones del Proyecto Panhispánico con muestras en torno a 400 informantes**

C. I.	Ten.	Alm.	Cád.	Cór.	Gra.	Jaén	Mál.	Sev.	Mel.	Ast.	Cant.	Arag.	Ali.	Val.	Media
01. CUE	0,063	0,066	0,082	0,075	0,085	0,072	0,059	0,073	0,087	0,074	0,080	0,074	0,075	0,078	0,075
02. ROP	0,067	0,065	0,079	0,060	0,094	0,059	0,037	0,058	0,070	0,086	0,054	0,071	0,064	0,070	0,067
03. CAS	0,045	0,053	0,078	0,038	0,042	0,030	0,040	0,054	0,056	0,062	0,044	0,049	0,045	0,057	0,050
04. MUE	0,034	0,033	0,081	0,033	0,057	0,045	0,032	0,043	0,042	0,063	0,039	0,043	0,036	0,047	0,045
05. ALI	0,033	0,041	0,047	0,042	0,054	0,044	0,032	0,039	0,043	0,042	0,040	0,044	0,039	0,049	0,042
06. MES	0,041	0,039	0,061	0,033	0,060	0,040	0,034	0,041	0,043	0,050	0,042	0,055	0,039	0,060	0,046
07. COC	0,032	0,031	0,045	0,031	0,055	0,037	0,033	0,034	0,036	0,049	0,037	0,039	0,033	0,053	0,039
08. ESC	0,029	0,031	0,047	0,034	0,043	0,035	0,035	0,035	0,041	0,044	0,032	0,036	0,040	0,046	0,038
09. ILLU	0,034	0,041	0,066	0,027	0,021	0,021	0,027	0,026	0,037	0,027	0,022	0,032	0,024	0,049	0,032
10. CIU	0,025	0,019	0,029	0,022	0,022	0,025	0,024	0,021	0,029	0,024	0,023	0,024	0,021	0,031	0,024
11. CAM	0,016	0,018	0,021	0,017	0,018	0,020	0,017	0,016	0,021	0,019	0,019	0,018	0,017	0,022	0,019
12. TRA	0,052	0,070	0,075	0,042	0,049	0,044	0,038	0,040	0,056	0,068	0,037	0,049	0,057	0,065	0,053
13. TRC	0,018	0,016	0,028	0,025	0,016	0,015	0,017	0,013	0,016	0,022	0,015	0,024	0,015	0,023	0,019
14. ANI	0,065	0,062	0,051	0,041	0,068	0,060	0,037	0,045	0,059	0,066	0,051	0,055	0,062	0,061	0,056
15. JUE	0,024	0,015	0,025	0,021	0,019	0,018	0,022	0,017	0,025	0,025	0,018	0,023	0,022	0,035	0,022
16. PRO	0,029	0,021	0,034	0,027	0,034	0,031	0,025	0,027	0,028	0,034	0,029	0,031	0,027	0,030	0,029
Media	0,038	0,039	0,053	0,036	0,037	0,037	0,032	0,036	0,043	0,047	0,036	0,042	0,039	0,049	0,040

**Cuadro 3.16. Centros de interés, ordenados por rangos, según el índice de cohesión en investigaciones del Proyecto Panhispánico con muestras en torno a 400 informantes**

C. I.	Ten.	Alm.	Cád.	Cór.	Gra.	Jaén	Mál.	Sev.	Mel.	Ast.	Cant.	Arag.	Ali.	Val.	Rango Medio	Media
01. CUE	3	2	1	1	2	1	1	1	1	2	1	1	1	1	1,36	1
02. ROP	1	3	3	2	1	3	4	2	2	1	2	2	2	2	2,14	2
14. ANI	2	4	8	5	3	2	4	4	3	4	3	3	3	4	3,71	3
12. TRA	4	1	5	3	8	5	3	7	4	3	8	5	4	3	4,50	4
03. CAS	5	5	4	6	10	11	2	3	4	6	4	5	5	6	5,43	5
06. MES	6	8	7	8	4	7	7	6	6	7	5	3	7	5	6,14	6
04. MUE	7	9	2	8	5	4	9	5	8	5	7	8	9	10	6,86	7
05. ALJ	9	6	9	3	7	5	9	8	6	10	6	7	7	8	7,14	8
07. COC	10	10	11	10	6	8	8	10	11	8	8	9	10	7	9,00	9
08. ESC	11	10	9	7	9	9	6	9	9	9	10	10	6	11	8,93	10
09. ILLU	7	6	6	11	13	13	11	12	10	12	13	11	12	8	10,36	11
16. PRO	11	12	12	11	11	10	12	11	13	11	11	12	11	14	11,57	12
10. CIU	13	13	13	14	12	12	13	13	12	14	12	13	14	13	12,93	13
15. JUE	14	16	15	15	14	15	14	14	14	13	15	15	13	12	14,21	14
13. TRC	15	15	14	13	16	16	16	16	16	15	16	13	16	15	15,14	15
11. CAM	16	14	16	16	15	14	16	15	15	16	14	16	15	16	15,29	15

A partir de la zona central de la jerarquía (rangos 5 al 11) se da una gran variación, de unas comunidades a otras, en el rango de cada centro de interés. Esto ocasiona que, en valor absoluto, dos o más centros de interés tengan el mismo rango medio: ‘Muebles de la casa’ y ‘Alimentos y bebidas’ (rango 7), ‘La cocina’ y ‘La escuela’ (rango 9), ‘Trabajos del campo y del jardín’ y ‘El campo’ (rango 15).

**Cuadro 3.17. Centros de interés, según la media general del índice de cohesión en investigaciones del Proyecto Panhispánico con muestras en torno a 400 informantes**

Rango	Rango Medio	Centros de interés	Promedio	Intervalo	Desviación
1	1,36	01. El cuerpo humano	0,075	> 0,070	0,0079
2	2,14	02. La ropa	0,067	0,050–0,070	0,0140
3	3,71	14. Los animales	0,056		0,0096
4	4,50	12. Medios de transporte	0,053		0,0126
5	5,43	03. Partes de la casa	0,050		0,0119
6	6,14	06. Objetos sobre la mesa para...	0,046		0,0097
7	6,86	04. Los muebles de la casa	0,045	0,030–0,050	0,0138
8	7,14	05. Alimentos y bebidas	0,042		0,0058
MEDIA			0,040		0,0059
9	9,00	07. La cocina	0,039		0,0082
10	8,93	08. La escuela	0,038		0,0058
11	10,36	09. Iluminación, calefacción	0,032	0,010–0,030	0,0126
12	11,57	16. Profesiones y oficios	0,029		0,0037
13	12,93	10. La ciudad	0,024		0,0034
14	14,21	15. Juegos y distracciones	0,022		0,0049
<u>15</u>	15,14	13. Trabajos del campo y del jardín	<u>0,019</u>		0,0047
<u>15</u>	15,29	11. El campo	<u>0,019</u>	0,0019	

Finalmente, la agrupación en intervalos de los valores medios del índice de cohesión, muestra la siguiente división entre los centros de interés (véase Cuadro 3.17):

a) el mayor grado de cohesión, con el conjunto léxico más claramente delimitado, caracteriza al campo léxico ‘El cuerpo humano’;

b) a continuación, los rangos 2 al 5 muestran áreas temáticas con un grado de cohesión muy alto (semánticamente muy compactas y cerradas): ‘La ropa’, ‘Los animales’, ‘Medios de transporte’ y ‘Partes de la casa’;

c) el tercer grupo de categorías léxicas se sitúan en torno a la media entre los rangos 6 a 10: ‘Objetos colocados sobre la mesa para la comida’, ‘Los muebles de la casa’, ‘Alimentos y bebidas’, ‘La cocina’, ‘La escuela’ e ‘Iluminación, calefacción y medios de airear un recinto’;

d) finalmente los campos léxicos con un grado de cohesión muy bajo, con un rango medio superior a 12 (semánticamente muy difusos y abiertos) son ‘Profesiones y oficios’, ‘La ciudad’, ‘Juegos y distracciones’, ‘Trabajos del campo y del jardín’ y ‘El campo’.



## IV. RELACIONES ENTRE LAS PALABRAS EN EL LÉXICO DISPONIBLE

La idea de que las palabras se asocian está implícita en la misma noción de *disponibilidad* establecida por Michéa (1953: 341)<sup>1</sup>, quien la definía como la “propiedad que posee una palabra de ser evocada de una forma más o menos inmediata en el curso de una asociación de ideas”. Este lingüista francés, pionero en el diseño de las pruebas para recabar el vocabulario temático, explicaba además que esta palabra disponible “viene inmediatamente y naturalmente a la mente en el momento en que se tiene necesidad de ella” y que “es una palabra que, formando parte de las asociaciones de ideas usuales, existe en potencia en el sujeto hablante en cuanto estas asociaciones entran en juego” (Michéa, 1953: 310).

Sin embargo, a pesar de referirse a los procesos mentales de asociación y de acceso al léxico, no ha sido hasta pasado el siglo XX cuando los lingüistas del ámbito hispánico se han ocupado de los aspectos psicolingüísticos en las listas aportadas en las pruebas de disponibilidad léxica<sup>2</sup>.

La vinculación de la disponibilidad léxica con la psicolingüística reside en la misma prueba de recogida de léxico disponible. Esta prueba es una tarea artificial en la que se fuerza al individuo a extraer palabras del lexicón mental dentro de una categoría amplia llamada *centro de interés*, que a su vez se divide en subcategorías (Ferreira y Echeverría, 2010: 137). El objetivo principal es fijar las palabras evocadas de forma más o menos inmediata en el curso de una asociación de ideas, a partir de un estímulo léxico. Al mismo tiempo, esta tarea se convierte en una manifestación externa del proceso interno de generación de unidades léxicas (Paredes García, 2012: 89), en las que unas palabras tiran de otras formando cadenas léxicas.

---

<sup>1</sup> Traducción del francés de Carcedo González (2000: 18).

<sup>2</sup> Entre estas primeras aportaciones cabe destacar a Romero Rubilar (2000), quien aplicó una prueba de disponibilidad léxica junto con una prueba de categorización, donde estudiaba el grado de prototipicidad de los ejemplos que aparecieron en las listas, y a Galloso Camacho (2002b), quien, teniendo en cuenta las bases teóricas de la teoría de la red (*cobweb*) como manera de representación del lexicón mental, estableció una taxonomía de las relaciones asociativas observadas en los corpus de disponibilidad.

En consecuencia, los análisis de léxico disponible ayudan a conocer la organización del lexicón mental, o conocimiento que un hablante tiene interiorizado del vocabulario, “pues plasman de modo claro las relaciones que los hablantes establecen entre los vocablos” (Samper Hernández, 2002: 12). Estos análisis muestran, como veremos más adelante, que “son los prototipos y las redes semánticas –que no son incompatibles– las que organizan el lexicón” (Manjón-Cabeza, 2008a).

Desde la perspectiva de la lingüística cognitiva, el punto de inflexión lo ha marcado Hernández Muñoz (2006) con la tesis titulada *Hacia una teoría cognitiva integrada de la disponibilidad léxica*. Esta estudiosa, tras analizar los índices de disponibilidad de las palabras en los diccionarios de léxico disponible y su posición en las listas individuales, llega a la conclusión de que la disponibilidad, en lo que se refiere a la palabra, depende de dos factores:

- “Las características extrínsecas de la palabra dictadas principalmente por el contexto semántico (centro de interés)”.
- “Las características intrínsecas de la palabra que atañen tanto al significado como a la forma” (Hernández Muñoz, 2006: 128).

Ambos factores hacen referencia a un doble eje asociativo que se da en el proceso de actualización de unidades léxicas. Según Paredes García (2012: 91), este doble eje asociativo está compuesto por un eje principal, que vincula la nueva unidad solo con el estímulo generador, y otro secundario, que resulta de las activaciones ocasionadas por la presencia de un término previo (rimas internas, derivación, composición, etc.).

Por tanto, para analizar las relaciones entre palabras en las pruebas de disponibilidad léxica, será necesario tener en cuenta estas dos características o ejes de los que hablan ambos lingüistas. Para ello, será necesario hacer referencia en ambos casos a las dos pruebas de la Psicología Experimental que comparten con la disponibilidad léxica los mismos principios en la activación del léxico. Estas pruebas son:

- 1) La fluencia semántica o fluencia de categoría (*semantic fluency* o *category fluency*).
- 2) La asociación libre de palabras (*free word association*).

## **4.1. La fluencia semántica y la teoría de los prototipos**

### **4.1.1. La elaboración de las listas**

Como recuerda Hernández Muñoz (2006: 29), la prueba de disponibilidad léxica es una situación experimental controlada que reproduce una de las pruebas típicas de la Psicología Experimental, la fluencia semántica. En esta prueba se da

al individuo el nombre de una categoría semántica (‘Partes del cuerpo humano’, ‘Ropa’, ‘Animales’, etc.) y éste tiene que enunciar en un tiempo limitado todas aquellas palabras que pertenezcan a esa categoría.

Por consiguiente, “el primer factor que determina la disponibilidad es la relación del término con su contexto categorial” (Hernández Muñoz, 2006: 129), el cual viene marcado por la etiqueta categorial. Nos situamos, de este modo, en el primer eje asociativo del que habla Paredes García (2012).

Diversos estudiosos de la disponibilidad léxica (Hernández Muñoz, 2006; Ávila Muñoz y Villena Ponsoda, 2010, Sánchez-Saus, 2011; etc.) han coincidido en señalar que a la hora de describir y cuantificar las relaciones establecidas entre los términos de una misma categoría semántica, la teoría más adecuada es la teoría de los prototipos. Basan su conclusión en la cercanía de los conceptos *disponibilidad* y *tipicidad*, y en la posibilidad de gradación de ambas cualidades (Hernández Muñoz, 2006: 129), y postulan “que cada centro de interés gira en torno a un prototipo creado a partir del concepto que determina ese centro de interés” Ávila Muñoz y Villena Ponsoda (2010: 55).

Los estudiosos malagueños observan la acción de los prototipos en el patrón de elaboración de la lista durante el proceso de respuesta a la prueba de disponibilidad léxica. Aplicando el concepto de “accesibilidad”, explican el proceso de respuesta a una prueba de disponibilidad léxica en los siguientes términos:

Cuando un individuo se somete al experimento de disponibilidad léxica, éste accede a su red léxica a partir del prototipo del centro de interés como puerta de entrada. Posteriormente, recorre la red léxica, acción que lo irá alejando, progresivamente, del núcleo prototípico. Ocurre que, de vez en cuando, el individuo percibe que, en un momento determinado, se ha alejado excesivamente del núcleo del centro de interés y busca entonces estrategias que le permitan reentrar en la red léxica original (Ávila Muñoz y Villena Ponsoda, 2010: 55).

En el siguiente ejemplo de una lista léxica del centro ‘Alimentos’, tomada del léxico disponible de la ciudad de Málaga (Ávila Muñoz y Villena Ponsoda, 2010: 55), se aprecia el modo en que estos procesos de reentrada cíclicos delimitan, en las cadenas léxicas producidas, agrupaciones de palabras en subconjuntos asociados<sup>3</sup>.

MÁLAGA. 11221 009 05 fruta, verdura, carne, **proteína, carbohidrato, mineral, vitamina**, cocina, aceite, legumbre, lenteja, huevo, pollo, **corral, campo**, trigo, pan, leche

Ejemplo: Acceso y desarrollo de una red léxica individual

<sup>3</sup> Manjón-Cabeza (2008a) ya había notado este fenómeno al determinar las redes semánticas naturales en las listas de léxico disponible del centro de interés ‘Los colores’, afirmando que “las palabras tienden a agruparse en conglomerados (*clusters*): núcleos de palabras muy relacionadas que se unen con otros de manera más laxa”.

Como se ve, el informante, cuando recibe el estímulo ‘Alimentos’, accede a su red léxica por medio de la palabra *fruta*. Continúa produciendo palabras prototípicas del nivel de base, los lexemas primarios y contextualmente neutros *verdura* y *carne*. Al llegar a este punto, realiza una asociación indirecta relacionada con las ‘propiedades nutricionales’ (*proteína, carbohidrato, mineral, vitamina*). Tras escribir *vitamina*, percibe que se ha alejado del estímulo inicial –‘Alimentos’– y, mediante una palabra puente –*cocina*– regresa a la red léxica de los alimentos *aceite, legumbre, lenteja, huevo, pollo*, alimentos que aparecen los cuatro últimos asociados en parejas, en relación de hiperonimia-hiponimia el primer par (la lenteja es una legumbre), y de metonimia el segundo (de los huevos salen los pollos). La palabra *pollo* aleja al informante del núcleo prototípico propio del centro de interés, y le hace escribir por asociación indirecta los lugares que son hábitat de un pollo (*corral, campo*). Como en el campo se produce el *trigo*, esta palabra que le sirve de nuevo de puente para entrar en la red léxica de los alimentos y escribir nuevamente elementos prototípicos del nivel de base: *pan, leche*.

#### 4.1.2. La categorización

Según la *teoría de los prototipos* de Rosch (1978), la etiqueta categorial de los 16 centros de interés clásicos (*Partes del cuerpo, La ropa, Partes de la casa, Los muebles, Alimentos y bebidas, etc.*) se corresponde con un prototipo del nivel supraordinado, que reúne los atributos comunes a los elementos inferiores. Los sujetos que responden a una prueba de disponibilidad léxica generalmente mencionan primero los elementos prototípicos del nivel de base, esto es, los lexemas que se consideran ejemplares idóneos o casos centrales, con los que los restantes miembros mantienen algún grado de similitud. Estos elementos prototípicos del nivel de base mencionados en las listas inmediatamente después de la etiqueta categorial (supraordinada), son categorizados más rápidamente que los no prototípicos, ya que “son punto de referencia cognitiva y son los primeros que aprenden los niños” (Hernández Muñoz, 2006: 136). De esta forma, “las palabras más disponibles son las más prototípicas” (Sánchez-Saus, 2011: 200).

El Cuadro 4.1, correspondiente a una investigación realizada en el liceo bilingüe polaco (López González, 2014: 52), muestra los diez vocablos más disponibles en cinco centros de interés diferentes. En cada caso, estos vocablos se pueden considerar ejemplares idóneos – prototípicos– de cada categoría léxica.

**Cuadro 4.1. Vocablos más disponibles en E/LE de alumnos polacos en cinco centros de interés**

La ropa	Partes de la casa	Medios de transporte	Los animales	Profesiones
pantalón	cocina	coche	perro	profesor
camisa	dormitorio	autobús	gato	médico
zapato	habitación	tranvía	caballo	abogado
camiseta	cuarto de baño	avión	vaca	azafata
falda	salón	bici(cleta)	gallina	policía
calcetín	baño	tren	pájaro	director
jersey	ventana	metro	cerdo	conductor
bragas	pared	barco	elefante	bombero
chaqueta	puerta	caballo	oveja	enfermero
vestido	comedor	a pie	león	actor

Se convierte así el índice de disponibilidad en un indicador del grado de prototipicidad que los vocablos poseen dentro de un centro de interés, según Ávila Muñoz y Villena Ponsoda (2010). En consecuencia, cuanto más próximo sea un vocablo al prototipo del núcleo temático, mayor índice de disponibilidad mostrará, como se aprecia en la lista de léxico disponible en español del liceo bilingüe polaco (Cuadro 4.2).

**Cuadro 4.2. Vocablos de ‘Medios de transporte’ con í.d. > 0,1 empleados por el 30% de la muestra**

12. Medios de transporte			
Nº	Palabra	Disponibilidad	% Aparición
1	coche	0,84626	99,174
2	(auto)bús	0,78827	97,521
3	tranvía	0,66348	85,950
4	avión	0,61491	90,909
5	bici(cleta)	0,61363	94,215
6	tren	0,55598	79,339
7	metro	0,40932	64,463
8	barco	0,36405	70,248
9	caballo	0,29030	61,157
10	a pie	0,22561	46,281
11	moto	0,19386	35,537
12	taxi	0,14474	24,793

A medida que el índice de disponibilidad de los vocablos disminuye y, por tanto, es empleado por menor número de hablantes, los elementos léxicos se van alejando del núcleo prototípico del centro de interés (ver Cuadro 4.3). Contrastan entonces estos vocablos altamente disponibles con los que en el mismo centro de interés obtienen menor índice de disponibilidad y que están más alejados del núcleo de categorización colectiva del centro de interés.

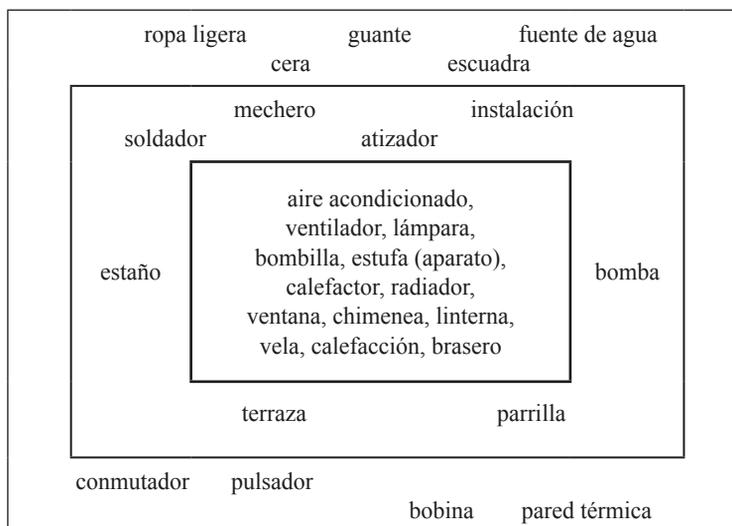
**Cuadro 4.3. Vocablos de ‘Medios de transporte’ con menor ID en la muestra polaca**

12. Medios de transporte			
Nº	Palabra	Disponibilidad	% Aparición
119	acorazado	0,00184	0,826
120	ruta	0,00184	0,826
121	trineo	0,00184	0,826
122	lento	0,00184	0,826
123	mar	0,00184	0,826
124	nave espacial	0,00167	0,826
125	tren rápido	0,00167	0,826
126	pasajero	0,00151	0,826
127	camino	0,00151	0,826
128	sendero	0,00137	0,826

Según este planteamiento, Ávila Muñoz y Villena Ponsoda (2010: 58) determinan la representación de la estructura del léxico en un centro de interés en función de los valores alcanzados por cada vocablo. Para ello establecen los estudiosos malagueños diferentes líneas de seguridad de categorización conceptual, como aparece en el Cuadro 4.4.

Sucesivamente, de dentro afuera, se distinguen tres conjuntos léxicos delimitados por tres líneas. La primera, interior, es la línea de seguridad conceptual del centro de interés, que abarca a los vocablos más próximos al núcleo prototípico del centro de interés. A medida que decrece el índice de disponibilidad, los vocablos se alejan del centro del estímulo y se acercan a una segunda línea, que estos autores llaman *línea de incertidumbre*. Este sería el límite a partir del cual el sujeto podría empezar a dudar de si la información que está aportando corresponde al estímulo propuesto. La tercera línea, exterior, es la línea de categorización conceptual del centro de interés, que marca el límite entre las palabras que guardan relación con la categoría léxica investigada y las que se hallan fuera de esta categoría.

**Cuadro 4.4. Representación de la categorización conceptual colectiva del Centro de interés 09. ‘Calefacción, iluminación y medios de airear un recinto’ (Ávila y Villena, 2010: 59)**



Continuando con este planteamiento, Ávila Muñoz y Villena Ponsoda (2010: 183–184) clasifican las unidades que aparecen en los listados según su mayor o menor compatibilidad semántica con el estímulo origen, a partir del análisis de las asociaciones léxicas producidas por los hablantes en los centros de interés. Ello supone una de las aportaciones teóricas y metodológicas más novedosas del estudio de la *Variación social del léxico disponible en la ciudad de Málaga*.

Asumiendo los principios de la teoría de los prototipos, estos autores recurren al concepto matemático de “conjunto difuso”, según el cual la relación entre un elemento y el conjunto no se define en términos tradicionales de pertenencia/ no pertenencia, sino según el grado de compatibilidad de elemento con respecto al conjunto con se mide. De este modo, los vocablos pueden agruparse en cuatro bloques según su grado de compatibilidad con el centro de interés:

1) Elementos nucleares. El grupo de los elementos nucleares está formado por las unidades que ocupan las posiciones más altas en las listas de disponibilidad y son las que más elevada accesibilidad han obtenido en las pruebas, lo que las convierte en los mejores ejemplares o ejemplares prototípicos (ej. –‘Partes del cuerpo’: *cabeza, pierna, mano...*; ‘El color’: *rojo, azul, negro...*).

2) Asociaciones denotacionales. Este segundo grupo está formado por elementos semánticamente vinculados con el prototipo correspondiente, que provienen de evocaciones o asociaciones individuales pero de carácter colectivo (ej. –‘Partes del cuerpo’: *gusto, tacto, amilasa...*; ‘El color’: *mate, luminoso, crudo...*).

3) Derivaciones. El grupo de las derivaciones está compuesto por unidades que contienen algún tipo de aumento cuantitativo o cualitativo del léxico mencionado

mediante procedimientos de derivación o de complementación (ej. –‘La cocina’: *cuchara de madera, cuchara de postre, cuchara del bebé...*; ‘El color’: *amarillento, negro mate, rojo burdeos oscuro...*).

4) Asociaciones individuales. El último grupo está formado por las asociaciones individuales, unidades léxicas que no tienen una relación directa ni colectiva con el centro de interés. Se trata del grupo de voces más alejadas de los elementos prototípicos nucleares, producto de evocaciones o asociaciones a partir de la referencia y no de la denotación (ej. –‘Partes del cuerpo’: *doctor, pendiente, ropa...*; ‘El color’: *noche, hojalata, crimen...*).

## 4.2. La asociación de palabras y las redes semánticas

### 4.2.1. La asociación de palabras en las listas de léxico disponible

La asociación de palabras en el interior de las listas de disponibilidad se corresponde con el eje asociativo secundario al que Paredes García (2012: 91) hace referencia. En este eje se vincula la nueva unidad con un término previo. Este aspecto acerca, en cierto modo, la prueba de disponibilidad léxica a la asociación libre de palabras (*free word association*). Desde el punto de vista de la Psicolingüística Experimental, el presupuesto que subyace a este procedimiento es que “la actualización sucesiva de dos palabras implica que estas dos formas se encuentran situadas en lugares próximos dentro del lexicón mental” (Hernández Muñoz, 2006: 87).

En consonancia con la *teoría de las redes semánticas* (Collins y Quilian, 1969; Collins y Loftus, 1975), que considera que las unidades de la memoria semántica están organizadas en redes asociativas, en las listas de disponibilidad existe una serie de conexiones entre las palabras que reflejan diferentes tipos de asociaciones: fonéticas, ortográficas, semántico-enciclopédicas, categoriales, personales, etc. (cf. Aitchison, 1987: 72–85).

Galoso Camacho (1997: 29–31; 2002b: 114–115), que ha estudiado los mecanismos asociativos que relacionan dos términos sucesivos en una cadena de respuestas de disponibilidad léxica, establece los siguientes tipos de asociaciones:

- 1) Por similitud de significantes, bien en la raíz de las palabras (*gaviota – gavilán*), bien por rima (*corazón – pulmón*);
- 2) Por similitud de significante y significado, en palabras derivadas (*fruta – frutero*) y en composiciones sintagmáticas (*sala de estar – sala de juegos*);
- 3) Por semejanza de significados, en sinonimia (*burro – asno; lavaplatos – lavavajillas*);

4) Por antonimia (*frío – calor; pobre – rico*);

5) Otros mecanismos entre los que destacan la hiperonimia e hiponimia (*infusión – manzanilla*), la gradación (*capitán – teniente – general*) y la sucesión espacial (*cabeza – oreja – nariz – boca; puerta – recibidor – cocina – salón*).

A continuación ejemplificamos estos tipos de asociaciones en tres listas producidas por alumnos polacos de español de nivel B2+ en el centro de interés ‘Medios de transporte’:

Lista 1. CRACOVIA. 21133 031 12 coche, avión, metro, tranvía, (auto)bús, tren, bici(cleta), a pie, furgoneta, vela, velero, barco, carro, esquí, snow-board, patera, globo, helicóptero

Lista 2. LÓDZ. 21123 001 12 coche, (auto)bús, tren, tranvía, metro, avión, barco, moto(cicleta), a pie, bici(cleta), cercanías, caballo, camello, burro, esquí

Lista 3. VARSOVIA. 11123 006 12 guagua, (auto)bús, coche, tren, ferrocarril, taxi, metro, avión, nave, carretera, autopista, mulo, caballo, bici(cleta), canal, mar, carro, camino, sendero

En estas listas, las asociaciones más frecuentes son de tipo categorial. Se observa que *coche* y *autobús* –los medios de transporte sobre ruedas más habituales en una ciudad– son palabras contiguas en las listas 2 y 3. En las listas 1 y 2, *metro* y *tranvía* –transporte sobre rieles– son vocablos contiguos, a los que se añade en la lista 2 *tren*, un vocablo de la misma subcategoría y con el mismo grupo fónico /tr/. En las listas 1 y 2, *bicicleta* y *a pie* –medios sin motor que requieren del esfuerzo humano para su movimiento– también son contiguos, y en la lista 3 *bicicleta* aparece también entre medios naturales (animales y accidentes geográficos). En la lista 1 se agrupa el transporte sobre el mar (*vela – velero – barco*), el que se realiza deslizándose sobre la nieve (*esquí – snowboard*) y los medios aéreos (*globo – helicóptero*); en la lista 2 aparecen agrupados los animales (*caballo – camello – burro*), al igual que en la lista 3 (*mulo – caballo*), en la que se agrupan también las vías asfaltadas del tráfico rodado (*carretera – autopista*) y las vías acuáticas (*canal – mar*).

En cuanto a otros mecanismos de conexión, destacan en la lista 1 las rimas internas (*furgoneta – vela; barco – carro*) y la derivación a partir del mismo significante y significado (*vela – velero*); en la lista 2 la similitud de significantes (*motocicleta – bicicleta; caballo – camello*); en la lista 3 la sinonimia (*guagua – autobús; tren – ferrocarril; camino – sendero*).

En las listas anteriores se observa la presencia de *clusters* o subcategorías semánticas, esto es, núcleos de “palabras que están más fuertemente relacionadas

y que compartan más rasgos entre sí” (Ferreira y Echeverría, 2010: 138). Estas unidades conceptuales conformadas de dos, tres, o incluso más elementos se repiten con más frecuencia de un individuo a otro, y son una prueba de la existencia de las redes semánticas.

Las categorías de tipo natural favorecen de manera especial la aparición de *clusters*. En el siguiente ejemplo de Varsovia, también de nivel C1 de español, se aprecian fácilmente las agrupaciones y los cambios de grupo en el centro *Los animales*:

VARSOVIA. 21123 013 14 [oveja, ratón, caballo, perro], [hormiga, mariposa, araña, mosca, bicho, mosquito], [gato, rata], [vaca, toro, cabra], [zorro, serpiente, cocodrilo, mono], [pájaro, ruiseñor], {elefante, [tigre, león, leopardo]}

En la enumeración se pueden identificar clara y sucesivamente: animales domésticos (*oveja, ratón, caballo, perro*), insectos (*hormiga, mariposa, araña, mosca, bicho, mosquito*), de nuevo animales del ámbito doméstico (*gato, rata*), animales de granja (*vaca, toro, cabra*), animales salvajes (*zorro, serpiente, cocodrilo, mono*), aves (*pájaro, ruiseñor*), y animales de África, comenzando por *elefante*, y con una especialización en depredadores de la sabana (*elefante, [tigre, león, leopardo]*).

En otras ocasiones los cambios de grupo no son excluyentes, sino que se van superponiendo y transitando de una manera gradual y dinámica de un grupo a otro, como en el siguiente informante también de Varsovia:

VARSOVIA. 21133 014 14 [perro, gato], [vaca, caballo, {pato, pollo, pavo}, cigüeña, ave, pájaro, águila, golondrina}, [mosca, abeja, avispa], ratón, [elefante, jirafa, toro, cebra]

En este caso se observa una transición suave de animales domésticos (*perro, gato*) a de granja (*vaca, caballo, pato, pollo, pavo*), los tres últimos elementos de este grupo junto con *cigüeña, ave, pájaro, águila, golondrina* conforman el grupo de las aves, después aparecen los insectos (*mosca, abeja, avispa*) y tras, el vocablo *ratón*, que sirve de transición, aparece un grupo de animales salvajes (*elefante, jirafa, toro, cebra*).

En esta asociación de palabras juega un papel importante el *priming* (memoria implícita), o facilitación semántica y formal, que se da de manera automática cuando dos conceptos están relacionados semánticamente. Hernández Muñoz (2006) ha estudiado las manifestaciones del *priming* en el léxico disponible. Esta autora explica del siguiente modo la naturaleza de este fenómeno:

La aparición de un término favorece la activación, selección y producción de otras formas semánticamente relacionadas en un variado conjunto de tareas psicolingüísticas. [...] El efecto que tiene la palabra/estímulo que se presenta primero (*prime*) en la palabra objetivo (*target*), facilitando su producción, puede estar basado en los diversos tipos de relación que se establece entre ellas: semántica, formal, morfológica, de traducción entre lenguas o incluso de repetición de la misma palabra (Hernández Muñoz, 2006: 208).

A continuación, Hernández Muñoz lista distintos indicios que señalan que el fenómeno de *priming* está activo en la evocación del léxico disponible a diversos niveles. Entre ellos se cuenta:

1. Cuando la etiqueta categorial la forman dos términos (ej. –‘Comidas y bebidas’), la mayor parte de los sujetos comienza respondiendo con elementos del segundo:

VARSOVIA. 21123 001 05 cerveza, cava, vino, vodka, zumo, agua, plátano, manzana, tomate, maíz, hamburguesa, chorizo, carne, sopa, cebolla, pepino, macarrón, salsa, ajo, especia, pastel, tarta, yogur, leche, melón, cacahuete, avellana, nuez, patata frita, patata brava, lechuga

2. Cuando aparece un término en los últimos lugares de un centro de interés, puede aparecer en los primeros del siguiente. Es más probable que esto ocurra en centros consecutivos relacionados, como ya notó Galloso Camacho (2002b: 115) entre 06 ‘Objetos sobre la mesa para la comida’ y 07 ‘La cocina y sus utensilios’.

3. Cuando un término o un conjunto de términos han sido evocados en un centro de interés, tienden a volver a evocarse en los sucesivos de manera sistemática. Para Hernández Muñoz (2006: 210) esta es la causa del léxico común de dos centros de interés consecutivos, aunque en uno de ellos los términos compartidos sean periféricos. Ello explica la alta disponibilidad de los términos *cuchillo*, *tenedor*, *cuchara* en ‘Objetos colocados encima de la mesa para la comida’ y ‘La cocina y sus utensilios’, o de la aparición de vocablos como *casa*, *coche* en ‘El campo’, categoría que aparece a continuación de ‘La ciudad’.

4. Un elemento de un subgrupo dentro de la categoría favorece la aparición de otro elemento de esa subcategoría con lo que se responde en orden no a la disponibilidad de las palabras del conjunto total, sino según la disponibilidad de los términos en cada uno de los diferentes subgrupos. De este modo, el *priming* es la causa de las agrupaciones de palabras anteriormente mencionadas.

LÓDZ. 21123 001 15 [bingo, fútbol, béisbol, voleibol], [cartas, naipes], [cine, teatro, museo], discoteca, galería, baile, [natación, correr], [quedar con (los) amigos, salir de fiesta], viajar, [ordenador, pelí(cula), tele(visión)], escoba

En el ejemplo anterior, se manifiesta, además, la manera en la que el *priming* añade fuerza a las conexiones indirectas, es decir, entre palabras que no se producen

de manera sucesiva, sino que están separadas entre sí por otros términos (Nelson *et al.*, 1997). Este es el caso de *museo* y *galería*, que comparten la idea de ‘arte, exposición, pintura’, y de *discoteca* y *baile*, unidas por la ‘música para bailar’.

#### 4.2.2. La representación gráfica de la red semántica

Como hemos expuesto en las páginas anteriores, diversos autores advierten en sus corpórea ciertas agrupaciones categoriales o conjuntos asociativos. De hecho, muchos de ellos afirman que los vocablos disponibles se organizan en forma de redes semánticas<sup>4</sup> (propias de un paradigma conexionista) como, por ejemplo, López Morales (1999a), Urrutia Martínez (2001), Galloso Camacho (2002b), Gómez Devis (2004).

A la hora de enfrentarse al problema de la representación del conocimiento, el conexionismo –un tipo especial de asociacionismo en la ciencia cognitiva– se basa en las asociaciones entre diferentes tipos de información modeladas en redes neuronales artificiales. En este contexto, de acuerdo con Echeverría *et al.* (2008: 83), la red semántica es la mejor solución que tiene ante sí el psicolingüista para caracterizar las relaciones que se establecen entre los elementos del léxico disponible, y así poder representar las unidades y al mismo tiempo dar cuenta de sus interrelaciones.

La teoría de las redes semánticas (*semantic networks*), desarrollada por Collins y Quilian (1969) y Collins y Loftus (1975), considera que las unidades de la memoria semántica –idealizaciones abstractas a las que llama nodos (*nodes*)– están organizadas en redes asociativas. “En estas redes, constituidas por conjuntos de nodos unidos por conexiones (*links*), la activación de la información semántica se propaga jerárquicamente, dependiendo de la fuerza intrínseca de cada una de las conexiones” (Hernández Muñoz, 2006: 66).

En lo tocante a la representación gráfica del conocimiento, Lehmann (1992) ha definido este constructo de la siguiente forma:

Una *red semántica* representa el conocimiento como un grafo en forma de red. Una idea, un evento, una situación u objeto tiene casi siempre una estructura compuesta; esto se representa en una red semántica mediante una correspondiente estructura de *nodos* que representan *unidades conceptuales*, y *aristas* direccionadas que representan las *relaciones* entre las unidades (Lehmann, 1992: 2).

A partir de este constructo teórico, Echeverría *et al.* (2008) crearon *Dispo-Grafo*, un programa informático que permite calcular la frecuencia con que dos términos aparecen unidos en los listados y representa mediante esquemas esas relaciones. Por consiguiente, *Dispografo* posibilita el análisis cuantitativo de las

---

<sup>4</sup> Recordemos que Manjón-Cabeza (2008a) afirmaba que tanto los prototipos como las redes semánticas participan en la organización del lexicón mental, y que, por tanto, no son incompatibles.

asociaciones que se establecen entre las unidades con objeto de determinar las asociaciones estables en el conjunto de un colectivo de encuestas. Los esquemas gráficos resultantes están formados por las palabras (nodos) unidas entre sí por líneas (aristas) sobre las que se indica el número de veces que aparecen relacionadas, como se puede ver en la Figura 4.1, tomado de Echeverría *et al.* (2008: 88). En dicho grafo *podá 1* significa que el grafo original ha sido podado de las aristas con peso 1 y eliminados los nodos huérfanos (sin aristas).



Figura 4.1. Centro de interés 'Alimentos' (23 sujetos, poda 1\*), Chile (Echeverría *et al.*, 2008: 88)

\* *Podá 1* significa que el grafo original ha sido podado de las aristas con peso 1 y eliminados los nodos huérfanos (sin aristas)

Sobre el grafo se pueden realizar observaciones cualitativas como la *categorización*, a través de las agrupaciones nodales (*clusters*) claramente visibles en el

grafo; la configuración de *núcleos* y *vecinos* inmediatos; la *fuerza de la relación* establecida por el peso de las aristas, etc. En el ejemplo, la estructura general muestra categorías fácilmente reconocibles tales como ‘frutas’ (superior izquierda), ‘frutas tropicales’ (superior derecha): *pomelo*, *guayaba*, *mango*. Como subconjuntos aparte aparecen *salmón*, *atún* (superior central), y *queque*, *torta* (medio derecha). Alrededor de *lechuga* y *tomate* se agrupan las ‘hortalizas’ (inferior izquierdo) (Echeverría *et al.*, 2008: 88).

Los grafos de un conjunto muy numeroso tienen una gran complejidad nodal. Para apreciar más claramente las relaciones semánticas de las palabras más disponibles, se puede prescindir de aquellos elementos cuyas aristas (conexiones) tengan un peso muy bajo en el conjunto. Así, si en el grafo anterior se podan las aristas con peso 2, se perciben aquellas relaciones entre elementos que han mostrado máxima estabilidad en el conjunto de las respuestas de los sujetos. Ello facilita igualmente el estudio de las agrupaciones más importantes, como demuestra el grafo de la Figura 4.2.

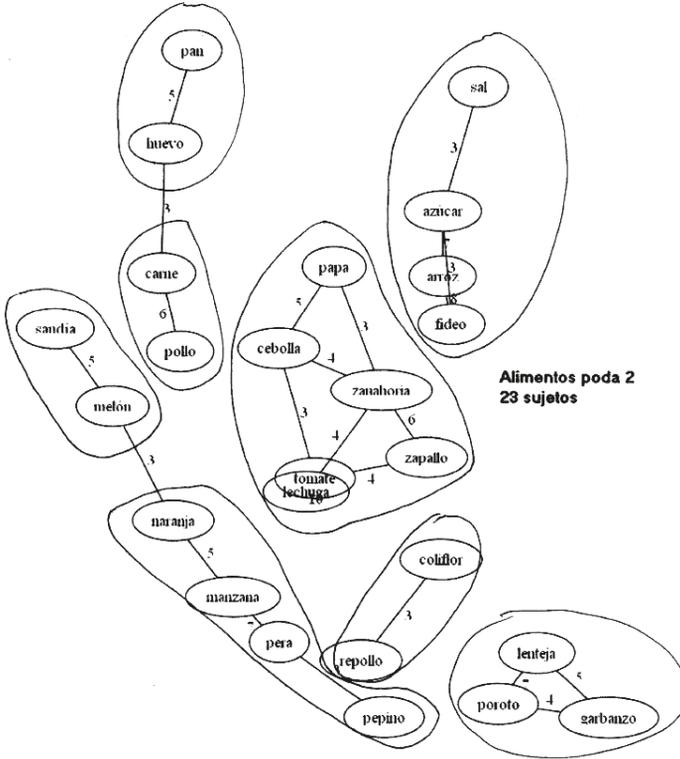


Figura 4.2. Centro de interés ‘Alimentos’ (23 sujetos, poda 2), Chile (Karelovic Burotto, 2010: 111)

Gracias a esta aproximación se puede visualizar mejor la categoría ‘hortalizas’ (en posición central) con *papa*, *cebolla*, *zanahoria*, *zapallo*, *tomate* y *lechuga*, con núcleo en estas dos últimas. A la derecha abajo aparecen claramente las ‘legumbres’, y sobre estas, a la izquierda se observa el par autónomo de *coliflor* y *repollo*. Otras agrupaciones delimitadas son: a la izquierda, el par *sandía* – *melón*, ambas frutas grandes con pulpa aguanosa de la familia de las cucurbitáceas; más abajo a la izquierda la conexión *naranja* – *manzana* – *pera* (frutas de tamaño parecido que suelen adornar las fuentes y canastos sobre la mesa), a las que se une *pepino* por similitud de significantes en el lexema; arriba a la derecha los condimentos *sal* y *azúcar*, unidos a *arroz* y *fideo*; el par *carne* – *pollo* fuertemente relacionado; además de *pan* y *huevo*.

### 4.2.3. Las relaciones significativas

Como se desprende del análisis anterior, la naturaleza de las relaciones entre los nodos puede ser semántica, fonológica o morfológica. A propósito de esta cuestión, Sánchez-Saus (2011) ha indagado en las asociaciones que se producen dentro de los centros de interés a la luz de los presupuestos de la semántica estructuralista y llega a la conclusión de que

la configuración de los mismos parte de la capacidad que las palabras tienen para asociarse con otras, no intralingüísticamente, es decir, no por las propiedades que tienen estos elementos dentro de la lengua, entendida como estructura –los centros de interés no son campos léxicos estructurados –, sino a partir de las asociaciones mentales que se producen en la mente de los informantes cuando se les propone un centro como estímulo Sánchez-Saus (2011: 323).

Esto lleva a la autora a entender los centros de interés como *configuraciones asociativas*. Según Coseriu (1977: 169), las *configuraciones asociativas* son “asociaciones de un signo con otros signos, asociaciones establecidas por similitud o contigüidad, tanto de los significantes como de los significados”, fundadas “también en asociaciones relativas a las *cosas* designadas, no a las unidades lingüísticas como tales”. Por lo tanto, la base de asociación no está en el nivel del significado, sino, fundamentalmente, en el de la designación, es decir, en una referencia a lo extralingüístico.

A la hora de determinar el carácter de las relaciones asociativas que ocurren en los centros de interés, Sánchez-Saus (2011: 326–341) distingue entre las que son de naturaleza lingüística y las de naturaleza extralingüística. Dentro de las de carácter lingüístico, tienen un papel activo en la composición de las listas de léxico disponible las relaciones léxicas, ya sean auténticas relaciones semánticas

desde el significado (sinonimia, parasonimia, antonimia, hiperonimia, hiponimia), ya sean pseudorrelaciones semánticas a partir del significante (polisemia y homonimia)<sup>5</sup>. Las relaciones extralingüísticas, denominadas “significativas”, se establecen mediante cualquier tipo de vínculo basado en el contenido (metáfora, metonimia, etimología popular, relaciones de contigüidad o meronimia, terminologías, asociaciones de carácter cultural o experiencial, etc.).

Para dar cuenta de la pluralidad de relaciones que se dan en los centros de interés, Sánchez-Saus (2011) propone utilizar el concepto de *relaciones “significativas”*, que Casas Gómez (2011: 88; 2005: 9–10) define como “todo ese conjunto de valores sustanciales del contenido que tienen que ver con las relaciones entre los componentes del signo: el significante, el significado/concepto y la referencia”. A continuación la autora aclara que estas relaciones pueden ser tanto lingüísticas como no lingüísticas:

estas relaciones “significativas”, cuando se establecen entre significantes o entre significantes y significados, se considerarán *lingüísticas* y serán la base de fenómenos como la paronimia, la etimología popular y de los campos semasiológicos o campos de significaciones y campos onomasiológicos o campos de designaciones; cuando, sin embargo, las relaciones se establezcan entre conceptos y realidades o cosas, tales relaciones “significativas” tendrán un carácter no lingüístico y explicarán las conexiones que se producen a partir de aspectos sustanciales de naturaleza *designativa* (relaciones asociativas, estilísticas, culturales o terminológicas, entre otras) (Sánchez-Saus, 2011: 327–328).

Sánchez-Saus (2011: 342–376) ha usado el programa *Dispografo* para estudiar el tipo de relaciones “significativas” y los vínculos entre palabras sobre los que están organizados los centros de interés, a partir de una muestra de 322 alumnos extranjeros de la universidades de Andalucía. A continuación mostramos dos ejemplos de las relaciones significativas establecidas en el interior de distintos tipos de centros de interés, a partir de una investigación que hemos realizado con una muestra de 121 alumnos polacos de Liceo de nivel B2+ (López González, 2014).

– ‘Partes del cuerpo humano’.

Sea el grafo de la Figura 4.3, donde se han eliminado las aristas con conexiones iguales o inferiores a 3 y los nodos huérfanos (sin aristas). En este, se comprueba que el centro de interés nº 1, ‘Partes del cuerpo humano’ es una categoría, en su estructura interna, natural (con elementos constituyentes fácilmente identificables) y bien definida (una lista cerrada de elementos similares), e inclusiva (todos sus elementos tienen una relación de pertenencia estricta). Por tanto, este centro está compuesto de elementos léxicos que, en su gran mayoría, pueden con-

---

<sup>5</sup> Como vimos anteriormente (cf. 4.2.1), este tipo de asociaciones en disponibilidad léxica ya había sido estudiado por Galloso Camacho (1997: 29–31; 2002b: 114–115).

siderarse cohipónimos, correspondiéndose su hiperónimo con la denominación del centro de interés. Ahora bien, aunque el vínculo de muchos de los elementos que componen el centro con la denominación del mismo sea una relación léxica de hiperonimia – hiponimia, las conexiones que más frecuentemente se producen dentro de este centro, entre estos elementos, son, como puede observarse, de tipo “significativo” designativo, fundamentalmente por contigüidad: así, *pierna-pie*, *mano-brazo-mano*, *mano-dedo-uña*, *boca-labio*, *labio-diente*.

En este centro se muestran también relaciones “significativas” lingüísticas entre significantes de tipo paronímico: ésta puede ser una causa, junto con la condición de cohipónimos y la cercanía referencial, de la alta frecuencia con la que aparecen unidos *ojo* y *oreja*, en ambas direcciones, al igual que el caso de *pie-pierna*.

Dentro de estas relaciones designativas es evidente que los informantes han creado un “subcentro”, el de la cabeza, que se observa muy claramente en la representación gráfica que nos facilita *DispoGrafo*, en cuyo interior los elementos funcionan como merónimos de *cabeza*, esto es, como partes de un todo. *Ojo*, *oreja*, *nariz* aparecen en el centro del gráfico muy estrechamente relacionados en una relación de órganos de los sentidos. A éstos también se unen *labio*, *boca* y, en otra dirección, *pelo*. A partir de *boca* y *labio* se lleva a cabo la asociación, también frecuente con *diente* y *lengua*, y, en una relación con menos conexiones, *pestaña*, *mejilla*, *cara*. Se trata, en todos los casos, de relaciones “significativas” designativas por contigüidad de los referentes.

Un segundo núcleo relacional también muy fuerte es el de las extremidades, que aparece en el gráfico bajo los vocablos designativos de los órganos de los sentidos. Se presenta la relación que une *mano-pierna* –así como *pierna-brazo*–, y a ambas designaciones con *dedo* como término común. Se trata de una relación de cohiponimia, en tanto que partes del cuerpo, pero también son elementos contrarios, “antónimos” del hablar, asociados polarmente a partir del conocimiento que se tiene de la realidad: extremidad superior frente a extremidad inferior. A partir de ellos, y de nuevo por meronimia, aparecen *mano-dedo-uña*, y las conexiones *dedo-brazo* y *dedo-codo*.

En el conjunto léxico, las relaciones por contigüidad de referentes y basadas en la ordenación visual son frecuentes. De este modo se asocian: *espalda-brazo* como unión de las extremidades superiores al tronco, *rodilla-codo* como articulaciones intermedias, *cabeza-cuello-pecho* en una ordenación de arriba abajo uniendo la cabeza al tronco, *culo-pecho* con evidente relación erótica. Además en la parte inferior, el gráfico muestra un itinerario relacional entre el vientre (*barriga-estómago*) y los órganos (*corazón-pulmón*), que en un nivel inferior se amplía a *riñón*.

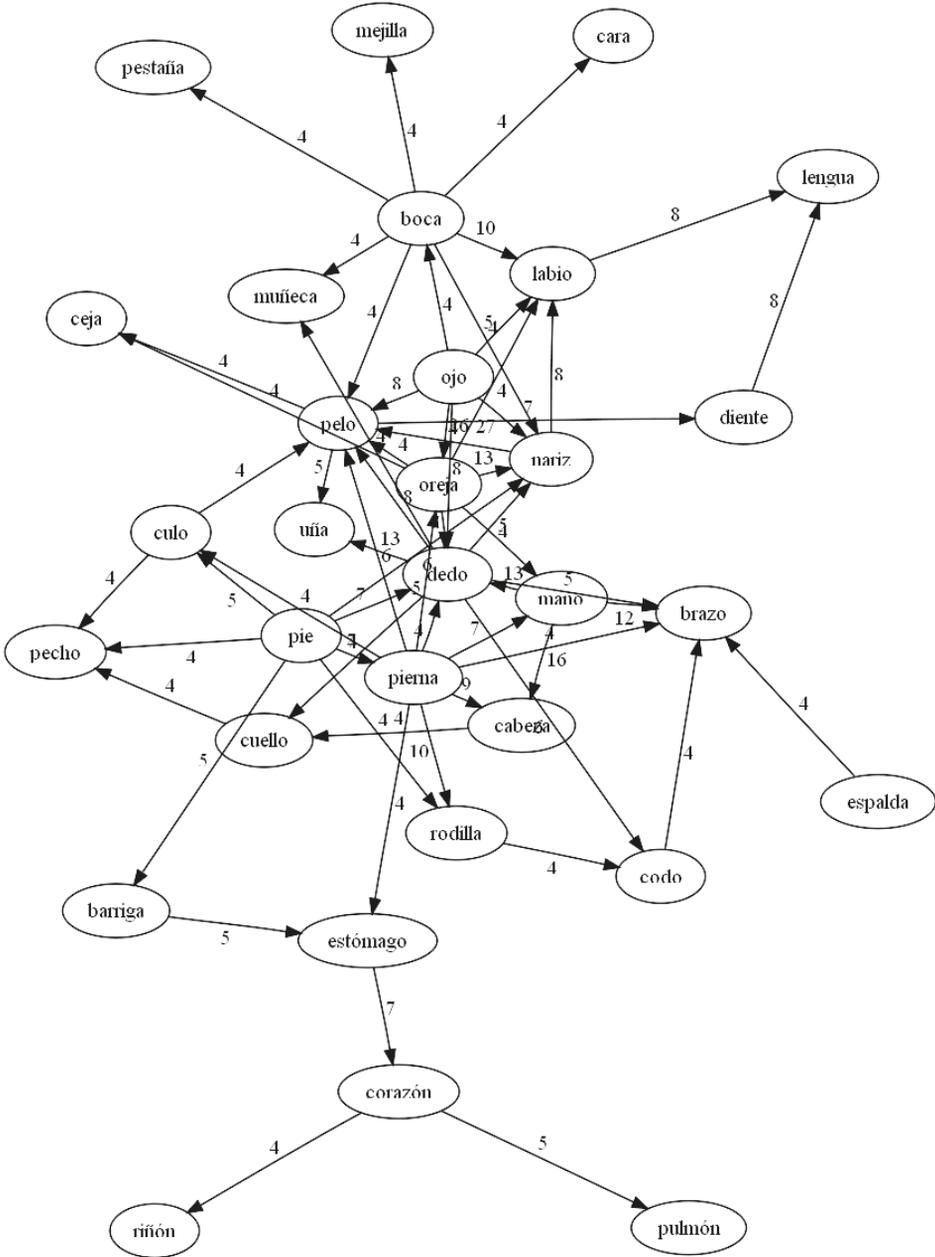


Figura 4.3. ‘Relaciones semánticas en el centro 01. ‘El cuerpo humano’ (poda 3)\*

\* Eliminadas las aristas con conexiones iguales o inferiores a 3 y los nodos huérfanos (sin aristas)

– ‘La ciudad’.

El centro nº 10 ‘La ciudad’ representa un tipo de configuración diferente, como se observa en el grafo de la Figura 4.4. Se trata de una categoría radial (con múltiples subcategorías en su interior) y con forma de esquema (que representa idealmente cómo es una ciudad) en su estructura interna. En cuanto a los niveles de inclusividad que se dan, es un campo relacional, en el que las palabras que forman el centro de interés ‘La ciudad’ se relacionan con el enunciado del centro en tanto que *cosas* que están en la realidad a la que este se refiere. De este modo, se forman configuraciones más restringidas en su interior.

Así, por un lado, encontramos el campo léxico de los medios de transporte, en el que la palabra *tranvía* aparece como la que vertebra la ciudad polaca y, por tanto, el conjunto léxico resultante. El campo léxico medios de transporte muestra la relación fuerte *tranvía-autobús*, y de manera más débil los medios que se mueven sobre raíles *tranvía-tren* y *tranvía-metro*. Por contigüidad designativa se relacionan los medios de transporte que usan ruedas: *coche* se une a *autobús* y *bicicleta*, y, desde *coche*, *calle*, *semáforo*, *gente*, *contaminación*, igualmente por contigüidad designativa. Desde *calle* sale, por un lado, otra vía, como *avenida*, con la que forma un campo léxico distinto, el de los tipos de vías públicas, así como otros espacios urbanos, como *parque* y *plaza*, y dos subconjuntos con ciertos elementos que se encuentran en la calle, como *edificio*, *banco* y ayuntamiento, y por otro lado el campo léxico de la vivienda (*casa-piso-bloque*). De otra parte, a través de otra relación designativa por contigüidad, se relaciona *gente* con diferentes espacios de entorno público: *escuela-universidad* en el ámbito educativo, que con *hospital* forma parte del sector de servicios públicos, y *teatro-cine-museo* pertenecientes al mundo del ocio y la cultura.

De forma independiente se presentan los miembros del campo de los establecimientos de restauración-diversión: *bar-restaurante-cafetería-discoteca*. También como subconjuntos independientes aparecen: los animales que viven en la ciudad (*perro-ciudad*), el par *tráfico-atasco* con evidente relación causal, y el campo léxico del comercio (*tienda-supermercado-mercado-almacén*). Por lo tanto, como ya notó Sánchez-Saus (2011: 360) en un estudio similar anterior, “las relaciones de los elementos de un campo léxico con su archilexema correspondiente son [...] de hiperonimia-hiponimia, siendo cohipónimos entre sí los miembros del campo”.



## V. LOS CONDICIONANTES EXTRALINGÜÍSTICOS DE LA DISPONIBILIDAD LÉXICA

Subrayan Samper Padilla y Samper Hernández (2006: 45) que “una de las aportaciones más relevantes de los estudios panhispánicos [de disponibilidad léxica] es la consideración de la influencia de los factores sociales en los índices de disponibilidad”. Para el valor de estos índices tan importante es la respuesta subjetiva e inherentemente propia de cada individuo como la pertenencia por razones sociales y culturales a una misma comunidad de habla. Ambos papeles de la persona –el de individuo con su propia experiencia y el de miembro de una comunidad– inciden en la estructura léxica de los centros de interés a un tiempo (Ávila Muñoz y Sánchez Sáez, 2011: 71).

La sociolingüística se ha constituido en el pilar básico de los estudios de disponibilidad. Efectivamente, la perspectiva sociolingüística ha sido adoptada en la mayor parte de los trabajos en este campo, hasta tal punto que la metodología del Proyecto Panhispánico incluye de manera obligatoria la estratificación y análisis de la muestra de acuerdo con las variables sociales ‘sexo’, ‘tipo de centro’, ‘ubicación del centro’ y ‘nivel sociocultural’. Aparte de estos factores sociales, otras investigaciones han integrado en su malla sociolingüística el estudio de las variables ‘edad’, ‘contacto lingüístico’, ‘lengua materna’, ‘modelo de enseñanza’, ‘zona geográfica’, etc.

Para la sociolingüística, la disponibilidad es “un instrumento que le permite establecer estratificaciones lingüísticas de las comunidades de habla” en función de diferentes factores sociales, con las que “determina la caracterización léxica de los sociolectos que integran la comunidad en cuestión” (López Morales, 1996: 252). Para el mismo autor, la principal ventaja de realizar análisis sociolingüísticos por medio de la disponibilidad léxica es la ausencia de problemas teóricos de relieve, ya que se trabaja principalmente con cuantificación y no con variación cualitativa, al contrario de lo que sucede al estudiar la auténtica variación léxica. Se evitan así las dificultades semánticas a la hora de establecer sinónimos, y metodológicas, relacionadas con “la intención comunicativa del hablante y su reacción en el oyente y la contaminación subjetiva que puede darse en el proceso de interpretación de los datos” (López Morales, 1989: 105).

Las dos variables que han tenido más rendimiento en los estudios de disponibilidad léxica son el ‘sexo’ y el ‘nivel sociocultural’. A ellas prestaremos una mayor atención en las próximas líneas. Completaremos el capítulo con un repaso a los resultados más destacados en relación con el resto de variables sociolingüísticas.

## 5.1. El sexo

La influencia del factor ‘sexo’ ha sido analizada con asiduidad en la disponibilidad léxica. Sin embargo, a pesar de esa presencia constante, en el Proyecto Panhispánico no se preestratifica este factor, sino que se sigue un procedimiento aleatorio de selección. En la práctica esto significa que, una vez seleccionados los centros de enseñanza, se pasa la encuesta a los grupos de alumnos que están disponibles en un momento determinado.

La consecuencia de este proceder es que, “todas las muestras, sin excepción, presentan una proporción mayor de mujeres que de hombres” (Samper Padilla y Samper Hernández, 2006: 46). Esta mayor presencia femenina es un reflejo de la realidad de las aulas españolas, en las que en el curso previo a la entrada a la universidad el número de alumnas supera al de alumnos, hecho presente en las estadísticas oficiales del Ministerio de Educación y Ciencia<sup>1</sup>.

En Samper Padilla *et al.* (2003) se señala que los resultados referidos a la variable ‘sexo’ ponen de manifiesto una gran homogeneidad general en las respuestas aportadas por hombres y mujeres. En efecto, diferentes investigaciones (Alba, 1995a; Benítez Pérez, 1994c; Mateo García, 1998; Carcedo González, 2001; Prado Aragonés y Galloso Camacho, 2005; Ávila Muñoz, 2006) en el ámbito hispánico concluyen que el análisis del factor ‘sexo’ no revela, por lo general, diferencias apreciables y significativas entre hombres y mujeres, lo cual vendría a coincidir con las conclusiones de los trabajos pioneros en francés (Gougenheim *et al.*, 1956, 1964) e inglés (Dimitrijevic, 1969).

Sin embargo, tres investigaciones (Galloso Camacho, 2002a; Gómez Devís, 2004; Hernández Muñoz, 2006), que han utilizado medios estadísticos, han encontrado que las diferencias cuantitativas entre la producción léxica del sociolecto masculino y el femenino sí son significativas. Así, Galloso Camacho (2002a) para la provincia de Zamora comprueba la significación estadística de las diferencias de producción léxica entre sexos por medio de un T-TEST. El resultado pone de manifiesto que la diferencia entre los promedios de respuestas de los hombres y de las mujeres es estadísticamente significativa ( $\rho = 0,006$ ) a favor del colectivo

---

<sup>1</sup> En el documento *Datos y cifras del curso escolar*, publicado anualmente por la Secretaría General Técnica del Ministerio de Educación y Ciencia, se puede comprobar que los porcentajes correspondientes a los alumnos matriculados en Bachillerato se sitúan, desde el año 2000 hasta el año 2013, entre el 45–47% para los hombres y el 53–55% para las mujeres.

masculino. Idéntica prueba estadística emplea Hernández Muñoz en el estudio de Castilla-La Mancha, en el que la diferencia también es estadísticamente significativa ( $p < 0,01$ ) también a favor del grupo masculino. Por su parte, Gómez Devís (2004), en Valencia, utiliza un análisis de la varianza simple (ANOVA), que le sirve para encontrar once diferencias significativas en los centros de interés. Los tres casos anteriores llevan a Hernández Muñoz (2006: 325) a replantearse la afirmación clásica de que el factor ‘sexo’ no tiene una influencia destacable en la conformación cuantitativa del léxico disponible.

**Cuadro 5.1. Promedio de palabras por informante, según sexo, en investigaciones de España**

Comunidad	Investigación	Hombres	Mujeres
Alicante	Martínez Olmos (2007)	18,74	19,53
Almería	Mateo García (1998)	20,47	20,45
Aragón	Lagüens Gracia (2008)	20,82	21,37
Asturias	Carcedo González (2001)	18,90	19,20
Cádiz	González Martínez (1997, 2002)	18,37	18,75
Castilla-La Mancha	Hernández Muñoz (2006)	19,41	18,63
Ceuta	Ayora Esteban (2003, 2006)	19,86	20,15
Chile	Valencia y Echeverría (1999a)	17,04	17,43
Córdoba	Bellón Fernández (2003, 2011)	19,97	19,89
Galicia	López Meirama (2008, 2011)	20,99	21,49
Gran Canaria	Samper Padilla y Hernández Cabrera (1995b)	20,51	20,95
Granada	Pastor Milán y Sánchez García (2008)	17,28	18,6
Huelva	Prado Aragonés y Galloso Camacho (2005)	20,50	20,00
Lérida	Serrano Zapata (2002)	17,91	19,43
Melilla	Fernández Smith <i>et al.</i> (2008)	17,30	16,78
Salamanca	Galloso Camacho (2003a)	20,94	19,94
Segovia	Santiago Guervós (2008)	20,08	20,35
Sevilla	Trigo Ibáñez (2007)	18,96	19,72
Soria	Bartol Hernández (2004a)	20,25	22,17
Valencia	Gómez Devís (2004)	19,56	20,78

Por lo demás, como se aprecia en el Cuadro 5.1, los resultados de las investigaciones apuntan más bien hacia una mayor riqueza léxica del grupo femenino. Y es que, aunque hay zonas en las que las diferencias no son apreciables y algunas pocas en las que se registra una superioridad numérica en el promedio de respuestas de los hombres (Puerto Rico, Almería, Castilla-La Mancha, Salamanca, Zamora), son muchas más (Chile, República Dominicana; Alcalá de Henares, Alicante, Aragón, Asturias, Ávila, Cádiz, Ceuta, Córdoba, Galicia, Granada, Gran Canaria,

Lérida, Sevilla, Soria, Valencia, etc.) aquellas en las que hay una ligera superioridad léxica en el promedio de respuestas de las mujeres (Samper Padilla y Samper Hernández, 2006: 47; Samper Padilla, 2006: 111).

Entrando en detalle en los resultados de cada centro de interés, algunos centros son más favorables para unos hablantes que para otros, dependiendo de su sexo. Así, hasta en dieciocho investigaciones diferentes<sup>2</sup> se detecta una diferenciación temática entre centros de interés masculinos y femeninos. De este modo, las chicas aventajan significativamente a los chicos en los centros ‘La ropa’, ‘Los muebles de la casa’, ‘La cocina’ y ‘Los colores’<sup>3</sup>. Otros centros destacados entre las informantes femeninas han sido ‘Comidas y bebidas’ y ‘Objetos sobre la mesa para la comida’. Por el contrario, las investigaciones de siete comunidades de habla<sup>4</sup> definen un grupo temático con superioridad masculina en ‘Calefacción e iluminación’, ‘Medios de transporte’ y ‘Trabajos del campo y del jardín’.

La diferenciación por sexo también se manifiesta en aspectos cualitativos comunes a las investigaciones que han abordado esta cuestión: Aragón (Lagüéns Gracia, 2008), Cuenca (Hernández Muñoz, 2004), Valencia (Gómez Devís, 2004) y Gran Canaria (Samper Hernández, 2003a). Estas características comunes a estas investigaciones son:

– El alto índice de vocablos compartidos entre ambos sexos, cercano al 50% (por ejemplo, un 46% en el total de vocablos de Aragón).

– Las mayores divergencias se presentan referidas a ‘Juegos y distracciones’, con mayor riqueza léxica en deportes, aventura, lucha y aparatos tecnológicos (*fútbolín, videojuego, consola (de videojuegos), puenting* en Aragón; *squash, motociclismo, consola, waterpolo, fútbol sala, paddle* en Valencia) para los chicos, frente a juegos de patio, juegos de mesa e interior, y actividades culturales para las chicas (*comba, escondite, goma (elástica), muñeco* en Aragón; *cuerda, patinar, aerobic, goma, gimnasia rítmica* en Valencia).

– Un mayor uso de tabúes entre los chicos frente a un vocabulario más cuidado de las mujeres. Por ejemplo, en Sevilla, Trigo Ibáñez y González Martínez (2011: 34) encuentran en el centro de interés ‘Partes del cuerpo humano’ entre los chicos once coloquialismos referidos a los genitales (*nabo, cojón, porra, pi-*

---

<sup>2</sup> Chile (Echeverría *et al.*, 1987; Valencia y Echeverría, 1999a), Puerto Rico (Butrón, 1989), Madrid (Benítez Pérez, 1994c), Gran Canaria (Samper Padilla y Hernández Cabrera, 1997; Samper Hernández, 2003a), Cádiz (González Martínez, 1997), Salamanca (Galoso Camacho, 2002b), Ceuta (Ayora Esteban, 2003), Córdoba (Bellón Fernández, 2003), Lérida (Serrano Zapata, 2004), Soria (Bartol Hernández, 2004), Valencia (Gómez Molina y Gómez Devís, 2004), Cuenca (Hernández Muñoz, 2004), Alicante (Martínez Olmos, 2007), Aragón (Lagüéns Gracia, 2008), Galicia (López Meirama, 2008), Sevilla (Trigo Ibáñez, 2011).

<sup>3</sup> En las comunidades en las que se ha contemplado en este último ámbito temático.

<sup>4</sup> Gran Canaria (Samper Padilla y Hernández Cabrera, 1997), Cádiz (González Martínez, 1997), Córdoba (Bellón Fernández, 2003), Lérida (Serrano Zapata, 2004), Soria (Bartol Hernández, 2004), Cuenca (Hernández Muñoz, 2004), Aragón (Lagüéns Gracia, 2008).

*cha, pirindola, gallitrompa, pito, tercera pata, pilila, manubrio y paquete*) frente a sólo tres (*pepe, coño, chochete*) entre el grupo femenino. Una mayor aceptación del tabú entre los informantes masculinos también es constatada por Lagüens Gracia (2008) en Aragón.

– Las mujeres presentan una mayor riqueza en expresiones de tipo metonímico referidas a los colores. Así, en Aragón, Lagüens Gracia (2008: 130) registra el uso del sustantivo con valor metonímico, especialmente abundante entre las mujeres: *champán y miel, berenjena, bronce, piel, platino, nácar, azafrán, borgoña, cereza y grosella, agua, arena, avellana, albaricoque, barniz, brillantina, caldero, camello, caramelo, chocolate, ciruela, coral, ébano, gazpacho, grava, hielo, humo, jade, lavanda, leche, mármol, nata, olivo, paja, piedra, pizarra, tabaco, terracota o yema*; y, además, *amarillo canario, verde moco, azul azafata, marrón tierra, rojo amapola, verde bosque, verde agua, verde lima y verde pino, amarillo pollito, azul océano, blanco lápida, gris topo, rojo tierra, rojo tomate, rojo sandía, rosa chicle, verde coliflor, verde hospital, verde limón, verde río, verde piscina, verde sapo o verde vidrio*. Al respecto, este autor advierte que entre ambos grupos, “se impone una diferencia no solo cuantitativa, sino también semántica: [ya que] el abanico de colores matizados por las mujeres cubre prácticamente todo el espectro cromático”.

Estas diferencias en los repertorios léxicos de ambos sexos parecen deberse a patrones psicosociales y roles sociales definidos por la comunidad, ya que, como hemos visto, con frecuencia “ciertos campos léxicos relacionados con el papel que ha desempeñado tradicionalmente la mujer obtienen un número de respuestas más elevado entre las alumnas que entre sus compañeros de aula” (Samper Padilla y Samper Hernández, 2006: 51), lo que confirma las apreciaciones sobre el sexismo léxico indicadas por García Moutón (1999: 70–71)<sup>5</sup>. Ante este hecho, Samper Padilla (2006: 112) se sorprende de “que en la población juvenil actual se sigan manteniendo esas diferencias cuantitativas en centros de interés asociados con los papeles masculino y femenino”. Para este autor, existe la posibilidad de “una correlación del factor ‘sexo’ con otros de los condicionantes contemplados en las investigaciones (por ejemplo, el factor urbano/rural” que pueda influir en las diferencias de producción léxica entre ambos sexos. Al respecto, Gómez Molina (2006) concluye que el distinto comportamiento léxico de hombres y mujeres está motivado tanto por causas sociolingüísticas (biológicas y culturales) como psicolingüísticas (categorización del entorno, memoria semántica, fluencia semántica...).

---

<sup>5</sup> García Moutón (1999: 70) se refiere a la frecuente especialización que sufre el léxico en campos de interés “que han sido tradicionalmente femeninos” (como los de los colores y las telas) del siguiente modo: “Hasta hace poco se habría considerado poco masculino a un hombre que empleara en una conversación normal *azul cielo, verde hoja o rosa palo*, porque la expresión de determinados matices y sentimientos sólo estaba permitida a las mujeres, hasta el punto de que algunas palabras se especializan semánticamente por sexos”.

## 5.2. El nivel sociocultural

Junto al ‘sexo’, la otra variable destacada dentro de la malla sociolingüística en los estudios de léxico disponible es el ‘nivel sociocultural’. Para su estudio se sigue el planteamiento metodológico introducido en la sociología en los años 50 y adoptado por la sociolingüística variacionista, por la influencia de W. Labov (1972). Se trata del *modelo estratificadorio*, basado fundamentalmente en la concurrencia de tres parámetros: la profesión, el nivel de estudios y los ingresos económicos.

En consonancia con este planteamiento, para establecer el ‘nivel sociocultural’ de los informantes, el Proyecto Panhispánico establece como pauta general la utilización de estos tres parámetros básicos: las profesiones de los padres, el nivel de estudios finalizados de éstos y los ingresos de la unidad familiar. Sin embargo, como señala Fernández Smith *et al.* (2008: 40), en la mayoría de investigaciones<sup>6</sup> “se ha descartado el parámetro de ingresos, fundamentalmente porque se trata de una información desconocida para muchos de los informantes, pero también porque se le supone una mayor probabilidad de error”.

Por tanto, la variable ‘nivel sociocultural’ es el resultado de sumar, sin ponderación, las puntuaciones obtenidas en los cuestionarios para el nivel de estudios culminados por los padres y la profesión desempeñada por éstos. De este modo, esta variable es el resultado de la post-estratificación de otras dos: ‘el grado de instrucción’ y ‘la profesión de sus padres’. Para la tabulación de ambas variables, por niveles y puntuación<sup>7</sup>, se suele aplicar la clasificación del Cuadro 5.2, de Samper Padilla y Hernández Cabrera (1997: 229–239).

Tras la suma de las diferentes puntuaciones, la mayoría de estudios establece tres estratos socioculturales, que se corresponden con las tres referencias clásicas en los estudios de estratificación social, en los que se separan los siguientes intervalos:

---

<sup>6</sup> Entre los investigadores que dan a la variable carácter multidimensional y añaden los ingresos familiares se hallan Alba (1995), López Morales (1999a), Carcedo González (2001). Por otro lado, algunos estudiosos como Galloso Camacho (2002b) y Arnal Purroy *et al.* (2004) sólo consideran los estudios de los padres, en tanto que Serrano Zapata (2003) añade además un parámetro de nivel cultural de la familia.

<sup>7</sup> Con respecto al ‘Nivel de estudios de los padres’, en los trabajos de González Martínez (2002), Bartol Hernández (2004: 16), Hernández Muñoz (2006: 255) y Fernández Smith *et al.* (2008: 43) se consideran solo cinco posibilidades, eliminándose el nivel de Enseñanza Secundaria Obligatoria. Por su parte, Carcedo González (2001: 30) distingue hasta ocho posibilidades en el nivel de estudios, ya que tiene en cuenta los casos en los que las personas no han concluido la educación primaria y el bachiller elemental, de manera que hay dos niveles de educación primaria y dos del bachiller elemental. Por lo que se refiere al ‘Nivel profesional’, Bartol Hernández (2004a: 17) y Hernández Muñoz (2006: 255) establecen, como en el nivel de estudios, cinco grupos profesionales, y además distinguen dentro de estos las situaciones de inactividad o desempleo separadamente del resto. Aquí también, Carcedo González (2001: 30) amplía el rango de posibilidades hasta siete, pues sitúa por debajo en la escala a los profesionales liberales respecto de los grandes empresarios y altos directivos.

**Cuadro 5.2. Baremación de la variable ‘nivel sociocultural’ en los estudios de disponibilidad léxica**

Nivel de estudios de los padres	Padre	Madre
Sin estudios	1	1
Enseñanza primaria	2	2
Enseñanza Secundaria Obligatoria	3	3
Bachillerato superior – Formación profesional	4	4
Universitario de grado medio	5	5
Universitario de grado superior	6	6
Nivel profesional de los padres	Padre	Madre
Obrero sin cualificar	1	1
Obrero especializado	2	2
Empleados medios	3	3
Personal técnico, pequeños empresarios	4	4
Personal técnico superior, empresarios medianos	5	5
Profesionales liberales, altos directivos	6	6

Tras la suma de las diferentes puntuaciones, la mayoría de estudios establece tres estratos socioculturales, que se corresponden con las tres referencias clásicas en los estudios de estratificación social, en los que se separan los siguientes intervalos:

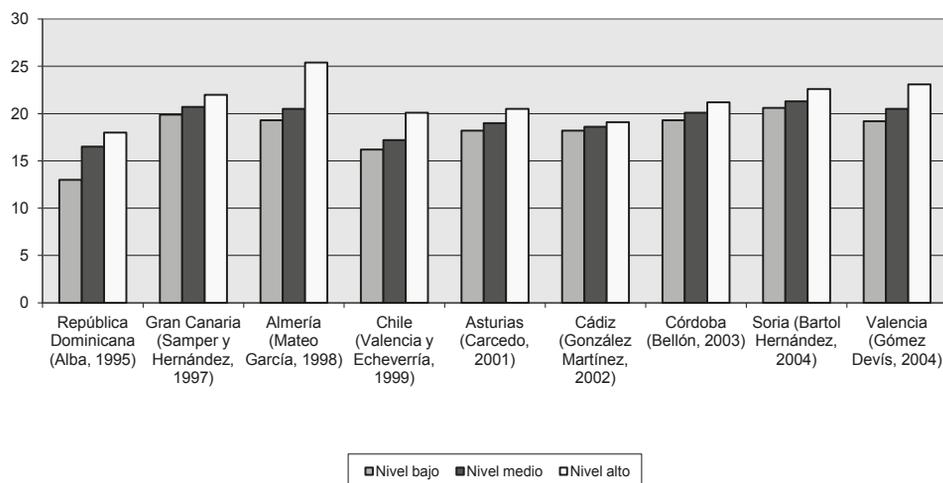
- nivel alto: 20–24 puntos,
- nivel medio: 9–19 puntos,
- nivel bajo: 4–8 puntos.

En cambio, otros autores (Arnal Purroy *et al.*, 2004; Bartol Hernández, 2004a; Hernández Muñoz, 2006; Santiago de Guervós, 2008; Ávila Muñoz<sup>8</sup>, 2006) hacen uso de una escala de cuatro niveles en la que se desdobra el nivel medio: en medio y medio-alto, o medio-bajo y medio-alto, según el caso. La razón esgrimida es que, a pesar de utilizar una escala que distribuye a los informantes en tres niveles, el perfil sociocultural de la muestra presenta una elevada concentración de informantes en torno a los valores centrales y una escasa representación de los valores extremos –más bajos y más altos (Bartol Hernández, 2004a: 17; Hernández Muñoz, 2006: 255). Una distribución en cuatro niveles reflejaría con mayor fidelidad las características económicas y demográficas de la zona en cuestión, donde el peso específico del nivel medio es mucho mayor que el de los restantes sociolectos.

Sea como fuere, la variable ‘nivel sociocultural’ es una de las que ofrece mayor rendimiento en los estudios de disponibilidad léxica. No en balde, “con

<sup>8</sup> Ávila Muñoz (2006) lo divide en medio-bajo, medio y medio-alto, desapareciendo, así, el nivel alto.

mayor o menor contundencia, los resultados han presentado, casi sin excepción<sup>9</sup>, una innegable relación entre el vocabulario disponible de los encuestados y el nivel sociocultural en el que se encuadran” (Carcedo González, 2001: 74). Así lo muestra el Gráfico 5.1.



**Gráfico 5.1. Disponibilidad léxica, según el nivel sociocultural**

Estas cifras se repiten con cierta regularidad y apuntan a la menor disponibilidad de los hablantes de niveles socioculturales bajos del espectro<sup>10</sup>. Esto ha hecho que los trabajos de disponibilidad léxica coincidan de alguna manera con la teoría del déficit elaborada por Bernstein (1971, 1975). Esta teoría postulaba que algunos sociolectos se distinguían de otros, superiores en el espectro social, no solo por el empleo de formas lingüísticas diferentes, sino por una dicotomía más acusada: la presencia/ausencia de elementos. Aunque también hubiese entre ellos diferencias de variación, su código restringido se oponía al elaborado por factores esencialmente cuantitativos: en el nivel léxico, la limitación del número de vocablos, el escaso empleo de sinónimos, o la selección rígida y limitada de adjetivos y adverbios (López Morales, 1996: 250; Romaine, 1996: 232).

<sup>9</sup> Una interesante excepción la constituye el trabajo de Benítez Pérez (1992). En su estudio sobre el área metropolitana de Madrid documenta que no siempre los porcentajes más altos correspondían a los informantes que pertenecían al nivel social superior del espectro.

<sup>10</sup> En el gráfico anterior llama poderosamente la atención, en relación con los datos de Almería, la distancia que separa a los informantes del nivel alto (25,4 palabras por informante) de los otros dos estratos medio (20,5 palabras) y bajo (19,3 palabras). Según Mateo García (1998: 222), esos datos “hacen pensar en una polarización sociocultural del tipo alto/no alto, como auténtico elemento responsable de la estratificación promovida por esta variable”. Una situación muy similar se repite en Ceuta (Ayora Esteban, 2006: 109), donde el nivel alto (25,75), se opone a los niveles medio (20,75) y bajo (21,48).

De hecho, la primera incursión de López Morales (1979) en el factor sociocultural estaba dedicada a establecer una posible correlación entre disponibilidad léxica y la teoría del déficit. Sin embargo, para López Morales (1979) la teoría del déficit de Bernstein, basada en las diferencias cuantitativas entre los ‘códigos amplios o elaborados’ de los niveles socioculturales altos y los ‘códigos restringidos’ propios de las clases bajas, no basta, por sí sola, para dar explicación de tales diferencias sociolectales. Para López Morales (1996: 256), lo único que está fuera de duda “es que existe base empírica para hablar de desniveles de disponibilidad léxica, y que estos están en clara asociación con factores sociales, entre ellos el nivel sociocultural de los sujetos”.

Las divergencias entre los niveles sociales no se limitan al aspecto cuantitativo. López Morales (1979) fue el primer autor que observó divergencias de tipo cualitativo. En su primer trabajo sobre Puerto Rico descubrió que algunos vocablos relacionados con el centro de interés ‘El cuerpo humano’ eran utilizados únicamente por el sociolecto más alto: *tórax, occipital, glándula pituitaria, epidermis, cromosoma, cordón umbilical, arteria, bazo, iris*.

En la misma línea, Gómez Devís (2004) en Valencia, y Samper Hernández (2003a) en Gran Canaria llegaron a conclusiones parecidas al analizar las palabras más disponibles<sup>11</sup> de cada sociolecto. Ambas autoras encontraron importantes discrepancias entre los sociolectos, con diferencias muy significativas en los campos ‘Juegos y distracciones’ y ‘Profesiones y oficios’. En ‘Juegos y distracciones’ es posible delimitar un vocabulario caracterizador de cada grupo sociolectal dentro del léxico altamente disponible. En Valencia son exclusivas del nivel alto palabras como *paddle, piragüismo, internet, hípica* (en Gran Canaria lo son *paddle, piano, hípica, navegar*); el nivel medio incluye términos como *videojuego* o *yudo*; en tanto que palabras como *tres en raya* o *cuerda* caracterizan al nivel bajo valenciano.

En cuanto al centro de interés ‘Profesiones y oficios’, en Valencia el nivel alto incorpora *notario, veterinario, cirujano, administrador, oftalmólogo* o *matemático*; el medio *administrativo, ganadero* o *químico*; y el bajo, *prostituta, piloto, militar, obrero* o *pescadero*. En Gran Canaria, las lexías exclusivas del sociolecto alto son *ministro, notario, traumatólogo* o *ginecólogo*; en tanto que el sociolecto bajo presenta como palabras propias *freganchín, vigilante de seguridad* o *limpiar*.

El análisis de las diferencias socioculturales en el léxico disponible permite estudiar, según Samper Padilla (2006: 116), “la competencia entre un dialectalismo y el término del español general o entre dos dialectalismos”, de manera que “la disponibilidad más inmediata de uno de ellos en el nivel más alto puede ser indicio de su aceptación social”. Así lo ejemplifica el estudio que Hernández Cabrera y Samper Hernández (2002–2004) dedican a la disponibilidad de algunas parejas de lexías en el centro de interés ‘La ropa’ y su relación con la norma culta

---

<sup>11</sup> Gómez Devís (2004) analiza las cincuenta primeras palabras de cada sociolecto, y Samper Hernández (2003b) las listas con el 75% del índice de frecuencia acumulado.

y popular en Gran Canaria. En estas encuestas, los niveles más altos muestran preferencia por el anglicismo *suéter*, frente a otro término también procedente del inglés, *pulóver*, más usado en los otros niveles. Por otro lado, aunque la voz *pendiente*, forma estándar en el español de España, tiene tendencia a imponerse en todos los sociolectos, aún se aprecia “la preferencia de los jóvenes del sociolecto bajo por la voz *zarcillo*” [vocablo dialectal canario], lo que “está en consonancia con los datos que ofrecen las encuestas de la norma popular urbana” (Hernández Cabrera y Samper Hernández, 2002–2004: 1352).

### 5.3. El factor espacial: urbano/rural

También los estudios de disponibilidad léxica se han ocupado de estudiar la incidencia de la procedencia geográfica sobre el léxico disponible. Dependiendo de las características físicas de las comunidades estudiadas y de los objetivos de la investigación, se han analizado en relación con el factor espacial las dicotomías urbano/rural, centro/periferia, litoral/interior, etc., en clara alusión no sólo a factores meramente geográficos sino también a condicionamientos económicos y culturales.

No faltan, sin embargo investigadores que, de acuerdo con la realidad de la comunidad que analizan, han considerado más relevante una división en varias zonas. Puerto Rico, Almería, Valencia y Alicante son las comunidades en las que se ha adoptado esta decisión. En Puerto Rico, López Morales (1999a), basándose en un criterio estrictamente demográfico, distingue cuatro zonas en función del número de habitantes: zona 1 (la metropolitana de San Juan de Puerto Rico, con una población superior a los dos millones de habitantes), zona 2 (entre 190.000 y 80.000 habitantes), zona 3 (79.000–25.000) y zona 4 (24.999–5.000). En Valencia, Gómez Devís (2004) efectúa una división muy parecida, al dividir la provincia valenciana también en cuatro zonas: la ciudad de Valencia (750.000 habitantes), el área metropolitana (500.000 habitantes), la zona semiurbana (poblaciones sobre 50.000 habitantes), y la zona rural (pueblos agrícolas de 10.000 habitantes). Por su parte, Mateo García (1998) en Almería prefiere tener la distinción geográfica entre zonas según sus condicionantes socioeconómicos, con lo cual divide en territorio almeriense en cuatro ámbitos: la capital, el interior, el litoral este y el litoral oeste. La división entre zona litoral y zona del interior también es usada en el estudio de Alicante por Martínez Olmos (2007).

De cualquier forma, la división más frecuentemente analizada ha sido la de medio urbano frente a medio rural. De hecho, formaba ya parte de los primeros trabajos para el francés (Gougenheim *et al.*, 1956, 1964, 1967). Una cuestión metodológica importante ha sido la definición de los núcleos urbanos frente a los rurales. Como hemos visto, en Valencia Gómez Devís (2004) califica como rural

a aquel núcleo poblacional con una población inferior a 10.000 habitantes. Coincide así con González Martínez (1997) que en su tesis doctoral sobre la disponibilidad léxica en Cádiz, siguiendo el criterio del padrón de enero de 1995 del Instituto Nacional de Estadística, consideró población urbana a toda aquella que superaba los 10.000 habitantes. Ahumada Lara (2006) eleva en Jaén el límite demográfico entre lo rural y lo urbano, considerando núcleos rurales a aquellos con menos de 25.000 habitantes.

Así las cosas, según indica Samper Padilla y Samper Hernández (2006: 64), la solución más empleada ha sido la de equiparar la diferencia urbano/rural a la que se produce entre la capital y el resto de la provincia. Además de este criterio general, esta división se ha complementado con la del lugar de residencia de los padres. Así se hace en los estudios de Zamora, Ávila y Salamanca (Gallos Camacho, 2003a), de Aragón (Arnal Purroy *et al.*, 2004), de Soria (Bartol Hernández, 2004a), de Huelva (Prado Aragonés y Gallos Camacho, 2005), de Castellón (Casanova Ávalos, 2006), de Castilla-La Mancha (Hernández Muñoz, 2006), etc.

Como en los primeros trabajos del ámbito francés, la mayoría de trabajos hispánicos documentan índices superiores de léxico disponible en entornos urbanos, como ocurre en Dorado, Puerto Rico (Román-Morales, 1985), Gran Canaria (García Domínguez *et al.*, 1994), Zamora (Gallos Camacho, 1997, 1998a), Chile (Valencia y Echeverría, 1999a), Asturias (Carcedo González, 2001), Cádiz (González Martínez, 1997), Soria (Bartol Hernández, 2004a), Huelva (Prado Aragonés y Gallos Camacho, 2005), Málaga (Ávila Muñoz, 2006), Aragón (Enguita Utrilla, 2008), etc.

El trabajo de García Domínguez *et al.* (1994: 68–71) sobre Gran Canaria, además de ser el primero en el ámbito español, constituye un buen ejemplo de todos los demás. Este estudio evidenció que los estudiantes urbanos superaban a los rurales, aunque las diferencias a favor de las zonas urbanas resultaban poco significativas. En opinión de estos autores, eran varios los factores que contribuían a limar las diferencias léxicas entre ambas zonas, y, por consiguiente, a nivelar el conocimiento y uso lingüístico en la isla: la escasa distancia geográfica entre los distintos núcleos de población de las islas, la escuela y, sobre todo, los medios de comunicación.

En otros casos, las diferencias llegan a ser más notorias, como en Soria (Bartol Hernández, 2004a), donde hay más de dos puntos de diferencia en 7 campos léxicos; en Huelva (Prado Aragonés y Gallos Camacho, 2005), con diferencias muy amplias de cuatro y cinco puntos en seis centros de interés, a favor de los estudiantes que residen en Huelva capital; o en Segovia (Santiago de Guervós, 2008), donde los centros urbanos superan a los centros rurales, por diferencias entre dos y tres puntos, según el campo léxico del que se trate. La única excepción la plantea el centro de interés de ‘Trabajos de campo y jardín’ (más propio del mundo rural), aunque con una diferencia no significativa de 0,29 puntos. Justamente

han sido los campos léxicos ‘Trabajos del campo y del jardín’ y ‘El campo’ los únicos que habitualmente presentaban porcentajes de superiores en los centros ubicados en zonas rurales. Así ocurría también en Huelva (Prado Aragonés y Galloso Camacho, 2005).

Sin embargo, también hay investigaciones en las que se presenta un comportamiento diferente de esta variable. Este es el caso de Gómez Devís (2004) en Valencia, donde globalmente son los sujetos de la zona rural los que obtienen el promedio general más alto (20,86), frente a los capitalinos (20,45). Los habitantes de las zonas semiurbanas y del área metropolitana –los más afectados por la inmigración– quedan por detrás (19,86 y 19,20, respectivamente). En diez centros de interés los alumnos de zonas rurales superan a los de la ciudad, destacando con tres o más unidades de ventaja sobre el área metropolitana en ‘Alimentos y bebidas’, ‘El campo’, ‘Animales’ y ‘Trabajos del campo y del jardín’. Gómez Devís (2004: 153–154) concluye que estadísticamente la variable ‘lugar de residencia’ tiene un valor discriminatorio medio, que “puede estar motivado por la configuración de la provincia de Valencia gracias a sus buenas comunicaciones (viarias, informativas, profesionales), así como por el grado de desarrollo de las distintas zonas geográficas”.

Con posterioridad, otros estudios también han arrojado resultados en los que los informantes de la zona rural también superan abiertamente a los que proceden de las ciudades. Samper Padilla y Samper Hernández (2006: 70) subrayan que “esa superioridad se constata de forma especial en los centros de interés relacionados con el campo y con las tradiciones populares. Así ocurre con los resultados de la investigación de Castilla-La Mancha (Hernández Muñoz, 2006), que coinciden en parte con los de Valencia. Entre los ámbitos que resultan estadísticamente significativos se encuentran ‘El campo’, ‘Los animales’, ‘Trabajos del campo y del jardín’, ‘Alimentos y bebidas’, ‘La cocina y sus utensilios’ y ‘Los muebles de la casa’.

Igualmente en Alicante, Martínez Olmos (2007: 190) encuentra que sólo cuatro áreas temáticas consiguen mejores promedios en el sociolecto urbano (‘Partes del cuerpo’, ‘Los muebles de la casa’, ‘Alimentos y bebidas’ y ‘Juegos y distracciones’. Otro caso similar se presenta en Granada, donde Pastor Milán y Sánchez García (2007: 37) notan cómo los alumnos de zonas rurales registran un promedio mayor en 6 centros de interés: ‘La ropa’, ‘Los alimentos’, ‘El centro educativo: muebles y material escolar’, ‘Trabajos del campo y del jardín’, ‘Medios de transporte’ y ‘Profesiones y oficios’.

Entre los estudios de tipo cualitativo, el más destacado es el que dedica Engrita Utrilla (2008: 77–101) al contraste norma léxica urbana frente a norma léxica rural de jóvenes aragoneses en torno al centro de interés ‘Profesiones y oficios’. A pesar de compartir un 50% del vocabulario, los modos de vida y las actividades laborales del medio urbano y del medio rural influyen en léxico disponible de sus habitantes. De este modo, palabras como *funcionario*, *relaciones públicas*, *celador* o *cerrajero* poseen índices más altos en el medio urbano, en tanto que

*granjero, practicante, gasolinero o cartero* aparecen con un índice de disponibilidad más alto entre los hablantes del medio rural. Por la misma razón, aparecen exclusivamente un buen número de vocablos entre los informantes urbanos que hacen referencia a actividades más propias o más frecuentes en la ciudad (*directivo, forense, guardaespaldas, geriatra*) o que implican una mayor especialización laboral (*auxiliar de ATS, auxiliar sanitario, auxiliar administrativo, subinspector, vicerrector*), mientras que los informantes rurales producen más vocablos que tienen presencia más viva en las áreas rurales (*pocero, tractorista, confitero o boticario*). También puede constatararse una mayor presencia de extranjerismos (*pizzero o parking*) y un superior empleo de morfemas derivativos (*realizador, baloncestista, confeccionista*) entre los estudiantes que viven en las capitales de las provincias.

#### 5.4. El tipo de centro escolar

Los estudios de disponibilidad léxica se han encargado asimismo de investigar la incidencia que puede tener optar por la enseñanza pública o por la enseñanza privada en la producción léxica de los escolares. Como afirma Mateo García (1998: 53), “implícitamente se considera que entre las instituciones educativas públicas y privadas existen diferencias significativas. Estas diferencias empiezan en las mismas posibilidades de acceso a ellas y [...] se manifiestan en relación con sus modelos de enseñanza y a las redes sociales en las que permiten desenvolverse a sus alumnos”.

En palabras de Samper Padilla y Samper Hernández (2006: 72), este factor social “se ha convertido en uno de los condicionantes extralingüísticos que mejor explica las divergencias en el vocabulario que tienen a su disposición los hablantes de las distintas comunidades”. Aunque se ha discutido su relación con la variable ‘nivel sociocultural’, la consideración del tipo de centro como una variable independiente se basa, según los mismos autores, “por un lado, en que las muestras según los dos condicionantes no son del todo paralelas y, por otro, en que no siempre hay coincidencia en los resultados de uno y otro factor”.

Metodológicamente hay que hacer notar que, dentro del Proyecto Panhispánico, las muestras han estado compuestas mayoritariamente por alumnos de centros públicos, en correspondencia con la composición de los universos estudiados. Ello se debe a que, en general, se ha aplicado un procedimiento de muestreo por cuotas proporcionales<sup>12</sup>. Samper Padilla y Samper Hernández (2006: 74) muestran en el artículo citado que el número de estudiantes de los centros públicos supera

---

<sup>12</sup> La excepción a esta norma la constituye el trabajo de O. Alba (1995a), en el que encuestó a 117 alumnos que habían cursado sus estudios previos en centros públicos y a 230 que lo habían hecho en instituciones privadas.

ampliamente al de los colegios privados. En concordancia con los niveles de renta en la Península Ibérica y con los datos del Ministerio de Educación y Ciencia, se observa que cuanto más al sur, más bajos son los porcentajes de alumnos de centros privados (inferiores al 20% en Castilla-La Mancha, Huelva y Extremadura), ascendiendo<sup>13</sup> en la zona centro (se supera el 30% en Valencia, Burgos y Madrid), hasta alcanzar el 41% en Barcelona.

Como muestra el Cuadro 5.3, los resultados casi siempre coinciden en señalar un mayor bagaje léxico para los alumnos de instituciones privadas. A los estudios desarrollados en Hispanoamérica de Alba (1995a), López Morales (1999a) y Valencia y Echeverría (1999a), se unen los desarrollados en España por García Domínguez *et al.* (1994), Samper Padilla y Hernández Cabrera (1995b), González Martínez (1997), Mateo García (1998), Carcedo González (2001), Ayora Esteban (2003), Bellón Fernández (2003), Galloso Camacho (2003a), Bartol Hernández (2004a), Gómez Devís (2004), Serrano Zapata (2004), etc. En muchos casos, los resultados son estadísticamente significativos a favor del grupo educado en colegios privados, mostrando desde la perspectiva cuantitativa una superioridad léxica casi constante del grupo que estudia en centros privados, como así ocurre en los estudios de Puerto Rico, República Dominicana y Chile, –y dentro de España– en Almería, Segovia, Valencia, Ceuta, Galicia, Salamanca, Gran Canaria, Huelva, Sevilla, Lérica, y otros.

**Cuadro 5.3. Promedio de palabras por informante, según centro escolar, en investigaciones de España**

Comunidad	Investigación	Público	Privado	Diferencial
1	2	3	4	5
Almería	Mateo García (1998)	8,65	13,63	+4,98
Segovia	Santiago Guervós (2008)	19,37	23,71	+4,34
Valencia	Gómez Devís (2004)	19,33	21,90	+2,57
Ceuta	Ayora Esteban (2003, 2006)	20,59	23,13	+2,54
Galicia	López Meirama (2008, 2011)	21,12	23,40	+2,28
Salamanca	Galloso Camacho (2003a)	19,62	21,56	+1,94
Gran Canaria	Samper Padilla y Hernández Cabrera (1995b)	20,28	21,76	+1,48
Huelva	Prado Aragonés y Galloso Camacho (2005)	20,43	21,75	+1,32
Sevilla	Trigo Ibáñez (2007)	19,38	20,63	+1,25

<sup>13</sup> Según el Ministerio de Educación y Ciencia, en el caso de Madrid, Barcelona y País Vasco, el porcentaje real de población estudiantil que cursaba estudios en centros privados en el curso 2003/04 rondaba el 50%.

1	2	3	4	5
Lérida	Serrano Zapata (2004)	18,21	19,34	+1,13
Castilla-La Mancha	Hernández Muñoz (2006)	18,82	19,63	+0,81
Asturias	Carcedo González (2001)	18,80	19,60	+0,80
Córdoba	Bellón Fernández (2003, 2011)	19,77	20,55	+0,78
Cádiz	González Martínez (1997, 2002)	18,44	18,94	+0,50
Soria	Bartol Hernández (2004a)	21,22	21,65	+0,43
Alicante	Martínez Olmos (2007)	19,13	19,36	+0,23
Castellón	Blas y Casanova (2003a)	21,43	20,80	-0,63
Granada	Pastor Milán y Sánchez García (2008)	18,00	17,00	-1,00

Como decíamos más arriba, esta prevalencia de los estudiantes de centros privados sobre los públicos en ocasiones se ha visto combinada con el nivel sociocultural al que éstos pertenecen. Sobre este particular, Carcedo González (2001: 69) recuerda que “los centros privados se nutren con frecuencia de un alumnado procedente mayoritariamente del nivel sociocultural alto”, y, de acuerdo con la mayoría de estudios, los índices de disponibilidad léxica aumentan a medida que se asciende en la escala social. Sin embargo, como también ya se ha apuntado, no siempre se produce un paralelismo claro entre un factor y otro (se pueden encontrar muchos niños de nivel alto en escuelas públicas –sobre todo en las céntricas– y, en menor medida, niños de los niveles bajo y medio-bajo en centros privados).

En cualquier caso, en el caso de Almería, dada la notoria diferencia en la producción de nuevos vocablos entre los colegios privados (68,31) y públicos (32,7), siempre a favor de los alumnos de colegios privados, Mateo García (1998: 188–189) afirma que “no es en absoluto descabellado enfocar la interpretación de esos datos desde la perspectiva de las teorías de Bernstein, sobre todos si tenemos en cuenta que [...] se establece una correlación bastante estrecha entre enseñanza privada y alumnado procedente de estratos altos o medio-altos”. De este modo, la escuela pública no habría sido capaz de nivelar por completo las diferencias sociales con respecto a los alumnos de los centros privados, lo que se viene manifestando en el inferior rendimiento escolar e intelectual de los alumnos de la enseñanza pública con respecto a la enseñanza privada, lo que es puesto de manifiesto por los diferentes informes que emite el Ministerio de Educación y Ciencia. Al respecto, Blas y Casanova (2003a: 19) hacen referencia al informe correspondiente al año 2000, en el que “se concluía que la media de alumnos de centros privados puntúa por encima de los alumnos de centros públicos en una serie larga de materias académicas, entre las que se encuentran asignaturas básicas del *currículum* escolar como las matemáticas o la lengua”.

En lo cualitativo, González Martínez (2002)<sup>14</sup> señala que los alumnos de centros públicos utilizan un léxico menos estandarizado, con más dialectalismos y vocablos populares, que los estudiantes de colegios privados. Del mismo modo, Gómez Devís (2004: 224), tras analizar los 50 primeros vocablos de los listados, concluye que, aunque la influencia decisiva de esta variable sobre la productividad léxica “queda algo neutralizada por el porcentaje de convergencia cualitativa” entre el grupo público y el privado, “puede comprobarse, en los vocablos diferentes de algunos centros de interés, una mayor riqueza, precisión y estandarización léxicas, así como cierta preocupación por el registro en la enseñanza privada”. Los ejemplos que aporta esta estudiosa se muestran en el Cuadro 5.4.

**Cuadro 5.4. Palabras aparecidas entre los 50 primeros vocablos de Valencia (Gómez Devís, 2004: 224)**

Centros	Público	Privado
‘Cuerpo humano’	<i>culo, testículo</i>	<i>tibia, peroné, páncreas</i>
‘Ropa’	<i>body, short</i>	<i>chubasquero, traje</i>
‘Partes de la casa’	<i>hall</i>	<i>hall, vestíbulo</i>
‘Iluminación...’	<i>abrir ventana, soplar, paipay, balcón</i>	<i>ojo de buey, candelabro, claraboya</i>
‘Colores’	<i>amarillo limón, amarillo canario, ....</i>	<i>ámbar, burdeos, marfil, pardo</i>

Pese a la tendencia general de una mayor competencia léxica de los alumnos de colegios privados, han aparecido trabajos en los que se llega a resultados diferentes. Este es el caso del estudio de Blas y Casanova (2003a), dedicado a investigar la influencia de la red educativa y del entorno sociocultural en la disponibilidad léxica. Tomando como base una muestra de 246 alumnos de COU<sup>15</sup> de la provincia de Castellón, procedentes de manera equilibrada de cinco centros públicos y cinco privados, llegan a la conclusión de que “el comportamiento de los alumnos de colegios públicos y privados apenas difiere, al menos en términos cuantitativos”, y de que “el nivel de congruencia<sup>16</sup> entre los ítems léxicos disponibles es también considerablemente elevado”, con lo que establecen “la escasa relevancia del factor que distingue entre alumnos de centros públicos y alumnos de centros privados como variable independiente aislada” (Blas y Casanova, 2003a:

<sup>14</sup> Citado en: Samper Padilla y Samper Hernández (2006: 74).

<sup>15</sup> Curso de Orientación Universitaria, en el nuevo sistema educativo, 2º de Bachillerato.

<sup>16</sup> No obstante, Blas y Casanova (2003) sí notaron diferencias en cuanto a la evocación de una serie de profesiones de distinto estatus social entre las 50 palabras más disponibles de cada grupo. Así, frente a vocablos como *farmacéutico, director, notario, dentista* o *filólogo*, mencionadas por los alumnos de centros privados, los estudiantes de centros públicos incorporan *prostituta, zapatero* o *tendero* (citado por: Samper Padilla y Samper Hernández, 2006: 78).

23–25). A la luz de los resultados de la provincia de Castellón, estos autores concluyen en el caso de Castellón “el propio sistema educativo ha podido contribuir a eliminar los puntos de partida diferentes y ha equilibrado los valores de la disponibilidad léxica en el conjunto de la población escolar (Blas y Casanova, 2003a: 27), lo que contrasta con la opinión de Mateo García (1998).

No es este el único caso, en el que las diferencias entre ambos grupos son muy reducidas, como se observa en el Cuadro 5.3. En las investigaciones de Cádiz, Soria y Alicante, el diferencial en el promedio de palabras entre ambos grupos es, de hecho, inferior a medio punto. Más aún, hay un estudio el que el diferencial en el promedio de palabras es de un punto a favor de los colegios públicos. Se trata del caso de Granada (Pastor Milán y Sánchez García, 2008), donde, de los veinte centros encuestados, dos son privados concertados.

Samper Padilla y Samper Hernández (2006: 78) se refieren, muy acertadamente, a la influencia que la reglamentación educativa puede tener en la composición del alumnado de los colegios privados, al condicionar la recepción de subvenciones estatales a criterios de admisión similares a los vigentes en las instituciones públicas. Esto hace aseverar a los estudiosos canarios que “no cabe duda de que este paso a concertados de muchos colegios privados puede repercutir en el tipo de alumnado que reciben“, lo que les lleva a preguntarse si “influirá esto en la disponibilidad léxica”. El caso de Granada parece corroborar esta aseveración, a la que Galloso Camacho y Prado Aragonés (2005) recurrían para justificar las diferencias entre Salamanca (centros auténticamente privados) y Huelva (centros concertados)

## 5.5. La edad

En opinión de Mateo García (1998: 49), “en la edad probablemente resida una de las grandes asignaturas pendientes en relación con la estratificación social del léxico disponible”, ya que son pocos los “estudios que se han ocupado de ella con cierta sistematicidad”. Efectivamente, el factor ‘edad’ suele quedar en segundo plano, al hallarse en función del ‘nivel de estudios’. Esta correspondencia se pone en relación con la didáctica de la lengua materna, y más concretamente con la evaluación del desarrollo de la competencia léxica.

De hecho, hasta la publicación del *Léxico disponible de Puerto Rico* (López Morales, 1999a) y la puesta en marcha del Proyecto Panhispánico, la mayoría de las investigaciones se había dedicado a realizar un seguimiento del desarrollo de la competencia léxica de los alumnos en distintos momentos del proceso educativo. Los niños habían sido, además, los elegidos para llevar a cabo las primeras investigaciones en Francia y Canadá; tanto Gougenheim *et al.* (1956, 1964) como Mackey (1971a), habían hecho sus investigaciones con escolares de 9 a 12 años de edad.

En el mundo hispánico, los estudios de disponibilidad léxica comenzaron con el análisis que López Morales (1973) lleva a cabo en San Juan de Puerto Rico acerca de la incorporación del vocabulario estudiado en tres diferentes niveles de enseñanza primaria: 1º, 3º, 5º, correspondientes a niños de 6, 8 y 10 años. Allí detecta irregularidades importantes en la competencia léxica de los escolares, ya que los resultados “parecen hablar en contra de una proporción gradual en el proceso de adquisición léxica” (López Morales, 1978: 82). Observa que “la adquisición léxica emprende una drástica baja cuando se esperaría que la línea se mantuviese en ascenso” y que “se produce una caída cuantitativa en los mismos niveles, de tercero a quinto” (López Morales, 1978: 83). López Morales achacaba esta anomalía a deficiencias en la programación de los contenidos léxicos, con carencias en cuanto al léxico frecuente y disponible.

Años después, López Chávez (1993, 1995a) mide en México la evolución de la disponibilidad léxica desde el nivel preescolar hasta sexto de primaria. Descubre sorprendentemente que “no parece que el incremento sea directamente proporcional al aumento de edad, ni cuantitativamente con relación al número de vocablos manejados por el niño, ni cualitativamente con relación a las distinciones semánticas que ya son capaces de incorporar a sus repertorios desde una edad relativamente temprana” (Mateo García, 1998: 49). Por ello mismo, con anterioridad, el mismo autor había llamado la atención sobre la necesidad de “presentar, proponer y aceptar nuevas perspectivas que ofrezcan al estudiante métodos más seguros y efectivos en el lento dominio de la lengua materna” (López Chávez, 1988: 5).

Otros estudios, por el contrario, sí han mostrado una progresión en el dominio léxico, en relación con la edad y curso escolar. Tal es el caso de Román-Morales (1985), que investiga la disponibilidad léxica en escolares de Dorado, Puerto Rico. Para ello, estudia a alumnos de los mismos cursos elegidos por López Morales en 1973: 1º, 3º y 5º. El análisis de los datos demuestra “que existe una relación asociativa entre el grado escolar de los estudiantes de primero, tercero y quinto grados del Distrito Escolar de Dorado y su disponibilidad léxica ya que a medida que aumenta el grado escolar, también aumenta la disponibilidad léxica” (Román-Morales, 1985: 61).

Por su parte, Mena Osorio (1986: 173) en un estudio a 300 escolares de los niveles 4º, 6º y 8º de primaria en Concepción, Chile, concluye que “el aumento del promedio de respuestas por alumno es bastante regular, puesto que entre uno y otro nivel escolar [se observa] un incremento aproximado de un 20%”: 60,1% para cuarto, 80,2% para sexto, 100% para octavo. Con respecto a las palabras diferentes, “el crecimiento es también muy regular: entre cada nivel escolar [se observa] un aumento de aproximadamente un 15% en la cantidad de palabras diferentes”, 69,9% para cuarto, 85,5% para sexto, 100% para octavo.

Echeverría (1991) también lleva a cabo un análisis de la disponibilidad en escolares chilenos. La muestra está formada por 300 estudiantes de 4º, 6º y 8º grado de la escuela básica, y 400 de los cursos primero y tercero de educación

media. Los resultados del análisis son igualmente positivos, ya que se percibe “un aumento gradual del número de vocablos, así como del promedio de respuestas a medida que asciende el nivel de escolaridad” (Carcedo González, 2000a: 34).

En la República Dominicana, Orlando Alba (1996: 749) constata que “a medida que se asciende en el nivel de escolaridad aumenta la disponibilidad léxica” por lo que se felicita, dado que ello “revela una efectiva acción de la escuela en el crecimiento lingüístico de los estudiantes”. Su investigación se basa en una muestra de 65 estudiantes de octavo curso de educación media (13–14 años) y del primer año universitario.

Por lo demás, en España han seguido esta línea de investigación, entre otros, estudiosos como Samper Hernández (2003a) en Gran Canaria, que analiza la evolución de la disponibilidad léxica en escolares de enseñanza primaria y secundaria; García Megía (2004) en Almería, que también estudia la evolución de la disponibilidad léxica entre los 9 y los 12 años; Prado Aragonés y Galloso Camacho (2006a, 2006b) en Huelva, que analizan dos niveles, 6º de primaria y 2º de bachillerato; García Casero (2013) en Santander, que se ocupa del léxico disponible en estudiantes de 4º de Enseñanza Secundaria Obligatoria, etc.

Yendo más allá de la etapa escolar, lo cierto es que, como señala Hernández Muñoz (2006: 123) “las investigaciones, [realizadas desde la psicolingüística], concuerdan en que la edad y el nivel educativo son las variables que mejor predicen la actuación léxica de los sujetos en las pruebas de fluencia de categoría”, a las que la disponibilidad léxica pertenece. Así, Urrutia Martínez (2001) en la Universidad de la Concepción de Chile, entre otras pruebas para valorar la competencia léxica de los sujetos, incluyó la disponibilidad léxica. Las entrevistas que realizó a 30 sujetos, divididos en tres grupos de edad (15–25, 35–40 y 65–70), con 10 informantes cada uno, le sirvieron para comprobar la influencia de la edad en el decaimiento de la producción, e identificar, asimismo, otros elementos que contribuyen a la mejora de las habilidades de producción: la creación de subgrupos dentro del campo nocional y la capacidad para cambiar entre ellos.

Desde otra perspectiva, en el Proyecto Panhispánico, el factor ‘edad’ ha quedado neutralizado debido al propósito principal de estas investigaciones: la descripción de la norma léxica adulta. Por ello, la mayoría de las investigaciones trabaja con estudiantes preuniversitarios, sujetos de 18 años con plena madurez léxica<sup>17</sup>, pero sin la especialización léxica de una profesión o de unos estudios universitarios (Samper Padilla *et al.*, 2003: 65).

No obstante, Borrego Nieto (2004: 65) se cuestiona “si representa realmente el informante preuniversitario la norma culta panhispánica que se busca” y afirma que “harían falta trabajos que compararan con los resultados de hablantes cultos de mayor edad”. A solventar esta carencia el propio Borrego Nieto (2002b, 2008)

---

<sup>17</sup> De hecho, al seleccionar informantes de último curso de enseñanza media, se busca también disminuir la influencia de la lengua infantil o adolescente (López Morales, 1999a: 28).

dedica dos estudios en los que coteja la disponibilidad léxica de una muestra de 21 jubilados de ambos sexos de una residencia de la tercera edad de la ciudad de Salamanca, mayores de 70 años<sup>18</sup> y con un nivel cultural predominantemente bajo, con dos muestras de informantes jóvenes: la primera, una muestra de 21 estudiantes preuniversitarios del Instituto de Bachillerato “Torres Villarroel” de Salamanca, que se incluía entre los preuniversitarios estudiados por Galloso Camacho (2002b), y la segunda, una muestra de 24 estudiantes del último curso de Filología Hispánica de la Universidad de Salamanca. Efectúa la comparación en los campos léxicos de: ‘La ropa’, ‘Alimentos y bebidas’, ‘La ciudad’, ‘Juegos y distracciones’, y ‘Profesiones y oficios’.

Los resultados del grupo de jubilados, expuestos en el artículo titulado “Edad y cultura léxicas” (Borrego Nieto, 2008), están muy alejados de los obtenidos por los universitarios y preuniversitarios salmantinos, además de los alumnos aragoneses también allí comparados, que entre sí muestran en ciertos aspectos más semejanzas que con sus conciudadanos mayores. Entre las diferencias cuantitativas que ambos trabajos detectan están el número de palabras –del orden de 23,6 entre los jóvenes frente a 9,64 entre los jubilados– y de vocablos –también bastante menos en los jubilados: 518, frente a las 826 de los universitarios y las 831 del «Torres Villarroel»–. Las cifras de disponibilidad y de densidad léxicas entre los universitarios salmantinos y los preuniversitarios de Salamanca y Aragón son igualmente mucho más altas que en el grupo de jubilados.

Al analizar las diez palabras más disponibles de cada campo léxico, las diferencias parecen ligadas a distintas causas:

– la separación generacional en ‘Alimentos y bebidas’, donde se muestran distintos hábitos alimenticios (*garbanzos, patatas, alubias, arroz, cocido, paella* en los jubilados; por *Coca Cola, pan, naranja, whisky, naranjada, macarrones, limonada* en los jóvenes), y ‘Juegos y distracciones’, con *baile, dominó, brisca, cine, billar* entre los jubilados, y *parchís, oca, tele(visión), tenis, trivial, escondite, damas* para los jóvenes;

– factores socioculturales en ‘Profesiones’, con preferencia por profesiones manuales (*carpintero, mecánico, labrador, fontanero, comerciante, sastre*) frente a profesiones más cualificadas del grupo joven (*ingeniero, abogado, secretario, arquitecto, policía, empresario, enfermero*). Por otra parte, los jubilados asocian el género al rol tradicional de las profesiones, y excepto en el caso de *maestro/a*, que aparece en ambos géneros, citan los oficios en masculino, salvo los casos siguientes: *costurera, planchadora, cocinera, peluquera, cajera, comadrona, monja, modista, limpiadora y azafata*. Los jóvenes aragoneses también tienen un comportamiento tradicional al nombrar solo en femenino: *ama de casa, ama de llaves, azafata, bordadora, mujer de la limpieza, prostituta, señora de la limpieza* (Arnal Purroy *et al.*, 2004: 47);

<sup>18</sup> Salvo dos de ellos, todos eran mayores de 70 años. El más veterano tenía 94 años.

– la diferente forma de conceptualizar y afrontar la encuesta en ‘La ciudad’: los jubilados ofrecen vocabulario vinculado a la ciudad donde viven (*Salamanca, plaza, ciudad, monumento, catedral, plaza mayor*), en tanto que los jóvenes ofrecen un vocabulario más general (*semáforo, parque, edificio, casa, contaminación, autobús*);

– el pudor en el tema de ‘La ropa’, ya que los jubilados no mencionan palabras como *braga* o *sujetador*<sup>19</sup>.

Por lo demás, “la pertenencia a un mismo grupo social parece primar, pues, sobre las circunstancias de ámbito geográfico” (Borrego Nieto, 2008: 232), aunque, en determinados centros de interés, aparecen “diferencias de cultura regional susceptibles de reflejarse en el léxico más disponible”, como la preferencia castellana por el término *alubias*, más extraño en Aragón, o en el ámbito de los juegos de cartas, con *guiñote* de Aragón, desconocido en Castilla, o los populares juegos castellanos (*tute, mus y brisca*), poco mencionados en Aragón.

Otro aspecto a tener en cuenta es que el trabajo con personas de edad avanzada o niños muy pequeños es un factor que motiva cambios en la metodología. En el estudio comentado, Borrego Nieto (2008: 243) detectó en los jubilados que respondían a la prueba escrita, malentendidos en relación con la tarea que se les pedía “por el desajuste entre tal tarea y sus hábitos textuales: en el catálogo de los géneros que tienen registrados no figuran las ristas léxicas de carácter asociativo”. Para evitar este u otros inconvenientes, otros estudios han utilizado encuestas orales grabadas, si “el colectivo de informantes estaba poco habituado a escribir y una tarea escrita podría o bien provocar un rechazo a colaborar o unos resultados escasos y poco fiables” (Hernández Muñoz, 2006: 109). Esto es lo que hacen Borrego y Fernández Juncal (2003a) al entrevistar a personas mayores de la mitad oriental de Cantabria en su estudio sobre las relaciones entre léxico disponible y Dialectología, o Sánchez Corrales y Murillo (1999:105) al recoger el léxico disponible de los niños costarricenses.

Finalmente, el estudio de la estratificación generacional en la disponibilidad léxica sirve para documentar situaciones de cambio sociolingüístico, entre los que se incluye la mortandad sociolingüística. Así, Fuentes González (1994) y Mateo García (1996) en estudios en relación con el léxico cañero-azucarero en Motril (Granada) obtuvieron resultados concluyentes. “La mortandad léxica seguía un itinerario creciente en el espectro generacional paralelo a la caída cronológica del sector en Motril” (García Marcos, 1999: 217).

Hernández Cabrera y Samper Hernández (2002–2004) han encontrado un comportamiento similar al comparar los datos de la disponibilidad léxica y los de la norma culta de Gran Canaria, recogida por Samper Padilla (1998b). Según este

---

<sup>19</sup> En este caso, este comportamiento de tabuización de un término coincide con el que muestran los jóvenes de la comunidad grancanaria, quienes relegan el término marcado (*sostén*) y favorecen el predominio de la voz más neutra (*sujetador*) (Samper Padilla, 2006: 117).

estudio, los jóvenes grancanarios conservan la norma comunitaria en un ejemplo de variación estable (como ocurre en *chaqueta/americana*), o se inclinan por la voz más frecuente entre las generaciones anteriores, con el abandono progresivo de la que ya era francamente minoritaria (*boina/bilbaína*). “En otros momentos puede constatarse el cambio generacional, con tendencia al triunfo de la forma estándar en el español de España, sobre todo si coincide con la preferida por la norma culta urbana: *pendiente/zarcillo*”. (Samper Padilla, 2006: 116). Igualmente, el cambio de hábitos sociales tiene su reflejo en la abundante presencia en el léxico disponible de palabras referidas a «ropa deportiva» (*playera, botín, calentador, malla, sudadera*), en consonancia con la importancia del deporte en la vida diaria de los jóvenes (Hernández Cabrera y Samper Hernández, 2002–2004: 1346).

## 5.6. Los condicionantes extralingüísticos en las comunidades bilingües

El precedente de los estudios de disponibilidad léxica en España se sitúa en una comunidad bilingüe, la ciudad de San Sebastián. Se trata de la tesis doctoral de Azurmendi Ayerbe (1983), que propuso un modelo para la descripción sociolingüística del bilingüismo en general, y del bilingüismo euskera-español en particular, a partir del estudio del léxico. Para el análisis del mismo incorpora el criterio de disponibilidad siguiendo –con modificaciones<sup>20</sup>– los trabajos canadienses de Mackey (1971a, 1971b). Los 1.224 alumnos de quinto y octavo de E.G.B. que componían la muestra respondieron a listas cerradas de 20 palabras disponiendo de 10 minutos por centro. Aunque la autora recabó el léxico disponible de quince áreas temáticas inventadas por ella misma, solo analizó cuatro de ellas: ‘Partes del cuerpo’, ‘Prendas de vestir y de calzar’, ‘Juegos –diversiones– entretenimientos, y ‘Agricultura y vegetales –ganadería y animales– pesca y peces’. Analizó la variable léxica disponible atendiendo a factores tales como ‘cantidad de producción’, ‘categoría léxica’, ‘distancia interlingüística’, ‘corrección lingüística’ y ‘divisiones semánticas’. Aparte de las variables sociolingüísticas ‘edad’, ‘sexo’, ‘lugar de residencia’, ‘tipo de centro (público/privado)’, estudia cuantitativa y cualitativamente otras cuatro variables estrechamente relacionadas con el carácter bilingüe de la comunidad de san Sebastián, que son: ‘lengua materna’, ‘lengua utilizada en casa’, ‘lengua empleada con los amigos’ y ‘lengua en la que se realiza la encuesta’.

---

<sup>20</sup> Precisamente las modificaciones metodológicas de la autora, que incluyen centros de interés diferentes a los habituales en el resto de las investigaciones y listas cerradas, le conllevaron fuertes críticas, en especial de Benítez Pérez (1991: 355), quien se queja de que con estas modificaciones “no tiene la posibilidad, siempre útil y necesaria en disponibilidad léxica, de comparar con los trabajos ya realizados o que puedan realizarse en el futuro”.

Ya en el Proyecto Panhispánico, con el correr de los años, la nómina de investigaciones ha ido cubriendo las diferentes comunidades bilingües españolas. Se cuenta así con investigaciones sobre el País Vasco (Etxebarria Arostegui, 1996, 1999), Navarra (Areta Lara, 2009; Saralegui y Taberner, 2008; Taberner Sala, 2008a, 2008b), Lérida (Serrano Zapata, 2002, 2003, 2004, 2006); Valencia (Gómez Devís, 2004, 2005; Gómez Devís y Llopis Rodrigo, 2010; Gómez Molina y Gómez Devís, 2004); Alicante (Martínez Olmos, 2007), Castellón (Blas Arroyo y Casanova, 2003a, 2003c; Casanova Ávalos, 2008), Galicia (Álvarez de la Granja y López Meirama, 2013; Blanco, 2011; López Meirama, 2008, 2011); Ceuta (Ayora Esteban, 2002, 2003, 2004, 2006; Fernández Smith, 2008), Melilla (Fernández Smith *et al.*, 2006, 2008), y Gibraltar (Escoriza Morera, 2003, 2004).

Como recogen Samper Padilla y Samper Hernández (2006: 79), de acuerdo con los criterios adoptados en la reunión de Bilbao de 1999, las investigaciones que se realizan en comunidades bilingües han de tomar en consideración para el análisis sociolingüístico, además de las variables comunes a todo el proyecto, otros tres factores específicos de este contexto, que son: la lengua materna, la lengua de uso familiar y el modelo de enseñanza (monolingüe/bilingüe). Veamos, a continuación, cómo se han reflejado estas variables en la distribución de la muestra en las comunidades estudiadas y qué resultados han arrojado.

### 5.6.1. 'Lengua materna'

Los investigadores en disponibilidad léxica han distinguido entre *lengua materna* y *lengua habitual* –este último, con un valor similar al de *lengua de uso familiar*. Dada la cercanía de estos conceptos, será necesario delimitar la aplicación de estos términos.

Según L. Dabène (1994), el término *lengua materna* se refiere, en la actualidad, a diferentes conceptos, según las circunstancias del contexto sociolingüístico o institucional al que se aplique<sup>21</sup>. Ello hace que el concepto por ella designado adopte en la bibliografía otras denominaciones, entre las que se incluyen: *lengua primera*, *lengua principal*, *lengua dominante*, *lengua preferida*, *lengua vernácula*, *lengua natural*, *lengua nativa*, *lengua de pertenencia*, etc.

---

<sup>21</sup> Según Santos Maldonado (2002: 20), la lengua materna (LM) recoge, principalmente, las ideas de:

- “la lengua de la madre”, según la etimología del calificativo *materna*;
- “la primera adquirida”, de acuerdo a su anterioridad en el orden cronológico de apropiación;
- “la lengua que mejor se conoce”, asociado a la valoración subjetiva de la competencia que un individuo posee en una lengua sobre otras lenguas también conocidas;
- “la lengua adquirida de forma natural”, es decir, mediante interacción con el entorno inmediato, sin intervención pedagógica ni reflexión lingüística consciente.

Dado que nos referimos a un contexto bilingüe o multilingüe, entendemos la *lengua materna* como lengua primera o L1 (Bailly, 1984; Giacobbe, 1990) por ser la lengua en que un individuo aprende a hablar. Esta denominación puede coincidir o no con la lengua que se usa preferentemente. De ahí que se la denomine también *lengua principal* (Siguán y Mackey, 1986), *lengua dominante* o *lengua preferida* (Postigo, 1998), lo cual supone que la lengua en la que se aprendió a hablar puede pasar a una posición secundaria frente a otra que se adquiere más tarde.

En lo que respecta ya al estudio de la disponibilidad léxica, este factor ha tenido una presencia muy dispar en las investigaciones de las comunidades donde se da el bilingüismo (ver Cuadro 5.5). Así, no deja de ser llamativo que el análisis de este factor esté ausente en las investigaciones realizadas en Valencia (Gómez Devís, 2004), Alicante (Martínez Olmos, 2007) y Galicia (López Meirama, 2008), todas ellas comunidades bilingües más castellanizadas o con un sentimiento nacionalista menos pujante. En el caso de Valencia, la autora justifica la exclusión de la variable ‘lengua materna’, al considerarla “menos relevante dada la edad de los informantes (final de la adolescencia; porque no se trata de una investigación sobre sujetos de Educación Primaria o de Secundaria Obligatoria), por los doce años de escolarización cursados y por el grado de formalidad que caracteriza una prueba o test de estas características” (Gómez Devís, 2004: 139). Las razones aducidas por la autora parecen indicar que los años de escolarización tienden a reducir el efecto de esta variable.

**Cuadro 5.5. Distribución de las muestras según el factor ‘lengua materna’ en las investigaciones realizadas en comunidades bilingües**

Comunidad	Investigación	Castellano	Propia	Bilingüe	Otra
Bilbao	Etxebarria Arostegui (1996)	100	euskera: 145	–	
Navarra	Areta Lara (2009)	352	euskera: 78		
Lérida	Serrano Zapata (2002, 2003)	66	catalán: 172	11	aranés: 4
Valencia	Gómez Devís (2004)	–	–	–	–
Alicante	Martínez Olmos (2007)	–	–	–	–
Castellón	Blas y Casanova (2003c)	104	valenciano: 106	36	
	Casanova Ávalos (2006)	128	valenciano: 159	57	
Galicia	López Meirama (2008, 2011)	–	–	–	–
Ceuta	Ayora Esteban (2003, 2006)	237	árabe: 29	–	
Melilla	Fernández Smith <i>et al.</i> (2008)	128	bereber: 9	–	francés: 2

Contrasta, sin embargo, esta postura en las provincias de Valencia y Alicante –los territorios de la Comunidad Valenciana más castellanizados– con la decisión

tomada en Castellón, provincia fronteriza con Cataluña, y con mayor presencia histórica de la lengua catalana. Aquí se hizo un estudio preliminar con una submuestra de la provincia de Castellón (106 valenciano-hablantes, 104 sujetos de lengua materna castellana y 36 bilingües) donde se demuestra que la lengua materna sí es un factor que repercute en el número de vocablos de los encuestados (Blas Arroyo y Casanova Ávalos, 2003b, 2003c). En este caso, los castellano-hablantes superan siempre a los de habla materna valenciana en el número de respuestas en español, aunque tal diferencia prácticamente desaparece cuando se examina la cantidad de vocablos aportado por los distintos grupos. Con el factor ‘lengua materna’ covaría el factor ‘sexo’, provocando un aumento de las divergencias entre las alumnas castellanohablantes y los varones de habla valenciana, y el factor ‘lugar de residencia’, que atenúa estas diferencias en la capital de provincia.

Por su parte, en Galicia, López Meirama (2008) descartó la variable ‘lengua materna’ por las características sociolingüísticas de esta comunidad de habla. Según la autora<sup>22</sup>, dado que en Galicia hay muy pocos “neofalantes” de gallego (muy pocos que hayan aprendido a hablar en español y luego pasen a emplear usualmente el gallego) y sí bastantes que han aprendido a hablar en gallego y luego se decantan por el uso del castellano, y dado que el uso generalizado (y estabilizado) del gallego está muy localizado en las poblaciones rurales, el único factor que consideró incluir fue el de la lengua habitual.

Sin embargo, en el proyecto de léxico disponible del gallego sí consideró necesario hacerlo, dado que la influencia del castellano sobre el gallego es muy importante. La lengua, en su uso familiar, está llena de castellanismos de todo tipo, y en este caso la lengua materna sí podría ser un dato significativo en relación con el léxico empleado por los informantes (más o menos “castellanizado”).

Por el contrario, la influencia del factor ‘lengua materna’ sí es analizada en el País Vasco, Navarra y Cataluña (en Lérida). Como indican Siguán y Mackey (1986), en muchas comunidades bilingües la lengua se convierte en un símbolo de identidad que, al ser utilizada o adoptada por los hablantes, adquiere un valor de integración y fidelidad al grupo considerado, así, como una marca de diferenciación frente a los demás. La dimensión política que este aspecto adquiere no es ajena a la importancia de la lengua materna en las comunidades mencionadas.

En cuanto a los resultados publicados de estas comunidades, Serrano Zapata (2004) realizó en Lérida un estudio preliminar sobre una muestra de 48 estudiantes preuniversitarios (19 con castellano como lengua materna, y 29 con catalán). La propia autora califica los resultados de llamativos, ya que “los informantes cuya lengua materna es el catalán sobrepasan en todos los campos léxicos a los castellano-hablantes con una diferencia media de 2,43 palabras por centro de interés” (Serrano Zapata, 2004: 161). Esta investigadora achaca estos resultados tan inesperados a la influencia de otras variables puesto que, en su muestra, el 62% de los informantes

---

<sup>22</sup> Agradezco muy encarecidamente a Belén López Meirama las explicaciones que sobre este punto muy amablemente me facilitó, y que aquí reproduzco.

cuya lengua materna es el catalán son del nivel social medio y el 58,6% acuden a un colegio privado. Sin embargo, al estudiar la incidencia de esta variable en la producción de interferencias, son los informantes de lengua materna catalana los que superan los que superan con creces a los castellanohablantes. Sobre 100 vocablos, los catalanohablantes producen 5,7 interferencias mientras que los castellanohablantes solo actualizan 2,7. A partir de estos datos, esta estudiosa formula la hipótesis de que los catalanohablantes “echan mano de ambas lenguas al contestar el cuestionario” y que, por consiguiente, “puede ser que posean un lexicón en el que mezclen o sumen el vocabulario castellano y el catalán” (Serrano Zapata, 2004: 162).

En todos los casos anteriores, independientemente de la lengua materna en cuestión (castellano, catalán, gallego, euskera), se trata de lenguas cooficiales en el estado español, y presentes en el sistema educativo de su comunidad. Una situación muy distinta es la que se da en las ciudades autónomas de Ceuta y Melilla, donde, junto a la mayoría hispanohablante, hay una importante minoría cuya lengua materna es el árabe o el bereber, lenguas ambas ausentes en esos territorios tanto en el sistema educativo como en la administración. Por ello, en estas dos ciudades, más que hablar de bilingüismo, hay que referirse a una situación de diglosia, es decir, una “situación de coexistencia de dos lenguas [en la que] se produce un desequilibrio a favor de una de ellas, lo que permite calificar a una de ellas de *lengua fuerte* y a la otra de *lengua débil*” (Siguán y Mackey, 1986: 45). La lengua fuerte o lengua A (español) es lengua de las funciones superiores de la vida social (escuela, administración, etc.) y la lengua débil o B (árabe-dariya en Ceuta, beber-mazigia en Melilla) es la de las funciones personales y cotidianas, vida familiar y, en general, la del lenguaje ordinario.

En Ceuta, la única lengua oficial es el español y la casi totalidad de los escolares musulmanes tienen como lengua materna un dialecto del árabe. Esto hace que en la escuela los musulmanes tengan que aprender este idioma tanto como vehículo de comunicación (oral) como de transmisión cultural (escrita, formal) y sean evaluados como si el español fuera su lengua materna. Ello “conlleva, entre otras cosas, un significativo índice de fracaso escolar en el ámbito educativo de la población ceutí de origen musulmán, que influye de alguna manera en un más bajo nivel en la totalidad de los estudiantes” (Ayora Esteban, 2006: 67–68). Este índice de fracaso escolar causado por las dificultades de los alumnos musulmanes con la lengua española, además de otros problemas de índole cultural, hace que pocos alumnos de la comunidad árabe lleguen al bachillerato. Esta fue una dificultad añadida con la que se encontró la autora, respecto a los criterios de representatividad en esta variable: lo irrelevante del número de alumnos musulmanes que cursan el nivel preuniversitario. Concretamente, en el curso 99–00, cuando hizo las encuestas en los centros públicos, de 597 alumnos matriculados en 2º de bachillerato, solo 53 (el 8,9%) eran musulmanes, y en el mes de marzo algunos de ellos lo habían abandonado. En el grupo entrevistado en el centro privado, no había ningún alumno musulmán (Ayora Esteban, 2006: 65, n. 5).

Ayora Esteban (2006: 106) descubre en su estudio que “hay bastante relación entre los informantes de lengua materna árabe con el nivel sociocultural bajo y el tipo de centro público, entrando en juego, de esta manera, tres variables sociolingüísticas”. Los informantes de lengua materna española aportan una media de 20,85 palabras por centro de interés y los de lengua materna árabe 19,17 palabras. En general hay una diferencia de dos o tres puntos por centro de interés entre los dos grupos y en ningún caso los alumnos de lengua árabe superan a los de lengua española. Por tanto, la *lengua materna* se presenta como un factor muy influyente: “el grupo de lengua materna árabe maneja un bagaje léxico notablemente menor al de lengua materna española” (Ayora Esteban, 2006: 203).

Fernández Smith *et al.* (2006: 772) presentan una situación parecida en Melilla, donde el colectivo bereber se encuentra en una “situación de bilingüismo precoz sucesivo de tendencia diglósica” que provoca igualmente altos índices de fracaso escolar, “sobre todo si estos alumnos viven en barrios donde constituyen la única población y donde, consecuentemente, la lengua hablada en su entorno es la mazigia”. De esta manera –continúan estos autores–, “las interacciones comunicativas con otros grupos sociales se reducen al ámbito escolar y a algunos encuentros extraescolares, lo que no es suficiente para garantizar la integración cultural y lingüística de este sector de la población”.

Con posterioridad, en el estudio del léxico disponible de la ciudad de Melilla, (Fernández Smith *et al.*, 2008: 144) obtuvieron resultados parecidos en relación al factor ‘lengua materna’ a los de Ayora Esteban (2006) en Ceuta. En la ciudad de Melilla, la *lengua materna* se presenta “como una variable susceptible de incidir en el dominio léxico de cada uno de sus subgrupos, [...] pues en la mayoría de los centros de interés, los hablantes de lengua materna española consiguen promedios más elevados que sus compañeros de lengua bereber”. Igualmente “existe una fuerte correlación entre la pertenencia al grupo de lengua materna bereber y una puntuación baja en el parámetro sociocultural”. Cualitativamente se presentan claras divergencias en función de la variable *lengua materna*, ya que se amplían las diferencias entre las posiciones que ocupan los centros de interés, que pueden llegar a ser de hasta seis puestos, como ‘Trabajos del campo y del jardín’, a favor de los hablantes de lengua materna-española, o ‘Partes del cuerpo’, a favor de los hablantes de bereber (Fernández Smith *et al.*, 2008: 144–145).

### 5.6.2. ‘Lengua de uso familiar’/‘Lengua habitual’

La variable ‘lengua de uso familiar’, también denominada como ‘lengua habitual’, hace referencia a la lengua predominante fuera del recinto escolar. En todos los estudios, a excepción del caso de Lérida (Serrano Zapata, 2003), la proporción de informantes que ha declarado usar el español como lengua familiar

o habitual ha sido superior a la de los hablantes de las otras lenguas del estudio (ver Cuadro 5.6). En Lérida los hablantes de catalán triplican a los de castellano, en tanto que en Valencia y País Vasco los porcentajes se invierten. En estas comunidades son los hablantes que dicen usar el español como lengua habitual los que casi triplican a los que dicen hablar habitualmente en catalán o vasco, respectivamente. En Galicia los hablantes de castellano duplican a los de gallego y en Navarra, hay cuatro hablantes de castellano por uno de euskera.

En el estudio de disponibilidad léxica de Valencia, Gómez Devís (2004: 139) se centra en el bilingüismo individual –ligado a la edad de adquisición de las lenguas–, “dado que el hablante es el punto de contacto de ambas lenguas, castellano y valenciano” en la provincia de Valencia. A la hora de considerar la variable ‘lengua habitual’, la autora establece dos subespecificaciones:

– “monolingües en castellano, aquellos informantes que tienen por lengua de uso en la familia, con los amigos y en el centro educativo, el castellano”, es decir, “castellanohablantes y bilingües pasivos en valenciano”;

– “bilingües, aquellos sujetos que tienen como lengua de uso familiar, con los amigos y en el centro escolar (patio, pasillos), mayoritariamente el valenciano o ambas”, o sea, “bilingües activos y, mayoritariamente, valenciano-hablantes” (Gómez Devís, 2004: 139).

El análisis de esta variable arroja promedios de respuestas muy similares: 20,3 palabras en los monolingües y 20,1 en los bilingües. En consecuencia, la investigadora concluye que “la variable ‘lengua habitual’ actúa como variable neutra, ya que presenta una incidencia muy escasa en las diferencias de productividad léxica de castellanohablantes y bilingües castellano-valenciano, y además, ofrece quince diferencias no significativas entre los promedios cuantitativos de las respuestas” (Gómez Devís, 2004: 143).

**Cuadro 5.6. Distribución de las muestras según el factor ‘lengua habitual’/‘lengua de uso familiar’ en las investigaciones realizadas en comunidades bilingües**

Comunidad	Investigación	Español	Propia	Bilingüe	Otra
Bilbao	Etxebarria Arostegui (1996)	125	euskera: 50	70	
Navarra	Areta Lara (2009)	355	euskera: 75		
Lérida	Serrano Zapata (2002, 2003)	56	catalán: 179	16	aranés: 2
Valencia	Gómez Devís (2004)	343		122 <sup>a</sup>	

<sup>a</sup> Hay que tener presente que el dato de hablantes bilingües en Valencia significa para Gómez Devís (2004) que mayoritariamente estos informantes usan la lengua propia: el catalán valenciano.

Los datos del estudio cualitativo de las 50 primeras palabras de los listados de los dos grupos muestra lo cercano del dominio léxico de ambos grupos, confirmando lo anticipado en el estudio cuantitativo: la variable ‘lengua habitual’ “es el factor social con mayor porcentaje de convergencia cualitativa en los cincuenta

primeros vocablos, ya que muestra el 86,35 % de coincidencia” (Gómez Devís, 2004: 244). No se reflejan, pues, “diferencias cognitivas en la actualización léxica y en las relaciones asociativas”, resultado más que probable de los doce años de escolarización, mayoritariamente, en castellano. Se notan, eso sí, más transferencias léxicas entre los bilingües, especialmente en el centro de interés ‘Trabajos del campo y del jardín’: *collir, empeltar*.

En Alicante, Martínez Olmos (2007: 84) prefiere utilizar la variable ‘comunidad de habla’, con cuatro subespecificaciones, “en función de la actuación comunicativa, dentro del ámbito familiar, en cada una de las lenguas cooficiales, catalán y castellano, de la Comunidad Valenciana”. Las cuatro áreas que distingue a partir de estudios sociolingüísticos previos, se caracterizan por diferentes grados de bilingüismo con predominio de una u otra lengua. Estas áreas, y el número de informantes que integran la muestra de Alicante, son:

- a) La Marina (Alta y *Baixa*)-El *Comtat-L'alcoià*: bilingüismo activo con predominio del valenciano (104 informantes);
- b) *L'Alacantí*-el Vinalopó: bilingüismo activo con predominio del castellano (210 informantes);
- c) El Bajo Segura, excepto Guadamar y Barba-roja: bilingüismo pasivo con dominio del castellano (127 informantes);
- d) El Alto Vinalopó (Villena y Sax): bilingüismo pasivo con dominio del castellano (19 informantes).

El promedio de respuestas por informante y centro de interés presenta una situación esperable: “se da un mayor rendimiento léxico en los estudiantes procedentes de la comunidad de habla más castellanizada de las cuatro que conforman la muestra: el Alto Vinalopó (Villena y Sax)” (Martínez Olmos, 2007: 207). Además, las diferencias más significativas las causa la confrontación de las dicotomías *l'Alacantí*-El Vinalopó vs. El Bajo Segura (excepto Guadamar y Barba-roja), y la de esta última con el Alto Vinalopó (Villena y Sax) (212), esto es, entre las zonas más castellanizadas. El estudio cualitativo posterior de los cincuenta primeros vocablos fija que “el factor ‘comunidades de habla alicantinas’ influye decisivamente en la convergencia asociativa”, puesto que solo 37 de los cincuenta primeros vocablos –los más disponibles– son comunes (Martínez Olmos, 2007: 369).

**Cuadro 5.7. Distribución de la muestra de Galicia según el factor ‘lengua habitual’ (López Meirama, 2008)**

Solo gallego	Más gallego	Bilingüe	Más español	Solo español	Otra
147	100	119	236	195	3

En Galicia, López Meirama (2008) estratifica la muestra según diferentes niveles de bilingüismo, según mayor o menor grado de preferencia por el gallego o el español como ‘lengua habitual’ (ver Cuadro 5.7). A partir de esta distribución

llega a conclusiones similares a los estudios valencianos. La ‘lengua habitual’ es un factor que no incide especialmente en la producción ni en la selección de palabras, especialmente entre los hablantes de gallego y los bilingües sin preferencia por una u otra lengua (gallego o castellano). De nuevo, el grupo castellanoparlante es el que obtiene promedios superiores (21,54 palabras los hablantes que se decantan por el castellano, 22,35 palabras de media los que dicen hablar solo en castellano). Únicamente en este último grupo se podría hablar de cierta relevancia desde el punto de vista estadístico, con respecto, eso sí, a las medias de producción léxica ofrecidas por los hablantes de gallego, como se muestra en el Cuadro 5.8<sup>23</sup>.

**Cuadro 5.8. Promedio de respuestas por informante según el factor ‘lengua habitual’ en el léxico disponible de Galicia (López Meirama, 2008)**

Centro Interés	Sólo gallego	Más gallego que castellano	Gallego o castellano	Más castellano que gallego	Sólo castellano
01. El cuerpo humano	28,04	28,16	28,93	29,78	30,51
02. La ropa	23,65	24,37	23,99	25,24	25,37
03. Partes de la casa	16,82	17,11	18,42	18,82	18,65
04. Muebles de la casa	15,70	16,27	16,62	16,67	16,79
05. Alimentos y bebidas	27,65	28,14	28,25	28,62	29,32
06. Objetos colocados en la mesa...	17,24	17,13	16,65	17,54	17,60
07. La cocina y sus utensilios	18,33	18,44	18,50	18,71	19,76
08. La escuela: muebles y materiales	24,47	23,97	24,19	24,59	25,73
09. Calefacción e iluminación	14,95	14,20	14,54	14,86	14,70
10. La ciudad	22,18	22,25	22,93	24,28	25,07
11. El campo	20,63	19,89	20,92	20,78	22,93
12. Medios de transporte	18,86	19,28	19,03	19,81	20,35
13. Trabajos del campo y del jardín	12,90	13,19	11,98	12,91	13,07
14. Los animales	28,87	28,09	28,42	29,73	32,15
15. Juegos y distracciones	17,44	17,03	17,11	18,47	18,84
16. Profesiones y oficios	22,00	20,93	21,24	22,73	24,37
17. Los colores	22,14	21,69	21,34	21,37	22,82
18. El mar	22,95	21,70	22,48	22,75	24,24
PROMEDIO	20,82	20,66	20,86	21,54	22,35

Para establecer el valor de la variable ‘lengua habitual’, López Meirama (2008) solicitó a los informantes que identificaran el idioma de uso cotidiano en

<sup>23</sup> Los datos de esta tabla han sido obtenidos a partir de la aplicación alojada en la dirección web: <http://gramatica.usc.es/disposal/>.

una de las seis categorías que aparecen en la tabla que adjuntamos. Un procedimiento similar empleó Serrano Zapata (2006) en Lérida para establecer el valor de esta variable, si bien solicitaba más ítemes de información para fijar las variantes correspondientes. En el cuestionario que los informantes contestaban incluía preguntas sobre la lengua que utilizaban los alumnos en su casa, con sus amigos y en la escuela. Al igual que el factor ‘lengua materna’, los condicionantes constaban de cuatro grupos: catalán, castellano, ambas lenguas y aranés.

Al respecto, Moreno Fernández (2012: 56) ha comentado las dificultades metodológicas que el propio contexto bilingüe ofrece, entre otros aspectos a la hora de definir la variable ‘lengua habitual’. Se trata de una variable que no es verificable por el investigador. Este estudioso comenta el caso concreto de una comunidad bilingüe hispana en EE.UU., en la cual la ‘lengua habitual’ se determina mediante tres opciones (español, inglés, o ambos). Moreno Fernández (2002: 56) opina que “en realidad se trata de hábitos que responden a factores comunicativos y contextuales de difícil simplificación”, y expresa sus reservas a la hora de abordar esta variable, ya que “la información sobre este factor refleja el juicio o la impresión del informante –recogida mediante formulario o cuestionario–, no un dato comprobado por el analista”. En contextos bilingües, por tanto, será de gran ayuda contrastar la información facilitada por los informantes con los estudios sociolingüísticos realizados previamente en esas mismas zonas por otros estudiosos.

### 5.6.3. ‘Modelo de enseñanza’ (monolingüe/bilingüe)

La especificidad de la situación sociolingüística de las comunidades autónomas donde la lengua vasca está presente –País Vasco y Navarra– ha hecho que en sus respectivos sistemas educativos coexistan diferentes modelos de enseñanza bilingüe, que faciliten el derecho a conocer y usar tanto el euskera como el castellano. De este modo, ambas lenguas constituyen asignaturas obligatorias en todos los niveles y centros educativos del País Vasco. En esta comunidad, los alumnos pueden elegir uno de los siguientes modelos educativos:

- Modelo A: la enseñanza se imparte básicamente en castellano. El euskera se imparte como asignatura.

- Modelo B: La enseñanza se imparte mitad y mitad, en euskera y en castellano, es decir, ambos idiomas son simultáneamente lenguas curriculares y vehiculares. Constituye una vía para transmitir al alumno de procedencia castellanohablante una mayor competencia en euskera.

- Modelo D: La enseñanza se imparte básicamente en euskera. El castellano constituye asignatura obligatoria. Está diseñado para alumnos de procedencia familiar vascófona.

- Modelo G: Adicionalmente, las zonas mixta y no vascófona de Navarra ofrecen, por su parte, el modelo G, que no incorpora la enseñanza en euskera ni del euskera.

La disponibilidad léxica se ha ocupado en las dos comunidades con bilingüismo euskera/castellano de estudiar la incidencia de la variable ‘Modelo de Enseñanza Bilingüe’ con el objetivo de “conocer si la producción y la disponibilidad léxica, en lengua española, varía según el modelo cursado y si resulta ser éste un factor diferenciador en la competencia y uso del español” (Etxebarria Arostegui (1999: 1484). La distribución de las muestras en las investigaciones realizadas en el área metropolitana del Gran Bilbao y en la provincia de Navarra queda como sigue:

**Cuadro 5.9. Distribución de las muestras según el factor ‘modelo educativo’**

Comunidad	Investigación	A. español lengua vehicular	D. euskera lengua vehicular	G. enseñanza en español
Bilbao	Etxebarria Arostegui (1996)	125 (51,02%)	120 (48,9%)	–
Navarra	Areta Lara (2009)	43 (10%)	141 (32,8%)	246 (57,2%)

En cuanto al modelo B, este era minoritario en el periodo (1993–1995) en el que se llevó a cabo la investigación en Bilbao, con una media de 17,7% de alumnos. Como se ve en el Cuadro 5.9, el modelo B no ha tenido representación en ninguna de las dos investigaciones. En Bilbao, Etxebarria Arostegui (1996, 1999) se concentró en la oposición Modelo A vs. Modelo D, es decir, basado en el castellano frente a uno con base euskera. Frente a la distribución equilibrada entre los dos modelos de la muestra bilbaína, casi mitad por mitad, más de la mitad de la muestra navarra (57,2%) pertenece al modelo de enseñanza exclusivamente en español, con ausencia de cualquier elemento de enseñanza bilingüe. Ello refleja el predominio del castellano en esta comunidad; según los datos del gobierno de Navarra de 2003, solo el 22% de los navarros tenía conocimientos de euskera, y la mayoría de individuos competentes son jóvenes, entre 15 y 34 años.

El estudio de la disponibilidad léxica de los escolares bilingües del área metropolitana de Bilbao de Etxebarria Arostegui (1996) muestra que las diferencias entre los resultados de los estudiantes del modelo A (principalmente en español) y D (principalmente en euskera) son pequeñas. De ello concluye la investigadora que el modelo de estudios D no parece afectar decisivamente al conocimiento, competencia y uso del español en el área del País Vasco examinada. En cualquier caso, si bien es cierto que las diferencias entre ambos grupos son pequeñas, también es verdad que los alumnos con mayor contacto con el euskera y menor con el castellano proporcionaron un promedio de respuestas (19,71) en castellano inferior en dos puntos a los alumnos con un modelo de instrucción lingüística inverso –menor contacto con la lengua vasca y mayor con el castellano (21,75).

También en Navarra, Taberero Sala (2008a) analiza las 430 encuestas que conforman la muestra de aquella comunidad, con el fin de medir la eficacia del sistema de modelos lingüísticos de enseñanza en la adquisición de la competencia

comunicativa en las dos lenguas en contacto. En un estudio que tiene carácter exploratorio, selecciona aquellos campos temáticos que parecen especialmente relacionados con los dos ámbitos fundamentales de adquisición de una lengua, el familiar y cotidiano, por una parte, con los campos ‘La cocina y sus utensilios’, ‘Trabajos del campo y del jardín’, y ‘Juegos y distracciones, y el ámbito académico, representado por los centros de interés ‘El cuerpo humano’ y ‘La escuela’.

El análisis estadístico según los modelos lingüísticos demuestra que “la actualización de términos vascos se observa en informantes de cualquiera de los modelos lingüísticos de enseñanza, A, D y G, si bien son escasas las apariciones en alumnos residentes en la zona no vascófona” (Tabernero Sala, 2008a: 561). Es este, pues, un comportamiento lógico en una zona de contacto lingüístico, que se da, sobre todo, en las zonas vascófona y mixta, independientemente del modelo de instrucción, A o D.

Las interferencias gráficas se dan entre los informantes de modelo D y zona vascófona con vasco o castellano como lengua materna y de uso habitual, quienes aportan un número representativo de palabras castellanas escritas con grafías del euskera. La investigadora “constata una correlación de los índices más bajos de disponibilidad léxica en castellano con los alumnos de modelo D, los que poseen el euskera como lengua materna y de uso habitual y los residentes en la zona de la Montaña”, “sobre todo en los centros relacionados con el ámbito doméstico y cotidiano” (Tabernero Sala, 2008a: 563). A raíz de estos datos, la autora se cuestiona a continuación, si con el programa educativo bilingüe de inmersión del modelo D, que utiliza el castellano como lengua vehicular y de aprendizaje en una sola asignatura frente al resto del currículo impartido en euskera, es posible alcanzar “la competencia léxica y, por tanto, comunicativa, de estos alumnos en español, entendiendo por esta última la producción de un discurso congruente, idiomáticamente correcto y adecuado al contexto situacional”. La autora considera necesario realizar las mismas encuestas sobre disponibilidad léxica en las dos lenguas de la comunidad bilingüe –aquí en vasco– para contar con una doble perspectiva que permita diseñar programas educativos bilingües que garanticen la competencia lingüística en las dos lenguas de forma adecuada y, sobre todo, eficaz.

Precisamente, Serrano Zapata (2003) ya había adoptado la doble perspectiva lingüística en Lérida, al iniciar un estudio comparativo sobre la disponibilidad léxica en español (con 253 informantes) y en catalán (con 255 informantes). Posteriormente, Serrano Zapata (2006), para facilitar la labor de comparación, redujo la muestra a 240 alumnos, que fueron los que realizaron la encuesta en dos lenguas. En los presupuestos de su investigación, la autora predice “en los resultados finales una influencia importante de la lengua catalana y, por lo tanto, una notable presencia de convergencias e interferencias lingüísticas” (Serrano Zapata, 2003: 920). En su estudio, la autora se plantea tener en cuenta la variable ‘modelo de enseñanza o lengua utilizada en la enseñanza’ para observar hasta qué punto la

lengua en que se reciben los conocimientos influye en el léxico disponible del alumno en ambas lenguas. Sin embargo, la situación sociológica y política de Cataluña hace que la única lengua vehicular en la educación en todos los niveles y modalidades sea el catalán, según lo establecido en la Ley 7/1983, de 18 de abril, de Normalización lingüística, y la Ley 1/1998, de 7 de enero, de Política lingüística, de la *Generalitat de Catalunya*.

El factor ‘modelo educativo’ se haya ausente también en las investigaciones realizadas en Galicia. Esto se debe a que el sistema educativo de esta comunidad otorga un tratamiento uniforme a la enseñanza lingüística. Así, en Galicia, grosso modo, hay unos porcentajes mínimos de presencia de cada una de las dos lenguas, que, por tanto, pueden variar ligeramente de unos centros a otros. Es decir, en todos los centros, fuera de las dos materias de lengua, hay algunas que se imparten en castellano y otras que se imparten en gallego, con lo que el sistema es mixto, con una ligera superioridad del castellano sobre el gallego desde la entrada en vigor del Decreto 79/2010, de 20 de mayo, para el plurilingüismo en la enseñanza no universitaria de Galicia. Por lo demás, en Galicia también se ha llevado a cabo una investigación paralela sobre la disponibilidad léxica en las dos lenguas habladas en la comunidad dentro del proyecto de *Disponibilidad léxica en Galicia*. Cada una de las dos muestras analizadas se compone de 800 encuestas realizadas a estudiantes preuniversitarios escolarizados en 50 centros de enseñanza de toda la geografía gallega. El estudio para el español fue publicado por López Meirama en 2008, hallándose el dedicado al gallego en fase de elaboración.

#### **5.6.4. Las transferencias léxicas en las comunidades bilingües**

El estudio de la disponibilidad léxica en las situaciones de bilingüismo permite cuantificar y analizar la presencia de transferencias léxicas entre las lenguas que conviven históricamente en las comunidades bilingües. Según Martínez Olmos (2007: 377), “la aparición en las pruebas de ciertas transferencias léxicas (interferencia, préstamo, cambios de código, y calcos) es garantía suficiente de su uso dentro de una determinada comunidad de habla, así como de su pertenencia a la competencia comunicativa bilingüe del informante”. Naturalmente, los resultados pueden variar en función de las características sociolingüísticas de la comunidad estudiada: tipos de bilingüismo, modelo educativo (monolingüe, bilingüe, de inmersión), la estructura social, prestigio sociolingüístico, etc.

Del estudio de las transferencias léxicas en las comunidades bilingües españolas se han ocupado, no solo las tesis doctorales de López Meirama (2008) para Galicia, Gómez Devís (2004) para Valencia, Martínez Olmos (2007) para Alicante, Serrano Zapata (2002) para Lérida, y Areta Lara (2009) para Navarra, sino también un número creciente de trabajos específicos sobre el mismo tema.

Tal es el caso de Escoriza Morera (2007) en Gibraltar, Gómez Devís (2010), Llopis Rodrigo (2005, 2008), Llopis Rodrigo y Gómez Devís (2010) en Valencia, Serrano Zapata (2006) en Lérida, y Tabernero Sala (2008b) en Navarra, trabajos todos dedicados al análisis de la presencia de vocablos pertenecientes a la lengua vernácula (catalán –con su variante valenciana–, vasco y gallego) en el léxico disponible español.

Entre los primeros intentos de abordar este aspecto, se halla la investigación realizada en Gibraltar, por un equipo de la Universidad de Cádiz, mediante encuestas de disponibilidad a 137 estudiantes preuniversitarios en inglés y a 129 en español. Escoriza Morera (2004)<sup>24</sup> lleva a cabo un análisis preliminar de los materiales acopiados, según el cual distingue dos tipos de elementos que muestran interferencias del inglés en las encuestas en español de los hablantes gibraltareños:

a) ejemplos en los que se reproducen aspectos estructurales del inglés en la palabra española (*air acondicionado, refrigeratorio*);

b) términos con alteraciones gráficas por influencia del inglés (sin consecuencias fónicas aparentes –*buffalo, elephanta*– o con ellas –*confession, receptionista*).

Por su parte, Gómez Devís (2004) dedica un apartado de su tesis al análisis de los catalanismos hallados en la disponibilidad léxica en español de los estudiantes valencianos. En su estudio, clasifica las transferencias léxicas en tres grupos: a) préstamos puntuales o esporádicos (transferencias); b) préstamos en vías de integración o ya integrados, en función del grado de adaptación lingüística y adopción social, y c) sustituciones de código, cuando el informante no recuerda el vocablo en castellano.

En la valoración global de los catalanismos registrados, la autora concluye que:

a) Se reconocen muchos préstamos culturales, es decir, vocablos en catalán para denominar objetos, partes de la casa, animales, comidas, labores del campo, juegos, etc., que son voces autóctonas, características de este entorno y sin equivalente en castellano (*garroffó, andana, porxada*). Se trata de préstamos puntuales y que en la conversación ordinaria serían cambios de código.

b) Se actualizan otros vocablos catalanes con un índice de disponibilidad superior a los correspondientes castellanos, lo cual refleja cierto desconocimiento u olvido temporal de su equivalente en castellano (*desllunat, pallisa, pastera, tramús, llar, carlota, carxofera*, etc.). Son ejemplos evidentes de transferencia léxica y de sustitución de códigos.

c) También se observa una alta proporción de dobles, con un índice de disponibilidad superior en castellano (*llaurador/labrador; fuster/carpintero*); lo que indica que varios informantes han mencionado ambas unidades léxicas, por

---

<sup>24</sup> Citado en: Samper Padilla y Samper Hernández (2006: 82).

lo que no puede ser considerado, globalmente, como casos de transferencia léxica (Gómez Devís, 2004: 336).

Por su parte, Taberero Sala (2008a: 552–557) analiza las transferencias del euskera en el léxico disponible español de los preuniversitarios navarros. Las áreas que facilitan el contacto lingüístico son, en su mayoría, las relacionadas con el ámbito doméstico y cotidiano, esto es: ‘La ropa’, ‘Partes de la casa’, ‘La comida y bebida’, ‘La escuela’, ‘El campo’, ‘Los trabajos de campo y jardín’ y ‘Los juegos y distracciones’. En el análisis de las transferencias, la autora distingue entre:

– *Code-shifting*: Considera “sustituciones de código aquellas voces que no guardan similitud alguna con la correspondiente designación castellana e incluso con cualquier otra de esta lengua y de diferente significado, lo que las hace difícilmente previsibles en un hablante sin competencia lingüística en euskera” (Taberero Sala, 2008a: 552). Son palabras producidas por hablantes de la zona vascofona y mixta de Navarra, casi todos hablantes con el euskera como lengua materna y de uso habitual, tales como: *beroki* ‘abrigo’, *chano* ‘sombrero’, *gona* ‘falda’, *sagardo* ‘sidra’, *somera* ‘viga’, *zato* ‘trapo de cocina’, *arbela* ‘pizarra’, *txirrina* ‘timbre’, *txakurrak* ‘perros’, *gorringos* ‘oronja (seta)’, *bostortza pasatu* ‘rastrillar’, *jorrar* ‘escardar’, y distintos tipos de juegos como *perrexil*, *txintxurri*, *ipur zikina*, o *xake* ‘ajedrez’.

– Interferencias gráficas, fónicas y morfológicas. Se trata de voces que, según su registro lexicográfico, pertenecen al caudal léxico del euskera, si bien su similitud casi exacta con las castellanas del mismo significado lleva a considerarlas como interferencias o préstamos gráficos y no como *code-shifting*. Son palabras producidas mayormente por informantes que cursan el modelo D de enseñanza (con el euskera como lengua vehicular), tales como: *bakero*, *kamiseta*, *motxila*, *pontxo*, *txaketa*, *txaleko*, *txandal*, *dutxa*, *entxufe*, *pizina*, *amaka*, *entsalada*, *le-txuga*, *txitxarro*, *txokolate*, *txorizo*, *txuleta*, *metxero*, *kartulina*, *kortxo*, *bizikleta*, *txipiron*, *musika*, *zine*, *fisiko*, etc.

– Préstamos culturales, es decir, cuando no existe esa palabra en español. Pueden presentarse de dos formas: a) sin adaptación, como las formas *txoko*, *kali-motxo*, *patxaran*, *txapela*, *txamarra*, *borroka*, *birika*, *marmitako*, *txakoli*, *gaztetxe*, *goiti-behera* o *goitibehera*, *pottoka*, *betizu*, *trikitixa* y *pelotari*, producidas en su mayoría por hablantes de la ciudad de Pamplona; y b) parcial o totalmente adaptados, tratándose de términos procedentes del ámbito de los usos y tradiciones como *matatxerri*, *baserris*, *sarde*, *zatas*, *larres*, *mandarra* y *chingles*. Se constatan romanceados: *chapela*, *chamarra*, *calimocho*, *chacolí*, *changurro*, *cocochas*, *layas*, *gabarra*, *goitibera*, *zulo(s)* y *chistorra*.

En Galicia, Álvarez de la Granja (2011) estudia la presencia del gallego en los listados de léxico disponible español. A tal fin, clasifica los distintos fenómenos relacionados con el contacto léxico desde una doble perspectiva (formal y funcional), ya se trate de una presencia directa (Cuadro 5.10) o indirecta (Cuadro 5.11).

**Cuadro 5.10. Presencia directa del gallego en el léxico disponible español**

Lapsus bilingües	garaxe
Cambios de código	invernadoiro

**Cuadro 5.11. Presencia indirecta del gallego en el léxico disponible español**

	Formas sin adaptar		Formas adaptadas	Calcos semánticos	Calcos morfoléxicos
	Integradas lingüísticamente	No integradas lingüísticamente			
Prestamos integrados	<i>silva</i>	<i>xouba</i>	<i>jurelo</i>	<i>vertedero</i> (‘pila para lavar’)	<i>rellamante</i>
Interferencias léxicas			<i>houciño</i>	<i>garfio</i> (‘tenedor’)	<i>lavalouza</i> (¿?)

Esta clasificación le sirve para analizar aquellas palabras proporcionadas por los informantes en sus listados que se corresponden con formas puramente gallegas o formas influidas por esta lengua. Así, desde el punto de vista formal, se da cuenta de una serie de fenómenos de contacto relativos al grado de adaptación de las formas a la ortografía, fonética o morfología del gallego (*cortejo-cortello*, *houciño-fouciño*, *mugir-mujir*), la presencia de calcos semánticos (*vertedero* ‘pila destinada a lavar la vajilla situada normalmente en la cocina’) y morfoléxicos (*lavalouza* por *lavavajillas*) y el tipo de integración lingüística que todos ellos tienen en el español de Galicia. Por otra parte, desde el punto de vista funcional, el estudio muestra los factores que motivan la aparición de ciertos fenómenos de contacto lingüístico en el léxico disponible: desde los préstamos del gallego integrados en el castellano (*pota*, *sachar*, *sacho*, *faiado*, *rapante*, *grelo*, *mariola*, *jurelo*, *gamela*, *chapapote*, etc.) hasta los casos de interferencias léxicas (*castiñero*, *cortejo*, *houciño*, *matoguera*, *paso de peones*, *peixera*, etc.), además de lapsus bilingües y cambios de código, propios de las comunidades con más de una lengua oficial.

En el tercer capítulo de este mismo volumen editado por López Meirama (2011), José A. Bartol Hernández presenta un análisis muy abarcador de las interferencias sobre el español en los léxicos disponibles de las zonas bilingües de España, que incluye los datos de todos los trabajos realizados hasta la fecha al respecto. Así, se sirve de los trabajos de: López Meirama (2008) para Galicia, Gómez Molina y Gómez Devís (2004) para Valencia, Martínez Olmos (2007) para Alicante, Serrano Zapata (2006) para Lérida, y Marta Areta (2009) para Navarra. Además aporta datos novedosos sobre la situación en dos provincias del País Vasco, Vizcaya y Álava. El trabajo se marca tres objetivos: 1) establecer los campos léxicos más propicios a la aparición de estos vocablos, 2) ver las diferencias entre

la situación sociolingüística de cada zona, y 3) valorar el grado de integración de las voces en la comunidad de habla.

Metodológicamente, Bartol Hernández (2011) toma como objeto de estudio las palabras cuya presencia en los listados de disponibilidad léxica sea explicable preferentemente por el contacto *actual* de lenguas, por lo que se ha venido llamando *interferencias, transferencias, cambios de código* (cfr. Gómez Molina y Gómez Devís, 2004). Por lo tanto, el autor deja de lado el análisis de los *calcos semánticos* (cfr. Serrano Zapata, 2006), “ya que las respuestas de los informantes de disponibilidad léxica sólo nos permiten intuir el significado con el que se actualizan las palabras”.

A partir del análisis del índice de disponibilidad y la frecuencia de aparición del léxico no español en los listados, muestra grandes diferencias en la densidad de los préstamos en las distintas zonas, siendo las comunidades que presentan un mayor porcentaje de palabras en lengua vernácula Lérida (9,21%), y Galicia (2,88%), apareciendo, a continuación, con valores semejantes, Valencia (2,3%) y Vizcaya (2,35%). Las otras tres zonas –Álava, Alicante y Navarra– no superan el 1% de media.

En cuanto a los centros de interés, aunque con diferencias entre unas y otras comunidades, parece que hay algunos que favorecen más que otros la presencia de elementos de la otra lengua. En este caso, predominan considerablemente en ‘El campo’ (sobre todo, en el léxico gallego de Galicia), ‘Trabajos del campo y del jardín’, ‘Los animales’ y, en menor medida, las ‘Las partes de la casa’ (especialmente en las comunidades de habla vasca).

Bartol Hernández (2011) analiza también el grado de integración de los préstamos léxicos. Para ello, toma como valor de referencia el índice de disponibilidad (ID) 0,01. En los estudios de disponibilidad léxica, las palabras que igualen o superen este valor se consideran palabras altamente utilizables en esa comunidad. En el caso de los préstamos, indica que el vocablo está completamente integrado. De este análisis resulta la siguiente distribución:

– Galicia. De los 555 vocablos que presentan algún tipo de interferencia del gallego sólo 19 (3,42%) obtienen un ID igual o superior a 0,01. Son: *pucho, faiado, rocho, cociña, grelo, rapante, pota*<sup>25</sup>, *plato fondo, sacho*<sup>26</sup>, *carballo, silva, legón, chimpín, sachar, gandería, mariola, trompicallo*.

– Valencia. De los 169 catalanismos sólo seis (3,55%) alcanzan un ID igual o superior a 0,01. Se trata de: *desllunat, palmito, empeltar, collir, truc, sambori*.

– Alicante. De los 58 catalanismos que aparecen en las listas del léxico disponible, solo 5 superan el 0,01. Son: *carlota, coca, ensulfatar, espionar, empeltar*.

– Vizcaya. Descontando las palabras mencionadas por un solo informante, de los 112 vasquismos documentados, 22 vocablos se hallan por encima del ID

<sup>25</sup> Mencionada en dos campos: ‘Objetos en la mesa para la comida’ y ‘La cocina y sus utensilios’.

<sup>26</sup> Mencionada en dos campos: ‘El campo’ y ‘Trabajos del campo y del jardín’.

0,01. Son: *tirrin, txapela, txoko, patxaran, marmitako, sue, euskera, borragoma, maisu, liburu, ikastola, Bilbo, Gazteiz, Donosti, Palacio Euskalduna, gaztete, baserri, baserritarra*<sup>27</sup>, *luxe, maldak, berango, goitibehera, pottoka, hegazkin, aizkolari*<sup>28</sup>, *artzaina, abeltzantza, zegan egin, txala, txintxu, ertzaina, etarra*.

– Navarra. De los 52 vocablos con interferencia del vasco, sólo cuatro alcanzan el índice del 0,01 (7,4%). Son: *txoko, patxaran, borragoma, gaztete*.

Finaliza su estudio Bartol Hernández (2011), comparando la disponibilidad de las palabras anteriores con la del sinónimo correspondiente en español (palabras vernáculas – palabras castellanas). Si se descuentan los préstamos culturales –para los que no existen palabras en español–, las palabras en lengua vernácula que superan a su sinónimo español son las que aparecen en el Cuadro 5.12.

**Cuadro 5.12. Palabras en lengua vernácula con mayor ID que su sinónimo español**

Comunidad	Palabra vernácula	Palabra castellana
Galicia	<i>grela rapante pota sacho legón sachar mariola faiado</i>	→ <i>nabiza, nabo</i> → <i>gallo</i> → <i>olla</i> → <i>azada</i> → <i>azadón</i> → <i>cavar</i> → <i>rayuela</i> → <i>sobrado</i> . La superan <i>ático</i> y <i>desván</i> .
Valencia	<i>desllumat</i>	→ <i>patio de luces</i>
Alicante	<i>coca</i>	→ <i>torta</i>
Vizcaya	<i>maisú aizkolari abeltzantza txoko txala txintxu</i>	→ <i>maestro</i> → <i>cortador de árboles</i> → <i>ganadería</i> → <i>rincón</i> → <i>ternera</i> → <i>rayuela</i>
Navarra	<i>txoko</i>	→ <i>rincón</i>

Como se ve, el trabajo de Bartol Hernández (2011) es muy completo y complejo. No obstante, quedan pendientes algunas cuestiones, que solo es posible tratar en un estudio detallado de cada comunidad bilingüe, como por ejemplo, el análisis de los datos en función de la lengua habitual de los informantes (monolingües en una u otra lengua o bilingües), y la influencia de otros factores sociales y cognitivos al evocar una palabra en lengua vernácula.

Valga como ejemplo de este tipo de análisis, el realizado por Martínez Olmos (2007) en las comunidades de habla alicantinas. El análisis cuantitativo y cualitativo de las transferencias léxicas del catalán en la disponibilidad léxica muestra que:

<sup>27</sup> Mencionada en dos campos: ‘El campo’ y ‘Trabajos del campo y del jardín’.

<sup>28</sup> Mencionada en dos campos: ‘Trabajos del campo y del jardín’ y ‘Profesiones y oficios’.

- Las transferencias léxicas son más comunes en las mujeres que en los hombres;
- Los alumnos asistentes a centros públicos actualizan más catalanismos que los alumnos de centros privados;
- Las transferencias aparecen preferentemente en centros ubicados en comunidades de habla con un bilingüismo activo, bien con predominio del valenciano (Callosa d'En Sarriá y Castalla), bien del castellano (Alicante, Elche, Guardamar y Novelda);
- El nivel sociocultural más permeable a la inclusión de transferencias es el medio-medio.

Estos hechos evidenciados por la investigación verifican completamente la hipótesis de trabajo de la autora: “a mayor competencia comunicativa bilingüe en la comunidad de habla, mayor transferencia lingüística del vernáculo, sobre todo en aquellos centros de interés más próximos al ámbito de uso familiar” (Martínez Olmos, 2007: 407). La situación de bilingüismo social sería así la que favorece que un informante de una comunidad de habla menos castellanizada acuda a su vernáculo valenciano durante la prueba, bien porque recurre a su propio vernáculo, o bien para conseguir una mayor actualización de términos dentro de un mismo centro de interés. Ambos casos suponen un cambio de código (404).

## VI. EL ANÁLISIS CUALITATIVO DEL LÉXICO DISPONIBLE

### 6.1. El análisis cualitativo: fines y posibilidades de estudio

Los objetivos y utilidades del léxico disponible no se limitan al análisis cuantitativo de los resultados por medio de datos estadísticos. El lenguaje no es solo número y estadística. El análisis descriptivo del léxico recogido en cada centro de interés permite observar diversos aspectos del comportamiento léxico de la comunidad de habla investigada. Gómez Devís (2004: 181) destaca las siguientes posibilidades de análisis cualitativo:

a) Las deficiencias lingüísticas de los escolares, tanto léxicas (vulgarismos, disfemismos, registros, etc.) como ortográficas (acentuación, grafías, adaptaciones, etc.), cuestión que ya ha sido investigada por diversos autores (Paredes García, 1999; Galloso Camacho, 2002a, 2008; Ávila Muñoz, 2007; Saura Ramí, 2008; Blanco, 2011; Ríos González, 2011), que se puede comprobar de forma explícita al aplicar los criterios de estandarización para los inventarios léxicos;

b) Las transferencias léxicas presentes en distintas áreas temáticas: cantidad y procedencia de extranjerismos, adaptación lingüística como préstamos, etc. (Benítez Pérez, 1993; Alba, 1995c, 1999, 2000a; Bartol Hernández, 1998, 2005; López Morales, 1999b; Paredes García, 2001b; Carrera y Bradley, 2004; Bellón Fernández, 2005; Casanova Ávalos, 2008; Prado Aragonés, 2008a; Samper Padilla y Samper Hernández, 2009; Samper Hernández, 2012);

c) Los diferentes procesos de lexificación: composición, derivación, sinonimia, antonimia, etc., es decir, descubrir los campos asociativos de las palabras; sustantivos abstractos, marcas comerciales, siglas, profesiones en femenino, etc. (Galloso Camacho, 1997, 2002b; Valencia, 1998b, 2000; Valencia y Echeverría, 1998; Fernández Smith *et al.*, 2008; Moliné Justé, 2008; Sánchez-Saus, 2011);

d) Los léxicos específicos, dialectalismos y palabras patrimoniales (Briz Gómez, 1985; Mateo García, 1998; González Martínez, 1999; Samper Padilla, 1999, 2000; Echenique, 2000; Hernández Cabrera y Samper Padilla, 2003, 2006, 2007;

Llorente Pinto, 2005; Samper Hernández, 2005b, 2005c, 2008; Hernández Muñoz, 2006, 2009; Arnal Purroy, 2008b, 2008c; Prado Aragonés, 2009);

e) El léxico como diagnóstico del desarrollo etnolingüístico: influencia del entorno, evolución cultural, etc. (Carcedo González, 2000b, 2000c; Ávila Muñoz, 2007; Enguita Utrilla, 2008);

f) El estudio de las categorías gramaticales así como de los grupos léxicos: nominales (sust.+adj., sust.+sust., sust.+prep.+sust., etc.), verbales, etc. (López Chávez, 1994; Mateo García, 1998; Carcedo González, 2000a; González Martínez, 2002; Sánchez Gómez, 2005; Fernández Smith *et al.*, 2008);

g) Averiguar el grado de convergencia conceptual –convergencia cualitativa– en los repertorios léxicos de los diferentes subgrupos sociológicos: estudio de los prototipos, redes de evocación, diagnóstico psicolingüístico, etc. (Cañizal Arévalo, 1991; Romero Rubilar, 2000; Borrego Nieto, 2002b; Hernández Muñoz, 2006; Hernández Muñoz e Izura, 2010; Gómez Devís, 2004; Paredes García, 2005; Ríos González, 2007; Manjón-Cabeza, 2008a, 2008b, 2009; Gómez Devís y Llopis Rodríguez, 2010).

Con todo, aparte de las posibilidades de estudio que menciona Gómez Devís (2004), “los análisis cualitativos más sugerentes han sido los que resultaban del establecimiento de comparaciones de distintos listados, tanto dentro de una misma muestra –lo que llamaremos comparaciones intramuestrales– como entre varias muestras –comparaciones intermuestrales”– (Bellón Fernández, 2011: 153). Por medio de estas comparaciones se pueden advertir diferencias geográficas y sociológicas en las listas obtenidas, así como descubrir el vocabulario usual y común entre dos o más grupos o comunidades.

Los análisis cualitativos intramuestrales, con diferente finalidad, suelen acompañar a la mayoría de las investigaciones. Entre ellos se cuentan los citados más arriba, a los que hay que sumar aquellos que adoptan una perspectiva sociolingüística al comparar el léxico disponible de los grupos sociales que integran la muestra en función del sexo, el nivel sociocultural, el factor geográfico, etc. Este es el caso de estudios como los de Mateo García (1998), Carcedo González (2001), Gómez Devís (2004), Hernández Muñoz (2004), Martínez Olmos (2007), Arnal Purroy (2008a), López Meirama (2011), Trigo Ibáñez (2011), etc.

Por su parte, los análisis cualitativos intermuestrales a partir del cotejo de los listados de léxico disponible de diferentes sintopías han servido para certificar, de un lado, la amplia base léxica común a las diferentes comunidades de habla, al tiempo que ha puesto de manifiesto el vocabulario diferente en virtud de razones dialectológicas y etnolingüísticas. Ejemplo de tales cotejos son los trabajos de López Chávez (1992, 1995b), Alba (1998, 2000b), Carcedo González (1999b, 2000a, 2003), Samper Padilla (1999), Hernández Cabrera y Samper Padilla (2007), Samper Padilla y Hernández Cabrera (2003), etc.

El análisis cualitativo puede ofrecer una visión fidedigna de la realidad, sobre todo, si se realiza a partir de las expresiones tal cual fueron anotadas por los

informantes, es decir, antes de serles aplicados los criterios de edición. Ello es especialmente necesario en el estudio sobre los errores ortográficos o el de las marcas comerciales. Sin embargo, el acceso a los materiales en bruto de las investigaciones no siempre es factible, amén de que su manejo puede ser engorroso.

Por ello, dado el alto nivel de homogeneidad en los criterios metodológicos del Proyecto Panhispánico, los diccionarios de léxico disponible resultantes de las distintas investigaciones garantizan una descripción efectiva del léxico disponible de la comunidad de habla. Así, los diccionarios de léxico disponible aportan la información necesaria, gracias al índice de disponibilidad léxica, para emprender análisis cualitativos con distintos propósitos lingüísticos.

## 6.2. La composición de los listados de léxico disponible

Los diccionarios de léxico disponible de cada comunidad, resultantes de las distintas investigaciones, se presentan por medio de listados, producidos por un programa informático. El programa *Dispolex*, el más usado en la actualidad, dada su capacidad para procesar grandes cantidades de datos, y su formato moderno, de fácil y rápida consulta y gestión, ordena las listas en columnas, que proporcionan sucesivamente información sobre los siguientes parámetros:

– *Número de orden*: El orden de aparición de los listados es el del índice de disponibilidad de más a menos (aunque también se pueden obtener listados por orden alfabético o por cada una de las otras medidas: frecuencia relativa y aparición).

– *Palabra*: Entendida como vocablo, es decir, cada una de las palabras diferentes aparecidas en la investigación. Se presentan como entradas de diccionario, lematizadas según los criterios de edición de la investigación.

– *Índice de disponibilidad*: Es la medida de la disponibilidad de un vocablo. Se calcula mediante una fórmula completa, ideada por López Chávez y Strassburger en 1991, en la que se tienen en cuenta tanto la frecuencia de aparición de la palabra en las encuestas como su lugar de aparición. El resultado es una medida entre 0 y 1. Sería 1 en el caso hipotético en que todos los informantes señalaran la misma palabra y la señalaran en el primer lugar.

– *Frecuencia relativa*: Se trata de la frecuencia en % de aparición de ese vocablo en el conjunto de respuestas dadas por los informantes en un determinado centro de interés.

– *Aparición en %*: Indica el porcentaje de informantes que mencionó el vocablo.

– *Frecuencia acumulada*: Es la suma de la frecuencia de un vocablo más las frecuencias de los vocablos anteriores de la lista. Es una medida que resulta muy interesante para las comparaciones entre conjuntos léxicos.

Aparte de la pertinencia de cada uno de los parámetros léxico-estadísticos aportados por los listados, qué duda cabe que el valor más genuino de los estudios de léxico disponible es el índice de disponibilidad léxica, una medida que une los criterios de frecuencia y orden de las respuestas producidas en el conjunto de encuestas realizadas. Este valor explica que, frente a una situación concreta, aquellas palabras que primero acuden a nuestra memoria son más disponibles en relación con dicha situación, que aquellas otras que no hacen su aparición de inmediato; las primeras tendrán un índice de disponibilidad mayor y aparecerán antes en los listados de léxico disponible (López Morales, 2009: 33).

En los próximos apartados analizamos el contenido de los listados de léxico disponible, tomando como ejemplo el listado íntegro de léxico disponible del centro de interés 02 ‘La ropa’, de la provincia de Córdoba (Cuadro 6.1), de Bellón Fernández (2011). El centro de interés ‘La ropa’, de acuerdo con la clasificación de Hernández Muñoz (2006: 134–156), es una “categoría natural”, en la que sus miembros responden a las características clásicas de una categoría semántica, esto es, posee límites más o menos definidos relacionados con características físicas o abstractas que nacen de la experiencia del mundo, y con identificación sencilla de los elementos que la constituyen. Es también una categoría “inclusiva”, ya que para todos sus miembros es posible enunciar una proposición lógica del tipo “x es Y” (por ejemplo, “el pantalón es ROPA”).

El listado íntegro del léxico disponible de ‘La ropa’ en Córdoba es el siguiente:

**Cuadro 6.1. Centro de interés 02 ‘La ropa’, Léxico disponible de Córdoba (Bellón Fernández, 2011)**

Nº	Vocablo	Disponibilidad	Frecuencia relativa %	% Aparición	Frecuencia acumulada
1	2	3	4	5	6
1	pantalón	0,74135	4,134	91,250	04,134
2	camisa	0,71529	4,157	91,750	08,291
3	camiseta	0,59471	3,976	87,750	12,266
4	jersey	0,56697	3,613	79,750	15,879
5	calcetín	0,50399	3,930	86,750	19,810
6	chaqueta	0,41606	3,319	73,250	23,128
7	falda	0,40909	2,900	64,000	26,028
8	zapato	0,35956	2,911	64,250	28,939
9	calzoncillo	0,34886	3,092	68,250	32,031
10	braga	0,34012	3,194	70,500	35,225
11	abrigo	0,29722	2,639	58,250	37,864
12	sujetador	0,26036	2,730	60,250	40,593

1	2	3	4	5	6
13	media	0,25701	2,594	57,250	43,187
14	guante	0,24110	2,367	52,250	45,554
15	bufanda	0,23024	2,209	48,750	47,763
16	chándal	0,22426	2,095	46,250	49,858
17	sudadera	0,22205	2,005	44,250	51,863
18	vestido	0,20633	1,665	36,750	53,528
19	zapatilla	0,16972	1,552	34,250	55,080
20	chaquetón	0,16530	1,597	35,250	56,677
21	(pantalón) vaquero	0,16529	1,201	26,500	57,877
22	chaleco	0,15815	1,303	28,750	59,180
23	rebeca	0,15103	1,325	29,250	60,505
24	sombrero	0,13970	1,506	33,250	62,012
25	corbata	0,13817	1,393	30,750	63,405
26	blusa	0,13680	1,019	22,500	64,424
27	bota	0,13133	1,359	30,000	65,783
28	gorro	0,12759	1,303	28,750	67,086
29	cazadora	0,09869	1,008	22,250	68,094
30	pijama	0,09194	1,121	24,750	69,215
31	malla	0,07406	0,872	19,250	70,087
32	polo	0,07390	0,725	16,000	70,812
33	cinturón	0,07338	0,838	18,500	71,650
34	bermudas	0,07286	0,668	14,750	72,318
35	pañuelo	0,07204	0,827	18,250	73,145
36	gorra	0,06950	0,782	17,250	73,927
37	bañador	0,06901	0,895	19,750	74,822
38	anorak	0,06780	0,702	15,500	75,524
39	body	0,06234	0,770	17,000	76,294
40	gabardina	0,06027	0,702	15,500	76,996
41	camisón	0,05960	0,725	16,000	77,721
42	pantalón corto	0,05647	0,578	12,750	78,299
43	camiseta interior	0,05511	0,612	13,500	78,910
44	chubasquero	0,05309	0,691	15,250	79,601
45	traje	0,05259	0,521	11,500	80,122
46	leotardo	0,04926	0,555	12,250	80,677
47	panty	0,04751	0,544	12,000	81,221
48	impermeable	0,04597	0,578	12,750	81,799
49	short	0,04180	0,453	10,000	82,252

Cuadro 6.1 (cont.)

1	2	3	4	5	6
50	pajarita	0,03991	0,510	11,250	82,761
51	suéter	0,03837	0,317	7,000	83,078
52	slip	0,03643	0,351	7,750	83,430
53	biquini	0,03397	0,487	10,750	83,917
54	meypa	0,03305	0,317	7,000	84,234
55	orejera	0,03145	0,351	7,750	84,585
56	pluma	0,03018	0,306	6,750	84,891
57	bata	0,02901	0,430	9,500	85,321
58	minifalda	0,02874	0,283	6,250	85,604
59	saquito	0,02854	0,181	4,000	85,785
60	top	0,02805	0,328	7,250	86,114
61	calzón	0,02693	0,238	5,250	86,352
62	mono	0,02657	0,340	7,500	86,692
63	calentador	0,02581	0,374	8,250	87,065
64	chancla	0,02381	0,306	6,750	87,371
65	peto	0,02380	0,306	6,750	87,677
66	pichi	0,02300	0,238	5,250	87,915
67	ropa interior	0,02292	0,261	5,750	88,175
68	liguero	0,02110	0,238	5,250	88,413
69	albornoz	0,02072	0,294	6,500	88,708
70	liga	0,02050	0,294	6,500	89,002
71	faja	0,02016	0,351	7,750	89,353
72	pantalón de pinzas	0,01854	0,170	3,750	89,523
73	deportiva	0,01790	0,215	4,750	89,738
74	correa	0,01783	0,227	5,000	89,965
75	esmoquin	0,01760	0,215	4,750	90,180
76	parka	0,01682	0,159	3,500	90,339
77	fular	0,01660	0,215	4,750	90,554
78	sandalia	0,01598	0,215	4,750	90,769
79	corpiño	0,01542	0,159	3,500	90,928
80	boina	0,01424	0,249	5,500	91,177
81	cordón	0,01330	0,102	2,250	91,279
82	pasamontañas	0,01326	0,204	4,500	91,483
83	pamela	0,01293	0,181	4,000	91,664
84	zapatilla de deporte	0,01269	0,147	3,250	91,811
85	tanga	0,01253	0,170	3,750	91,981

1	2	3	4	5	6
86	chal	0,01152	0,091	2,000	92,072
87	bolso	0,01139	0,102	2,250	92,174
88	reloj	0,01110	0,113	2,500	92,287
89	falda corta	0,01109	0,079	1,750	92,366
90	plumón	0,01102	0,136	3,000	92,502
91	traje de chaqueta	0,01094	0,113	2,500	92,615
92	maillot	0,01082	0,147	3,250	92,762
93	falda larga	0,01065	0,091	2,000	92,853
94	americana	0,01045	0,125	2,750	92,978
95	tejano	0,01042	0,068	1,500	93,046
96	jeans	0,00984	0,079	1,750	93,125
97	combinación	0,00948	0,125	2,750	93,250
98	levita	0,00933	0,091	2,000	93,340
99	niqui	0,00916	0,079	1,750	93,419
100	(bota) katuska	0,00881	0,079	1,750	93,499
101	capa	0,00880	0,113	2,500	93,612
102	botín	0,00842	0,079	1,750	93,691
103	sostén	0,00836	0,102	2,250	93,793
104	mocasín	0,00829	0,091	2,000	93,884
105	chupa	0,00821	0,068	1,500	93,952
106	falda pantalón	0,00809	0,079	1,750	94,031
107	tirante	0,00796	0,147	3,250	94,178
108	de deporte	0,00783	0,079	1,750	94,258
109	camiseta de tirantes	0,00776	0,079	1,750	94,337
110	calzona	0,00741	0,079	1,750	94,416
111	blusón	0,00700	0,068	1,500	94,484
112	culote	0,00698	0,068	1,500	94,552
113	gemelo	0,00686	0,068	1,500	94,620
114	pantalón largo	0,00668	0,034	0,750	94,654
115	bomber	0,00631	0,079	1,750	94,733
116	chaqué	0,00626	0,079	1,750	94,813
117	frac	0,00588	0,091	2,000	94,903
118	manopla	0,00568	0,079	1,750	94,982
119	manga	0,00559	0,034	0,750	95,016
120	botón	0,00551	0,045	1,000	95,062
121	collar	0,00550	0,057	1,250	95,118
122	(abrigo de) tres cuartos	0,00537	0,057	1,250	95,175

Cuadro 6.1 (cont.)

1	2	3	4	5	6
123	bamba	0,00529	0,045	1,000	95,220
124	pendiente	0,00526	0,034	0,750	95,254
125	pantalón de vestir	0,00516	0,034	0,750	95,288
126	algodón	0,00505	0,045	1,000	95,334
127	babucha	0,00489	0,045	1,000	95,379
128	cremallera	0,00478	0,034	0,750	95,413
129	camisa interior	0,00469	0,045	1,000	95,458
130	ejecutivo	0,00468	0,068	1,500	95,526
131	pantalón de pana	0,00461	0,034	0,750	95,560
132	trenca	0,00460	0,068	1,500	95,628
133	camiseta corta	0,00457	0,045	1,000	95,673
134	castellano	0,00455	0,045	1,000	95,719
135	minishort	0,00453	0,034	0,750	95,753
136	mantón	0,00436	0,068	1,500	95,821
137	calcetina	0,00435	0,034	0,750	95,855
138	visera	0,00434	0,057	1,250	95,911
139	zapatilla de casa	0,00431	0,045	1,000	95,957
140	hombreira	0,00430	0,057	1,250	96,013
141	camiseta larga	0,00420	0,023	0,500	96,036
142	calceta	0,00415	0,068	1,500	96,104
143	pañuelo de cuello	0,00407	0,045	1,000	96,149
144	quimono	0,00406	0,068	1,500	96,217
145	poncho	0,00405	0,034	0,750	96,251
146	muñequera	0,00404	0,057	1,250	96,308
147	tienda	0,00381	0,023	0,500	96,330
148	calzado	0,00363	0,023	0,500	96,353
149	traje largo	0,00362	0,034	0,750	96,387
150	rodillera	0,00361	0,057	1,250	96,444
151	pulsera	0,00341	0,034	0,750	96,478
152	corsé	0,00337	0,045	1,000	96,523
153	velo	0,00333	0,034	0,750	96,557
154	(forro) polar	0,00328	0,045	1,000	96,602
155	tacón	0,00327	0,045	1,000	96,647
156	lazo	0,00317	0,045	1,000	96,693
157	pana	0,00315	0,023	0,500	96,715
158	cuello	0,00314	0,023	0,500	96,738

1	2	3	4	5	6
159	pareo	0,00311	0,045	1,000	96,783
160	torera	0,00309	0,034	0,750	96,817
161	sobrecamisa	0,00308	0,045	1,000	96,863
162	camisa corta	0,00305	0,034	0,750	96,897
163	chaqueta vaquera	0,00304	0,057	1,250	96,953
164	zapatilla deportiva	0,00301	0,023	0,500	96,976
165	gayumbo	0,00301	0,023	0,500	96,999
166	tela	0,00301	0,023	0,500	97,021
167	bota de agua	0,00301	0,034	0,750	97,055
168	pantalón clásico	0,00300	0,023	0,500	97,078
169	lencería	0,00298	0,034	0,750	97,112
170	camisola	0,00295	0,045	1,000	97,157
171	pelliza	0,00287	0,034	0,750	97,191
172	gafas	0,00287	0,045	1,000	97,236
173	lana	0,00284	0,023	0,500	97,259
174	jersey de cuello alto	0,00275	0,034	0,750	97,293
175	camisa de tirantes	0,00261	0,023	0,500	97,316
176	vestido largo	0,00254	0,023	0,500	97,338
177	moda	0,00250	0,011	0,250	97,350
178	armario	0,00250	0,011	0,250	97,361
179	mono de trabajo	0,00246	0,023	0,500	97,384
180	camisa de manga larga	0,00246	0,034	0,750	97,418
181	husky	0,00244	0,045	1,000	97,463
182	bota de fútbol	0,00242	0,023	0,500	97,486
183	felpa	0,00234	0,034	0,750	97,520
184	cuello de la camisa	0,00233	0,011	0,250	97,531
185	bolsillo	0,00233	0,011	0,250	97,542
186	muestrario	0,00233	0,011	0,250	97,554
187	perchero	0,00233	0,011	0,250	97,565
188	jersey de pico	0,00232	0,023	0,500	97,587
189	delantal	0,00219	0,034	0,750	97,621
190	camisa larga	0,00218	0,023	0,500	97,644
191	túnica	0,00217	0,034	0,750	97,678
192	pantalón liso	0,00217	0,011	0,250	97,689
193	bota de piel	0,00217	0,011	0,250	97,701
194	bombín	0,00214	0,023	0,500	97,723
195	zarcillo	0,00206	0,023	0,500	97,746

Cuadro 6.1 (cont.)

1	2	3	4	5	6
196	chistera	0,00202	0,011	0,250	97,757
197	imperdible	0,00202	0,011	0,250	97,769
198	picardías	0,00201	0,034	0,750	97,803
199	toga	0,00197	0,023	0,500	97,825
200	corchete	0,00193	0,023	0,500	97,848
201	casaca	0,00190	0,023	0,500	97,871
202	textura	0,00188	0,011	0,250	97,882
203	ligero	0,00188	0,011	0,250	97,893
204	aguja	0,00188	0,011	0,250	97,905
205	náutico	0,00187	0,023	0,500	97,927
206	cinta	0,00179	0,023	0,500	97,950
207	gamberro	0,00177	0,023	0,500	97,973
208	abrigo de piel	0,00175	0,034	0,750	98,007
209	calidad	0,00175	0,011	0,250	98,018
210	braslip	0,00175	0,011	0,250	98,029
211	pesado	0,00175	0,011	0,250	98,041
212	babero	0,00175	0,011	0,250	98,052
213	puño	0,00175	0,011	0,250	98,063
214	hilo	0,00175	0,011	0,250	98,075
215	pantalón de deporte	0,00175	0,034	0,750	98,109
216	toalla	0,00173	0,023	0,500	98,131
217	calza	0,00172	0,023	0,500	98,154
218	deportivo	0,00165	0,023	0,500	98,176
219	kilt	0,00162	0,011	0,250	98,188
220	braga pañal	0,00162	0,011	0,250	98,199
221	broche	0,00162	0,011	0,250	98,210
222	tutú	0,00162	0,011	0,250	98,222
223	cárdigan	0,00158	0,023	0,500	98,244
224	cazadora vaquera	0,00155	0,023	0,500	98,267
225	chapela	0,00153	0,023	0,500	98,290
226	poliéster	0,00151	0,011	0,250	98,301
227	minipull	0,00151	0,011	0,250	98,312
228	tul	0,00151	0,011	0,250	98,324
229	camiseta de verano	0,00151	0,011	0,250	98,335
230	pelele	0,00151	0,011	0,250	98,346
231	malla corta	0,00151	0,011	0,250	98,358

1	2	3	4	5	6
232	colgante	0,00151	0,011	0,250	98,369
233	escote	0,00151	0,011	0,250	98,380
234	camisa vaquera	0,00144	0,023	0,500	98,403
235	traje de noche	0,00143	0,023	0,500	98,426
236	jubón	0,00143	0,023	0,500	98,448
237	vestido de fiesta	0,00141	0,011	0,250	98,460
238	tenis	0,00141	0,011	0,250	98,471
239	sisa	0,00141	0,011	0,250	98,482
240	minimedia	0,00141	0,011	0,250	98,494
241	sintético	0,00141	0,011	0,250	98,505
242	pantalón de pitillo	0,00136	0,023	0,500	98,528
243	fajín	0,00132	0,023	0,500	98,550
244	loneta	0,00131	0,011	0,250	98,562
245	lycra	0,00131	0,011	0,250	98,573
246	calzoncillo largo	0,00131	0,011	0,250	98,584
247	bajo	0,00131	0,011	0,250	98,596
248	visón	0,00131	0,011	0,250	98,607
249	manguito	0,00131	0,011	0,250	98,618
250	espinillera	0,00128	0,023	0,500	98,641
251	austriaca	0,00122	0,011	0,250	98,652
252	bata de casa	0,00122	0,011	0,250	98,663
253	pantufla	0,00122	0,011	0,250	98,675
254	máquina de coser	0,00122	0,011	0,250	98,686
255	elástico	0,00122	0,011	0,250	98,697
256	alzacuello	0,00122	0,011	0,250	98,709
257	pulóver	0,00122	0,011	0,250	98,720
258	calzoncillo corto	0,00122	0,011	0,250	98,731
259	viso	0,00119	0,023	0,500	98,754
260	codera	0,00116	0,023	0,500	98,777
261	pantalón de género	0,00113	0,011	0,250	98,788
262	punto	0,00113	0,011	0,250	98,799
263	vestido corto	0,00113	0,011	0,250	98,811
264	gabán	0,00113	0,011	0,250	98,822
265	topless	0,00113	0,011	0,250	98,833
266	toca	0,00112	0,023	0,500	98,856
267	traje de novia	0,00106	0,011	0,250	98,867
268	cómodo	0,00106	0,011	0,250	98,879

Cuadro 6.1 (cont.)

1	2	3	4	5	6
269	gomilla	0,00106	0,011	0,250	98,890
270	camisa de manga corta	0,00106	0,011	0,250	98,901
271	patín	0,00106	0,011	0,250	98,913
272	incómodo	0,00098	0,011	0,250	98,924
273	pantalón fino	0,00098	0,011	0,250	98,935
274	sastre	0,00098	0,011	0,250	98,947
275	traje corto	0,00098	0,011	0,250	98,958
276	pinky	0,00098	0,011	0,250	98,969
277	sayón	0,00091	0,011	0,250	98,981
278	muda	0,00091	0,011	0,250	98,992
279	wonderbra	0,00091	0,011	0,250	99,003
280	ajustado	0,00091	0,011	0,250	99,015
281	paraguas	0,00091	0,011	0,250	99,026
282	calzón de deporte	0,00091	0,011	0,250	99,037
283	playera	0,00091	0,011	0,250	99,049
284	anillo	0,00088	0,023	0,500	99,071
285	rebecón	0,00086	0,023	0,500	99,094
286	bajera	0,00085	0,011	0,250	99,105
287	permeable	0,00085	0,011	0,250	99,117
288	jersey de cuello	0,00085	0,011	0,250	99,128
289	suit	0,00085	0,011	0,250	99,139
290	plumífero	0,00085	0,011	0,250	99,151
291	diadema	0,00085	0,011	0,250	99,162
292	boxer	0,00079	0,011	0,250	99,173
293	zapato de charol	0,00079	0,011	0,250	99,185
294	kiowa	0,00079	0,011	0,250	99,196
295	pasador	0,00079	0,011	0,250	99,207
296	suspensorio	0,00079	0,011	0,250	99,218
297	polaina	0,00079	0,011	0,250	99,230
298	suave	0,00079	0,011	0,250	99,241
299	polo de manga larga	0,00079	0,011	0,250	99,252
300	rugoso	0,00074	0,011	0,250	99,264
301	equipación	0,00074	0,011	0,250	99,275
302	blazer	0,00074	0,011	0,250	99,286
303	bota de tacos	0,00074	0,011	0,250	99,298
304	calientacuellos	0,00074	0,011	0,250	99,309

1	2	3	4	5	6
305	pañoleta	0,00074	0,011	0,250	99,320
306	polo de manga corta	0,00074	0,011	0,250	99,332
307	alpargata	0,00074	0,011	0,250	99,343
308	bota de fútbol sala	0,00069	0,011	0,250	99,354
309	camiseta de invierno	0,00069	0,011	0,250	99,366
310	manta	0,00069	0,011	0,250	99,377
311	salto de cama	0,00069	0,011	0,250	99,388
312	braga para la cara	0,00069	0,011	0,250	99,400
313	macuto	0,00069	0,011	0,250	99,411
314	llamativo	0,00069	0,011	0,250	99,422
315	bota de montar	0,00069	0,011	0,250	99,434
316	calcetín media	0,00064	0,011	0,250	99,445
317	sábana	0,00064	0,011	0,250	99,456
318	chaqueta de pana	0,00064	0,011	0,250	99,468
319	extravagante	0,00064	0,011	0,250	99,479
320	peto vaquero	0,00064	0,011	0,250	99,490
321	pantalón corto de deporte	0,00064	0,011	0,250	99,502
322	chaqueta de pingüino	0,00064	0,011	0,250	99,513
323	camiseta de manga larga	0,00064	0,011	0,250	99,524
324	pantalón ancho	0,00059	0,011	0,250	99,536
325	colorido	0,00059	0,011	0,250	99,547
326	chaleco sin mangas	0,00059	0,011	0,250	99,558
327	ropero	0,00059	0,011	0,250	99,570
328	enagua	0,00059	0,011	0,250	99,581
329	abrigo largo	0,00055	0,011	0,250	99,592
330	cinta del pelo	0,00055	0,011	0,250	99,604
331	mantón de Manila	0,00055	0,011	0,250	99,615
332	guardarropa	0,00055	0,011	0,250	99,626
333	twinset	0,00055	0,011	0,250	99,638
334	batín	0,00055	0,011	0,250	99,649
335	zueco	0,00055	0,011	0,250	99,660
336	mascota	0,00055	0,011	0,250	99,672
337	braga para el cuello	0,00055	0,011	0,250	99,683
338	jersey de cuello bajo	0,00055	0,011	0,250	99,694
339	zapato plano	0,00051	0,011	0,250	99,706
340	chupa de cuero	0,00051	0,011	0,250	99,717
341	tentación	0,00051	0,011	0,250	99,728

Cuadro 6.1 (cont.)

1	2	3	4	5	6
342	barbour	0,00051	0,011	0,250	99,739
343	zapato de tacón	0,00048	0,011	0,250	99,751
344	mantilla	0,00048	0,011	0,250	99,762
345	traje de gitana	0,00048	0,011	0,250	99,773
346	pantalón de ciclista	0,00048	0,011	0,250	99,785
347	uniforme militar	0,00045	0,011	0,250	99,796
348	lunar	0,00045	0,011	0,250	99,807
349	ojal	0,00045	0,011	0,250	99,819
350	traje de flamenca	0,00045	0,011	0,250	99,830
351	uniforme	0,00045	0,011	0,250	99,841
352	taparrabo	0,00045	0,011	0,250	99,853
353	camiseta de canalé	0,00041	0,011	0,250	99,864
354	cuadro	0,00041	0,011	0,250	99,875
355	media etiqueta	0,00041	0,011	0,250	99,887
356	lista	0,00039	0,011	0,250	99,898
357	goma elástica	0,00039	0,011	0,250	99,909
358	mono vaquero	0,00036	0,011	0,250	99,921
359	lino	0,00033	0,011	0,250	99,932
360	camiseta de cuello alto	0,00033	0,011	0,250	99,943
361	mitón	0,00033	0,011	0,250	99,955
362	talla	0,00031	0,011	0,250	99,966
363	guardapolvo	0,00031	0,011	0,250	99,977
364	percha	0,00029	0,011	0,250	99,989
365	gargantilla	0,00027	0,011	0,250	100,000

### 6.3. Clasificación del léxico disponible

Dos son los criterios principales empleados a la hora de analizar cualitativamente los listados de léxico disponible:

a) El valor del índice de disponibilidad léxica de cada vocablo. El listado léxico está ordenado –y, a su vez, jerarquizado– según el índice de disponibilidad léxica, el cual combina los parámetros de frecuencia y orden. El valor del índice de disponibilidad se ha relacionado con la teoría de los prototipos (véase 4.1.2). Asimismo, el índice de disponibilidad es tenido en cuenta a la hora de determinar la cantidad de vocablos que se van a comparar –bien entre los listados generales de muestras distintas, bien entre dos listas particulares de una única muestra –.

No obstante, ha habido diferentes propuestas sobre cuál debe ser el punto de corte entre los vocablos que se comparan y los que no, considerando distintos valores del índice de disponibilidad, combinado o no con el rango (las 50, 100 primeras palabras) y la frecuencia acumulada<sup>1</sup>.

b) La descripción lexicológica de las unidades que componen los listados léxicos, catalogando el vocabulario según criterios lexicológicos, semánticos y gramaticales. Mediante este análisis se describen de manera detallada los diferentes fenómenos lingüísticos y el comportamiento de las palabras o grupos de palabras: categorías gramaticales predominantes, procesos de formación de palabras (derivación, composición, parasíntesis, siglas, acrónimos), origen de los vocablos, dialectalismos, extranjerismos, vulgarismos, incidencia de las marcas comerciales, etc.

Mención aparte merece la propuesta de Borrego Nieto (2009), quien incorpora criterios estilísticos explícitos al análisis del léxico disponible. Así, aparte de la variación diatópica y diastrática, este autor analiza la variación diafásica en los listados de léxico disponible, con variantes que incluyen diferentes niveles de formalidad (formal/informal), vulgarismos, tecnicismos de distinto grado, amén de los términos neutros, no marcados.

### 6.3.1. Análisis según el índice de disponibilidad léxica

Tal como explicamos al analizar la perspectiva cognitiva aplicada a los estudios de léxico disponible (cf. 4.1.2), la ordenación de los vocablos según el índice de disponibilidad léxica puede ser explicada satisfactoriamente mediante la *teoría de los prototipos* (Rosch, 1978). Según esta teoría, la organización de las listas de léxico disponible muestra una estructura vertical, en la que se aprecian tres niveles de inclusividad:

- Un nivel supraordinado, que coincide con la etiqueta categorial del centro de interés –en este caso, ‘La ropa’–. Todas las etiquetas de los centros de interés clásicos (‘Partes del cuerpo humano’, ‘La ropa’, ‘Partes de la casa’, etc.) de la disponibilidad léxica se sitúan en el nivel supraordinado, que reúne los atributos comunes a los elementos inferiores.

- Un nivel de base, cuyos elementos léxicos poseen, según Kleiber (1995: 84–88), tres características principales: a) son los que se identifican de manera más rápida, al estar asociados a una imagen mental simple y global; b) suelen corresponder a las palabras más cortas y más frecuentes, siendo términos contextualmente neutros que se identifican con las primeras palabras que entienden y emplean los niños; c) en lo referente a la organización del conocimiento, la mayor parte de los atributos de la categoría se memorizan en este nivel. Los primeros

---

<sup>1</sup> Véase un análisis de las diferentes propuestas en Bellón Fernández (2011: 153–199).

vocablos de las listas de léxico disponible se corresponden con los elementos prototípicos del nivel de base, de tal forma que “las palabras más disponibles son las más prototípicas” (Sánchez-Saus, 2011: 200). De este modo, son casos ejemplares de la categoría ‘La ropa’ los veinte primeros vocablos de la lista cordobesa: *pantalón, camisa, camiseta, jersey, calcetín, chaqueta, falda, zapato, calzoncillo, braga, abrigo, sujetador, media, guante, bufanda, chándal, sudadera, vestido, zapatilla, chaquetón*.

- Un nivel subordinado, “más específico que el básico, [...] formado por elementos menos disponibles, *hipónimos* o *merónimos* (si la relación es naturaleza designativa) de elementos del nivel básico” (Sánchez-Saus, 2011: 200). A medida que el índice de disponibilidad de los vocablos disminuye, los elementos léxicos se van alejando del núcleo prototípico, dando lugar a entradas de todo tipo de niveles de inclusividad. Coexisten así, junto a vocablos del nivel de base, otros de nivel subordinado, como *pantalón corto* (42)<sup>2</sup>, *camiseta interior* (43), *calzón* (61), *pantalón de pinzas* (72), *zapatilla de deporte* (84), *falda corta* (89), *falda larga* (93), *bota katuska* (100), *falda pantalón* (106), *camiseta de tirantes* (109), *pantalón de vestir* (125), etc. En este nivel aumenta la presencia de construcciones de naturaleza sintagmática (sustantivo + adjetivo o sustantivo + preposición sustantivo/verbo).

En otro orden de cosas, y, con vistas a efectuar comparaciones interdiatópicas, un valor concreto del índice de disponibilidad léxica también puede ser tomado como punto de corte en el interior de un listado léxico. Este valor fija el límite del subconjunto comparado. Bartol Hernández (2008: 208–209) distingue dos subgrupos de vocablos diferenciados, en función del índice de disponibilidad, que marca, además, la representatividad de los vocablos dentro del campo léxico:

- a) El primer subgrupo está constituido por palabras altamente representativas de ese campo léxico, con índices de disponibilidad altos –de al menos el 0,01–, cuyas frecuencias acumuladas deben superar el 75% del total. Son palabras mencionadas por un porcentaje importante –al menos del 2%– de la muestra y que, por todo ello, no ofrecen dudas sobre su adscripción a la norma léxica de la comunidad. Se trata, en fin, de palabras altamente utilizables en esa comunidad, lo cual significa que en situaciones normales un individuo de ese colectivo al hablar del tema de que se trate usará con un porcentaje de posibilidades de al menos del 75% una de las palabras de ese grupo. Son palabras que se adquirieron a una edad temprana, y/o son para los informantes típicas del campo léxico, y/o les son muy familiares (de uso muy próximo).

En el caso de la lista cordobesa de ‘La ropa’ –que manejamos como ejemplo– integran este grupo 95 vocablos, con una frecuencia acumulada de 93,046%. Las últimas diez palabras de este subgrupo en la lista cordobesa son: *chal, bolso, reloj,*

---

<sup>2</sup> Entre paréntesis, el número de orden en la lista de léxico disponible de Córdoba (Bellón Fernández, 2011).

*falda corta, plumón, traje de chaqueta, maillot, falda larga, americana, tejano.* Entre estas palabras se hallan complementos de la indumentaria (*chal, bolso, reloj*), especificaciones de prendas básicas (*falda larga, traje de chaqueta, falda corta, [abrigo de] plumón*), sinónimos cultos (*americana* de *chaqueta*, *tejano* de *vaquero*) y una prenda deportiva (*maillot*).

La alta representatividad y frecuencia de uso de estas palabras contrasta con la de los términos que aparecen inmediatamente a continuación, ya por debajo del punto de corte. Entre los diez primeros se cuentan: *jeans, combinación, levita, niqui, (bota) katiuska, capa, botín, sostén, mocasín, chupa*. Estos primeros vocablos del segundo subgrupo se caracterizan por la especificidad de su función (*combinación* como ropa interior femenina), su menor uso social (*levita* como prenda de etiqueta masculina frente a *traje* o *frac*, *capa* como prenda elegante tradicional sustituida generalmente por *chaqueta, abrigo, gabardina*, etc., *mocasín* como calzado de piel moderno que imita al que usaban los indios), su carácter de extranjerismos (*jeans* del inglés como *vaqueros*, *niqui* del alemán Nicki como *camiseta*, *katiuska* del ruso *Katjuša* como *botas*), el rendimiento de su denominación (*botín* como media bota denominada también simplemente *bota*), su carácter marcado (*sostén* frente a *sujetador*, *chupa* frente a *cazadora de cuero*).

b) El segundo subgrupo, como acabamos de ver a partir de sus primeros diez vocablos, lo forma un variado repertorio de palabras cuyo índice de disponibilidad es bajo –inferior a 0,01–, que fueron mencionadas por pocos informantes. La variedad de este grupo se debe a muy diversos factores, entre los que Bartol Hernández (2008: 208–209) cita:

– Palabras poco representativas –típicas– del centro de interés, que no responden adecuadamente al estímulo propuesto, lo que hace que sean evocadas por pocos informantes. Son por ello palabras que presentan poca estabilidad en las encuestas, en unas pueden aparecer y en otras no. En el ejemplo analizado hallamos términos como: *capa* (101), *falda pantalón* (106), *manga* (119), *botón* (120), *collar* (121), *pendiente* (124), *cremallera* (128), *mantón* (136), *hombreira* (140), *velo* (153), etc.

– Palabras de incorporación reciente. A este grupo pertenecen muchos de los anglicismos no integrados: *jeans* ‘vaqueros’ (96), *bomber* ‘chaqueta de aviador’ (115), *minishort* ‘pantalones extra-cortos femeninos’ (135), *husky* ‘chaqueta acolchada’ (181), *braslip* ‘tipo de calzoncillos’ (210), *kilt* ‘falda escocesa’ (219), *minipull* ‘minijersey’ (227), *wonderbra* ‘sujetador realzador’ (279), etc.

– Palabras en proceso de desaparición. Su análisis es muy interesante para ver la evolución del léxico. En el grupo de las palabras que se hallan en desuso, en proceso de desaparición suelen estar muchos dialectalismos: *babucha* ‘zapato sin tacón árabe’ (127), *calceta* ‘labor de punto’ (142), *gayumbos* ‘calzoncillos’ (165), *pelliza* ‘abrigo reforzado con pieles en el cuello y las mangas’ (171), *zarcillo*

‘pendiente, arete’ (195), *loneta* ‘lona delgada’ (244), *pantufla* ‘calzado de casa’ (253), *gabán* ‘chaquetón’ (264), *alpargata* ‘zapato de lona con suela de esparto’ (307), etc.

– Palabras “atípicas”, mencionadas ocasionalmente por uno o varios informantes que han estado recientemente en contacto con ellas a través de una película, una clase, un programa de televisión, una canción, etc. Aquí se englobarían las palabras relacionadas con el mundo de la hípica y la tauromaquia –*calzona* ‘calzón de los picadores y vaqueros’ (110), *torera* ‘chaquetilla ceñida sin abotonar’ (160), *bajera* ‘manta bajo la silla de montar’ (286), *polaina* ‘media calza de cuero abotonada hasta la rodilla’ (297), *bota de montar* (315)–, ropa de bebé –*babero* (212), *braga pañal* (220), *pelele* ‘pijama de bebé de una sola pieza’ (230)–, ropa interior de carácter erótico –*combinación* ‘camisón corto’ (97), *culote* ‘braga femenina que resalta las nalgas’ (112), *corsé* (152), *picardías* ‘camisón corto transparente’ (198), *suspensorio* ‘minicalzoncillo que cubre únicamente el miembro viril’ (296), *salto de cama* ‘bata ligera de mujer para el momento de levantarse de la cama’ (311), *tentación* ‘camisón corto transparente’ (341)–, ropa deportiva –*zapatilla deportiva* (164), *bota de fútbol* (182), *pantalón de deporte* (215), *tenis* ‘zapato deportivo’ (238), *espinillera* ‘protector de la espinilla’ (250), *bota de tacos* (303), *bota de fútbol sala* (308), *pantalón de ciclista* (346), etc.–, ropa de época (*calza* ‘prenda ceñida o no que cubría el muslo y parte de la pierna’ (217), *jubón* ‘prenda ajustada al cuerpo que cubría de los hombros a la cintura’ (236), *toca* ‘prenda de tela con que se cubrían la cabeza las mujeres, y ahora las monjas’ (266), *sayón* ‘[túnica de] cófrade’ (277), *enagua* ‘falda interior’ (328)–, ropa ligada a una actividad u oficio muy específicos –*toga* ‘traje de ceremonia de magistrados, letrados, catedráticos’ (199), *tutú* ‘faldellín usado por las bailarinas de danza clásica’ (222), *malla corta* ‘vestido muy fino ajustado al cuerpo usado por bailarinas y artistas de circo’ (231), *alzacuello* ‘tira suelta rígida se ciñen al cuello los sacerdotes’ (256), *traje de novia* (267), *uniforme militar* (347), etc.

– *Creaciones* individuales más o menos afortunadas y graciosas de los informantes que recurren a ellas cuando no recuerdan alguna palabra en concreto. Su aparición pone de manifiesto la vitalidad de los recursos de creación léxica de la lengua. Confundidas con distintas especificaciones de uso corriente (*traje de flamenca*, *traje de gitana*, *traje de novia*, *vestido de fiesta*, *zapato de charol*), “aparecen creaciones personales de un único hablante que, en muchas ocasiones, pueden ser circunstanciales, ya que el objeto de aportar ese término probablemente sea rellenar una línea más en el cuestionario” (Bellón Fernández, 2011: 163). Tal es el caso de *camisa de tirantes* (175), *pantalón liso* (192), *camiseta de verano* (229), *pantalón fino* (273), *calzón de deporte* (282), *jersey de cuello* (288), *polo de manga larga* (299), *polo de manga corta* (306), *peto vaquero* (320), *chaqueta de pingüino* (322), *pantalón ancho* (324), *chaleco sin mangas* (326), *jersey de cuello bajo* (338), *mono vaquero* (358), *camiseta de cuello alto* (360).

### 6.3.2. Análisis descriptivo lexicológico

Otra manera de enfocar el análisis cualitativo de los listados de léxico disponible ha sido a partir de criterios estrictamente lexicológicos y gramaticales. Buenos ejemplos de este tipo de análisis los encontramos en las investigaciones de Almería (Mateo García, 1998), Cádiz (González Martínez, 2002), Melilla (Fernández Smith *et al.*, 2008), o la que Sánchez-Saus (2011) realiza sobre los estudiantes extranjeros de las universidades andaluzas.

Dado que la disponibilidad léxica se centra en las palabras temáticas, los listados de léxicos disponible recogen un alto porcentaje de sustantivos, y sólo en los centros de interés ‘Trabajos del campo y del jardín’ y ‘Juegos y distracciones’ ganan protagonismo los verbos. No obstante, al analizar la naturaleza del léxico obtenido en las encuestas, es posible realizar un estudio descriptivo de las unidades léxicas pertenecientes a cada centro de interés, que no depende necesariamente de los resultados aportados por los índices de disponibilidad. Este tipo de análisis nos aproxima a algunas de las preguntas tradicionales de la lingüística en relación al material léxico recogido empíricamente:

¿qué extranjerismos aparecen, en qué proporción y de qué lengua proceden? ¿hasta qué punto predomina el léxico patrimonial de origen dialectal o, por el contrario, se impone una cierta estandarización léxica y, de ser ello así, cuál es el volumen e influencia de la misma? ¿qué repercusión tienen en los procedimientos de designación mediante siglas y acrónimos? O, entre otros, ¿cuál es la implantación de las marcas y firmas comerciales en los repertorios del vocabulario de las comunidades de habla y hasta qué punto llegan a sustituir a las formas léxicas hasta ahora establecidas para designar a los objetos que estas nombran? (Mateo García, 1998: 95)

Este planteamiento hace que, dentro de la metodología sociolingüística adoptada en el Proyecto Panhispánico, y siguiendo una perspectiva variacionista, sea posible prestar una atención individualizada a los factores lingüísticos que se dejan ver en las listas de léxico disponible, además de analizar la influencia de los factores sociales en el léxico disponible, claro está. Para ello, González Martínez (2002: 47) clasifica en dos grandes grupos el léxico disponible: por un lado, el léxico heredado, que constituye el grueso del corpus, y por otro, los extranjerismos. Tanto en la investigación de Cádiz como en la de Almería los autores estudian también los andalucismos, así como las siglas y las marcas comerciales lexicalizadas.

Ambos autores completan el análisis léxico con el estudio de los grupos léxicos, aspecto relacionado con la delimitación de las unidades léxicas en la lexicometría, puesto que “los hablantes asocian vocabulario de una determinada temática con un conjunto ciertamente diversificado de posibilidades léxicas, que cubrirían desde palabras en sentido estricto hasta diversas clases de lo que tradicionalmente se han llamado grupos nominales y verbales” (Mateo García, 1998: 98). Estos grupos, expresiones de más de un término que tienen como núcleo

a un sustantivo o a un verbo, “son utilizados por los encuestados para especificar o determinar más claramente la significación de los vocablos recogidos en las encuestas” (González Martínez, 2002: 84). Constituyen, además, una parte muy importante del léxico disponible de la comunidad; por ejemplo, en el léxico disponible de Cádiz, los grupos nominales suponen el 13,8% del total de sustantivos recogidos y los grupos verbales el 41,9% de los verbos computados en las encuestas. En el análisis de los grupos nominales y verbales que ambos autores llevan a cabo, se tiene en cuenta las diferentes formas de agrupación de los grupos, y, dentro de estos, las sucesivas subdivisiones según el tipo de elemento que interviene en los mismos y la presencia o ausencia del nexos.

Por su parte, Sánchez-Saus (2011) dota a su estudio de una perspectiva léxico-semántica. En consecuencia, analiza los distintos núcleos temáticos que se descubren en el interior de los centros de interés, ofreciendo a continuación la cuantificación de las categorías lingüísticas mayoritarias, que en todos los trabajos sobre disponibilidad suelen ser los sustantivos. Junto a esto, analiza los fenómenos lingüísticos de mayor relevancia que afectan a este léxico, centrándose en cuatro ámbitos: la variación (diatópica, diafásica y especializada), los cambios semánticos (por el empleo de eufemismos y disfemismos, por metáforas y metonimias y por elipsis originada en combinatoria léxica), la importación léxica (extranjerismos, préstamos y calcos) y la formación y creación léxicas (derivación, composición y creación a partir de acortamientos).

Según los criterios lexicológicos y semánticos expuestos hasta ahora, el análisis cualitativo del listado de léxico disponible de Córdoba correspondiente al centro de interés 02. ‘La ropa’ queda como sigue:

a) Campos léxicos

Las mayoría de palabras registradas en este centro –y entre ellas las más disponibles– pertenecen al campo léxico de las prendas de vestir. Dentro de este campo, pueden distinguirse campos más reducidos (ropa femenina, ropa interior, calzado, etc.), otros más alejados del núcleo central como los complementos, que incluye elementos que no son prendas, y otros campos que no son prendas, pero que tienen una relación muy estrecha con ellas, como materiales y formas.

– ‘Prendas de vestir’: *pantalón, camisa, camiseta, jersey, calcetín, chaqueta, falda, zapato, calzoncillo, braga, sujetador, media, guante, bufanda, chándal, etc.*

– ‘Ropa femenina’: *falda, media, vestido, rebeca, blusa, camión, bikini, minifalda, top, pichi, ligero, liga, falda pantalón, traje de novia, etc.*

– ‘Ropa interior’: *calzoncillo, braga, sujetador, body, camiseta interior, slip, panty, corpiño, tanga, combinación, sostén, lencería, etc.*

– ‘Calzado’: *zapato, zapatilla, bota, chancla, sandalia, zapatilla de deporte, botín, mocasín, bamba, babucha, zapatilla de casa, tacón, bota de agua, etc.*

– ‘Prendas exteriores’: *chaqueta, abrigo, sudadera, chaquetón, rebeca, cazadora, anorak, gabardina, chubasquero, impermeable, bata, albornoz, chal, capa, etc.*

– ‘Complementos’: *guante, bufanda, sombrero, corbata, gorro, cinturón, pañuelo, gorra, pajarita, correa, boina, cordón, bolso, reloj, collar, pendiente*, etc.

– ‘Tejidos’: *pluma, algodón, pana, tela, lana, sintético, poliéster, tul, lycra, elástico, punto, lino*, etc.

– ‘Partes de una prenda’: *peto, cordón, manga, cremallera, visera, cuello, bolsillo, puño, escote, sisa, bajo, ojal*, etc.

#### b) Léxico heredado

Considerando como tal todos los vocablos que aparecen en el *DRAE*, sea cual sea su origen y su momento de entrada en el español. Se trata de un léxico arraigado entre los hablantes y de uso generalizado. Según su origen, puede ser:

– Léxico procedente del latín: *abrigo, aguja, ajustado, anillo, armario, austriaca, bajera, bajo, bañador, bolsillo, bolso, bombín, bota, calza, calceta, calcetín, calentador, calidad, calzado, calzón, calzona, calzoncillo, capa, castellano, cazadora, cinta, cinturón, codera, colgante, collar, colorido, combinación, (in) cómodo, corpiño, correa, cuadro, cuello, delantal, deporte, deportivo, ejecutivo, elástico, espinillera, extravagante, faja, fajín, gemelo, goma, hilo, hombrera, imperdible, lana, lazo, lencería, liga, liguero, lino, llamativo, lunar, manga, máquina, manopla, mantilla, mantón, media, muda, muestrario, náutico, ojal, orejera, pajarita, pañoleta, pañuelo, paraguas, pareo, pasador, pasamontañas, pelliza, pendiente, (im)permeable, pesado, peto, playera, pluma, plumífero, plumón, polo, pulsera, punto, puño, rodillera, rugoso, sábana, salto de cama, saquito, sayón, sombrero, sostén, suave, sudadera, sujetador, suspensorio, talla, tejano, tela, tentación, textura, tienda, toga, torera, traje, tres cuartos, túnica, uniforme (militar), vaquero, velo, vestido, visera, viso, zarcillo, zueco.*

– Léxico procedente del griego: *diadema, polar, reloj, sandalia, sintético.*

– Léxico procedente del árabe: *albornoz, algodón, alpargata, babucha, bata, chaleco, chupa, gabán, gabardina, mono.*

– Léxico procedente del francés: *alzacuello, blusa, blusón, botín, botón, broche, chal, chándal, chaqué, chaqueta, chaquetón, corchete, cordón, corsé, cremallera, culote, equipación<sup>3</sup>, escote, etiqueta, fular, frac, leotardo, levita, ligero, maillot, malla, mascota, mitón, moda, pana, pantalón, pantufla, patín, percha, perchero, polaina, sisa, tul, tutú, visón.*

– Léxico procedente del italiano: *camisola, corbata.*

– Léxico procedente de lenguas americanas: *anorak, enagua, macuto, pichi, poncho, tanga.*

– Léxico de procedencia prerromana: *muñequera<sup>4</sup>.*

<sup>3</sup> Este término es sinónimo de *equipaje* o de *indumentaria*, especialmente con el significado de ‘conjunto de ropas y cosas de uso particular de un futbolista de un determinado equipo’.

<sup>4</sup> *Muñequera* es un derivado de *muñeca*. Según el *DRAE*, la palabra *muñeca* es de origen prerromano, relacionado con ‘moño’ y el euskera *munu* (‘colina’). Esta versión también está respaldada

– Léxico de procedencia germánica: *equipación*<sup>5</sup>, *falda*, *guante*, *guarda*, *guardarropa*, *lista*, *minifalda*, *niqui*, *ropa*, *ropero*, *taparrabo*, *toalla*.

– Léxico procedente del portugués: *chubasquero*.

– Léxico procedente del catalán: *sastre*.

– Léxico procedente del vasco: *chapela*, *chistera*.

– Léxico procedente de lenguas asiáticas: *chal*<sup>6</sup>, *chaleco*, *pijama*, *quimono*, *zapatilla*, *zapato*.

– Léxico de origen incierto: *braga*, *boina*, *camisa*, *camiseta*, *camisón*, *casaca*, *felpa*, *gafas*, *gamberro*, *gayumbo*, *gorra*, *gorro*, *pelele*, *picardía*, *tacón*, *tiran-te*, *toca*, *trenca*.

– Léxico de origen onomatopéyico: *babero*, *bufanda*, *chancla*, *gargantilla*.

– Léxico de origen metonímico: *bermudas*<sup>7</sup>, *biquini*<sup>8</sup>, *katiuska*<sup>9</sup>, *kiowa*<sup>10</sup>, *lona*<sup>11</sup>, *pamela*<sup>12</sup>, *rebeca*<sup>13</sup>.

### c) Andalucismos

Se consideran como andalucismos los que vienen como tales en el *DRAE*, los que aparecen en el *Vocabulario Andaluz* (Alcalá Venceslada, 1998) [1951] y los considerados así en estudios dialectológicos.

– *Babero*, *manopla*.

### d) Extranjerismos

---

por Corominas (1973), quien agrega que originalmente muñeca significaba ‘hito y mojón’, de ahí a ‘protuberancia’ y más tarde a ‘articulación abultada de la mano con el antebrazo. No fue hasta el año 1400 que se le dio la acepción de ‘juguete de trapo con figura femenina con el cual juegan las niñas’.

<sup>5</sup> La palabra *equipo* y sus derivados, que el castellano adoptó del francés ‘équipé’, tenían en la Edad Media una relación con las actividades marítimas. ‘Équiper’ era ‘embarcar’ y luego ‘proveer una nave de todo lo necesario’ antes de zarpar del puerto, ya que la voz procedía del germánico ‘skip’ (barco) (alemán: *Schiff*; inglés: ‘*ship*’), a través de una forma existente en zonas del norte de Francia.

<sup>6</sup> Del fr. *châle*, y este del persa *šāl*.

<sup>7</sup> Se refiere a las Islas Bermudas, donde los soldados ingleses usaban este tipo de prenda. El nombre de estas islas es en honor del navegador español Juan de Bermúdez, quien las descubrió en 1505.

<sup>8</sup> De la Isla Bikini, en las Islas Marshall.

<sup>9</sup> El nombre proviene de una famosa zarzuela escrita por Emilio González del Castillo, que llevaba por título ‘Katiuska, la mujer rusa’ y cuya protagonista principal aparecía en escena provista de unas botas altas de media caña, las cuales recordaban a las utilizadas comúnmente en los días de lluvia.

<sup>10</sup> Zapatos tipo mocasín, cuyo nombre se debe a la tribu de los indios kiowa, que fabricaban su calzado partiendo de una sola pieza de piel que envuelve toda la horma, sobre la cual, posteriormente, se monta la suela y la pala (la parte superior del zapato que va desde el empeine hasta la punta de los dedos).

<sup>11</sup> De *Olonne*, población marítima de Francia, donde se tejía esta clase de lienzo.

<sup>12</sup> De Pamela, nombre de la heroína y título de la obra del novelista inglés S. Richardson.

<sup>13</sup> Del nombre propio Rebeca, título de un filme de A. Hitchcock, basado en una novela de D. du Maurier, cuya actriz principal usaba prendas de este tipo.

Entendidos como tales aquellos vocablos procedentes de otras lenguas que no estaban presentes en la edición del *DRAE* de 1992, o que mantienen la grafía original.

– Anglicismos: *blazer, body, bomber, bóxer, braslip, cardigan, esmoquin, jeans, jersey, minipull, minishort, mocasín, panty, pijama, poliéster, pulóver, short, slip, suéter, suit, tenis, top, topless, twinset, wonderbra.*

– Otros: *parka*<sup>14</sup>, *husky*<sup>15</sup>, *kilt*<sup>16</sup>.

e) Siglas

En este centro de interés, no aparecen.

f) Marcas comerciales

– *bamba* (de Wamba), *barbour, lycra, meyba, pinky, twinset.*

g) Grupos léxicos

– Grupos nominales:

• Grupos sin nexos

– Nombre + Adjetivo: *abrigo largo*; {*calzoncillo, pantalón, traje, vestido*} + (*corto, largo*); {*camisa, cazadora*} + *vaquera*; {*camisa, camiseta, falda*} + (*corta, larga*); {*camisa, camiseta, ropa*} + *interior*; *forro polar, goma elástica, malla corta*; {*mono, peto*} + (*corto, largo*); *pantalón* + (*ancho, clásico, fino, liso*); *uniforme militar*; *zapato* + (*deportivo, plano*).

– Nombre + Nombre: *bota katiuska, braga pañal, calcetín media, falda pantalón.*

– Determinante + Nombre: *media etiqueta, tres cuartos.*

• Grupos con nexos

– Nexos “de”: *bata de casa, bota (de agua, de fútbol, de montar, de piel, de tacos), calzón de deporte, camisa (de manga larga/corta, de tirantes), camiseta (de canalé, de cuello alto, de invierno, de manga larga, de tirantes, de verano), chaqueta (de pana, de pingüino), chupa de cuero, cinta del pelo, cuello de la camisa, jersey (de cuello alto/bajo, de pico), mantón de Manila, máquina de coser, mono de trabajo, pantalón (de ciclista, de deporte, de género, de pana, de pinzas, de pitillo, de vestir), pañuelo de cuello, polo de manga larga/corta, salto de cama, traje (de chaqueta, de flamenca, de gitana, de noche, de novia), vestido de fiesta, zapatilla (de casa, de deporte), zapato (de charol, de tacón).*

– Nexos “para”: *braga (para la cara, para la cara).*

– Grupos verbales.

En este centro no aparecen.

<sup>14</sup> Prenda de abrigo de material impermeable por fuera y acolchada por dentro. Del ruso, y este a su vez de las lenguas de Siberia komi y nenets.

<sup>15</sup> Término general para mencionar a diferentes tipos de perros utilizados como perros de trineo. Deformación del término derivativo “eskie”, utilizado para mencionar a las tribus esquimales que entraron en contacto con los europeos que realizaron las primeras expediciones a sus tierras.

<sup>16</sup> Falda masculina escocesa tradicional.

### 6.3.3. Análisis descriptivo según niveles de lengua (registros)

El planteamiento metodológico del Proyecto Panhispánico de disponibilidad léxica es el de recabar el léxico que conforma la norma hispánica culta de un hablante medio. Por ello, los sujetos que componen las muestras de cada comunidad de habla son obligatoriamente estudiantes del nivel preuniversitario. Se parte, por consiguiente, de la idea “de que los informantes elegidos ya han alcanzado la formación y madurez de un adulto cultivado, pero sin haber sido expuestos todavía a la contaminación técnica propia de las facultades universitarias, que los convertiría en menos aptos para representar la norma léxica no especializada” (Borrego Nieto y Fernández Juncal, 2003: 167). Además de considerar el alcance de esta precaución por evitar el vocabulario excesivamente especializado en los léxicos disponibles, también se pregunta Borrego Nieto (2004: 61) hasta qué punto la situación de la encuesta de léxico disponible “influiría en la «formalidad» de las respuestas y coartaría la aparición de coloquialismos, vulgarismos, dialectalismos, términos malsonantes y de argot, sin duda disponibles en buena medida en los sujetos”.

En cualquier caso, la presencia de tecnicismos y de términos marcados formalmente en las listas ha sido constatada abiertamente en el léxico disponible, en interferencia frecuentemente con variables de tipo social y geográfico. De este modo, Samper Hernández (2003a: 203) constata que en determinadas edades, ‘partes del cuerpo’ como *tobillo*, *espalda*, *ombligo* tienen una disponibilidad menor que otras menos cotidianas aprendidas en los libros de texto. Por otro lado, distintos autores, entre los que se hallan López Morales (2005) en San Juan de Puerto Rico, Galloso Camacho (2003a) en Ávila, Zamora y Salamanca, Samper Hernández (2003a) en Canarias, han destacado como las mujeres prefieren en menor proporción que los hombres los términos malsonantes o vulgares para referirse a *realidades tabuizadas* como las partes genitales.

Finalmente, Borrego Nieto (2009), quien ha dedicado dos trabajos a clasificar las variantes diafásicas (Borrego Nieto, 2001, 2002a), ha enriquecido el análisis lingüístico anterior, al incorporar criterios estilísticos explícitos. Este autor ha observado que en la composición de los listados de léxico disponible no sólo se deja notar la influencia de las características sociológicas de los informantes (situación cultural y económica de los padres, tipo de centro escolar, sexo...) y, en menor medida, la procedencia geográfica de los sujetos, sino que la otra gran fuente de variación, los registros, estilos o niveles de lengua, interviene decididamente en la configuración final de ese léxico. A esta conclusión llega al analizar los valores diafásicos en las series sinonímicas o cuasi-sinonímicas que se dan en el centro de interés ‘Partes del cuerpo humano’, en una amplia muestra de trabajos inscritos en el Proyecto Panhispánico, en el que se incluyen los léxicos disponibles de Asturias (Carcedo González, 2001), Aragón (Arnal Purroy *et al.*, 2004), Ávila,

Salamanca y Zamora (Galloso Camacho, 2003), Burgos (Fernández Juncal, 2008), Soria (Bartol Hernández, 2004a), Castilla-La Mancha (Hernández Muñoz, 2006), Valencia (Gómez Devís, 2004), Jaén (Ahumada Lara, 2006), Cádiz (González Martínez, 2002), Puerto Rico (López Morales, 1999a) y República Dominicana (Alba, 1995a). Se trata de parejas como *anginas/amígdalas*; *boca/morros*; *caratez*; *nariz/fosas nasales*; *piel/dermis*; *pupila/iris*; *rabadilla/coxis*; *vaginal/vulva*; *barriga/panza*; etc.

Borrego Nieto (2009) clasifica las variantes diafásicas del centro de interés ‘Partes del cuerpo humano’ en los siguientes tipos:

– Términos de referencia neutra, no marcados, “de aparición posible, sin disonancia ni connotación alguna, en todas las situaciones, registros y grados de formalidad” (Borrego Nieto, 2009: 63), que suelen situarse sobre todo entre los 50 primeros vocablos de las listas, como: *cabeza*, *brazo*, *pierna*, *mano*, *pie*, *nariz*, *oreja*, *boca*, *cuello*, *pelo*, *diente*, etc.

– Tecnicismos crudos, es decir, con poca difusión fuera del ámbito de los especialistas, y sin variantes, como *esternocleidomastoideo*, *bronquiolos*, *trompas de Falopio*, *metatarso*, *cardias*, etc.

– Tecnicismos extendidos, asimilados al vocabulario general culto: *antebrazo*, *glúteos*, *fémur*, *testículos*, *cráneo*.

– Términos informales o coloquiales, como *barriga* ‘vientre’, *sobaco* ‘axila’, *espinazo* ‘columna vertebral’, *bola* ‘biceps’, *mollete* ‘mejilla’, *tripas* ‘intestinos’, *piños* ‘dientes’, *panza* ‘vientre’, *sesos* ‘cerebro’, *zancajo* ‘pierna’, *napias* ‘nariz’, *pinreles* ‘ojos’, *lomo* ‘espalda’, *ternilla* ‘cartílago’.

– Términos transgresores, vulgares o malsonantes, sobre todo correspondientes a los órganos genitales, para el ‘pene’ *polla*, *pija*, *nabo*, *carajo*, *chorra*, *verga*, *capullo*, *bicho*; para la ‘vagina’ *coño*, *chocho*, *conejo*, *potorro*, *almeja*, *chumino*; además de *ojete* ‘ano’, *cojones* ‘testículos’, *tetas* ‘pechos’.

– Eufemismos de las realidades tabuizadas: *trasero*, *pompis*, *posaderas*, *derrriere*, *sentadura*, *pechos*, *senos*, *cola*, *rabo*, *pilila*, *lo que tiene la mujer*, *pubis*, *sexo*, *partes*, *entrepierna*.

– Términos peyorativos o de intención humorística, como *zarpa*, *pecuezo*, *pata*, *morros*, *olla*, *tarro*, *jamones*, *melones*, *domingas*, *pechuga*, etc.

– Términos de registro elaborado (arte, moda, literatura): *busto*, *torso*, *cabello*, *tez*, *faz*, *rostro*.

En el listado de léxico disponible del centro de interés ‘La ropa’ de Córdoba, que hemos venido analizando en este capítulo, también se confirma la presencia de series de sinónimos. En este, como ya constata Borrego Nieto (2009: 73), “las variantes diafásicas tienden a formar series estables, con un *término de referencia* neutro y el resto dispuesto de acuerdo con un orden fijo”:

– *anorak* (38), *parka* (76).

– *bata* (57), *albornoz* (69).

- *braga* (10), *panty* (47).
- *calzoncillo* (9), *slip* (52), *calzón* (61), *gayumbo* (165), *braslip* (210), *calzoncillo corto* (258), *bóxer* (292).
- *camiseta* (3), *camiseta interior* (43), *niqui* (99).
- *camisón* (41), *combinación* (97), *picardías* (195), *salto de cama* (311), *tentación* (341).
- *cazadora* (29), *chupa* (105), *chupa de cuero* (340).
- *cinturón* (33), *correa* (74).
- *corpiño* (79), *corsé* (152).
- *chancla* (64), *bamba* (123), *playera* (283).
- *chaqueta* (6), *americana* (94).
- *chaquetón* (20), *gabán* (264).
- *frac* (117), *chaqueta de pingüino* (322).
- *jersey* (4), *suéter* (51), *pulóver* (257).
- *media* (13), *panty* (47), *calza* (217).
- *pantalón* (1), *pantalón de pinzas* (72), *pantalón de vestir* (125), *pantalón clásico* (168), *pantalón liso* (192), *pantalón fino* (273).
- *pantalón corto* (42), *short* (49).
- (*pantalón*) *vaquero* (21), *tejano* (95), *jeans* (96).
- *pendiente* (124), *zarcillo* (195).
- *plumón* (90), *plumífero* (290).
- *sujetador* (12), *sostén* (103).
- *ropa interior* (67), *muda* (278).
- *traje* (45), *suit* (289).
- *vestido* (18), *traje largo* (149), *vestido largo* (176), *traje de noche* (235), *vestido de fiesta* (237).
- *zapatilla* (19), *zapatilla de casa* (139), *pantufla* (253).
- *zapatilla* (19), *zapatilla deportiva* (164), *deportivo* (218), *tenis* (238).

El orden de la serie sinonímica al que se refiere Borrego Nieto (2009) estipula que el segundo término suele situarse a gran distancia del primero. En el listado que analizamos se aprecian separaciones considerables en las series de los vocablos *calzoncillo*, *camisón*, *cazadora*, *corpiño*, *chancla*, *chaqueta*, *chaquetón*, *frac*, *jersey*, *media*, *pantalón*, *pendiente*, *plumón*, *sujetador*, *ropa interior*, *traje*, *vestido*, *zapatilla*.

En cuanto a la naturaleza de los vocablos de la lista, aunque el segundo término tiende a no ser informal, en un buen número de casos es marcado, aumentando esta probabilidad para el resto de los sinónimos, como también había notado Borrego Nieto (2009: 73). De esta manera, en el listado aparecen extranjerismos (*parka*, *panty*, *slip*, *braslip*, *bóxer*, *niqui*, *corsé*, *suéter*, *pulóver*, *short*, *jeans*, *suit*, *tenis*, *pantufla*), cultismos (*alzacuello*, *gabán*, *mitón*, *salto de cama*), dialectalismos (*manopla*, *playera*, *zarcillo*), tecnicismos y lenguaje especializado (*babero*,

*pelele*, *suspensorio*, *tutú*), lenguaje jergal (*chupa*), coloquialismos (*calzón*, *chistera*, *gayumbo*, *picardías*, *tentación*, *chaqueta de pingüino*), eufemismos (*muda*), marcas comerciales (*bamba* de la marca de zapatillas *Wamba*), arcaísmos léxicos por haber perdido vigencia la realidad a la que hacen referencia (*alpargata*, *enagua*, *polaina*), etc.

#### 6.4. Los extranjerismos en el léxico disponible del español

Dentro de los análisis sobre el léxico recopilado por la disponibilidad léxica, destacan los análisis lexicológicos que miden el índice de aparición de extranjerismos —en especial anglicismos— en los listados de léxico disponible. Con estos estudios, “no sólo señala el conocimiento y uso de palabras prestadas de otras lenguas que tiene un individuo, sino que además localiza el préstamo en los dominios más básicos de la comunicación” (Mateo García, 1998: 53), aportando información importante sobre el grado de penetración en la lengua.

Mackey (1965, 1971a, 1971b) fue el primer autor que hizo un recuento de los anglicismos que aparecían en sus estudios sobre el francés canadiense. Entonces concluía la gran influencia del inglés en la cultura francocanadiense, dado que de las 10.511 unidades analizadas en Acadie, 1.397 eran anglicismos.

Por lo que se refiere al español, el primer trabajo en este campo lo realizó López Chávez (1991), quien llevó a cabo el análisis de dos centros de interés (‘Partes del cuerpo’ y ‘Medios de transporte’) en las muestras de preescolares y de alumnos mexicanos con primaria terminada. Ello le permitió comprobar que el primer campo —‘Partes del cuerpo’— rechaza el anglicismo mientras que ‘Medios de transporte’ recogía muchas voces procedentes del inglés.

Desde entonces, el estudio de “la presencia de anglicismos [o, en general, de extranjerismos] en los diferentes centros de interés se ha convertido en uno de los principales temas de análisis del vocabulario recogido en las distintas comunidades de habla” (Samper Padilla y Samper Hernández, 2009: 869), así como de manera comparativa entre ellas. De manera paralela a los estudios, se ha desarrollado una rica reflexión teórica y metodológica acerca de cómo determinar el corpus objeto de análisis, es decir, delimitar cuáles son las palabras que deben ser incluidas bajo el rótulo de *extranjerismo*. En esta reflexión “no se trata solo del alcance e interpretación de conceptos como *préstamo* o *extranjerismo*, sino de otros factores, como el criterio de admisión de palabras dentro de cada campo, las marcas comerciales aceptadas o rechazadas, la interpretación e inclusión de las siglas, las unificaciones efectuadas, etc.” (Castañer Martín, 2008: 52).

No es esta una cuestión baladí, ya que como dice la misma autora, “es evidente que las diferencias detectadas entre unos y otros listados no responden

exclusivamente a variaciones diatópicas sino que también reflejan la diferente postura del investigador” (Castañer Martín, 2008: 53). Amén de la influencia de los criterios de edición en el análisis realizado, de capital importancia es la definición del concepto de *anglicismo* o *extranjerismo*, de la que se parta, ya que la frontera entre préstamo y extranjerismo es borrosa y se basa en criterios tan dispares como la asimilación gráfica, fónica y gramatical a las pautas de la lengua receptora, por un lado, y la conciencia lingüística de los hablantes, por otro (Gómez Capuz, 2005: 14).

Así, hay un contraste muy marcado entre la actitud laxa de Alba (1995c: 17), quien en su estudio sobre anglicismos léxicos en el español dominicano, parte del concepto de *anglicismo* como “todo elemento léxico proveniente del inglés, independientemente de que esté o no adaptado fonéticamente al español, o de que se encuentre o no registrado en el Diccionario de la Real Academia Española”, ya la excesivamente restrictiva de González Martínez (2002: 47), quien en su investigación sobre el léxico disponible de Cádiz, considera extranjerismos “aquellos vocablos procedentes de otras lenguas que no están presentes en la última edición del *DRAE* de 1992”.

Ha sido López Morales (1999b) el que ha zanjado la cuestión de la clasificación de los anglicismos, en su estudios del léxico disponible de Puerto Rico. En este, establece una tipología que ha sido seguida por la mayoría de los estudios posteriores, en la que distingue:

a) *préstamos*, voces plenamente adaptadas al español, tanto fonética como morfosintácticamente, o que, sin estar adaptadas, tienen un índice de disponibilidad alto (por encima del 0,01) o al menos superior a sus competidores patrimoniales (lo que hace pensar en un uso no ocasional);

b) *calcos*, construcciones que siguen “un modelo, generalmente semántica o morfemático”, de otra lengua (en su caso, creaciones paralelas al modelo inglés);

c) *extranjerismos*, que no están adaptados al español ni se caracterizan por tener un índice de disponibilidad elevado (lo que significa que tienen un empleo ocasional).

En la selección del corpus es fundamental fijar el límite temporal a partir del cual podemos hablar de préstamos recientes, esto es, “los que realmente deben ser objeto de análisis lingüístico, ya que se encuentran en el proceso de integración” (Gómez Capuz, 2005: 23). Al respecto, Bartol Hernández (2005) y Samper Hernández (2012) limitan el análisis a los extranjerismos incorporados después de 1900, para lo que toman como referencia la información contenida en el CORDE, en tanto que Castañer Martín (2008) se inclina por incluir aquellos extranjerismos que se incorporan al español a partir del primer cuarto del siglo XX (fijando el límite en la 15ª edición del diccionario académico, por lo que se justifica la ausencia de voces como *tranvía* o *yate*, recogidas en ediciones anteriores del *DRAE*).

Entre los primeros trabajos en el mundo hispánico, cabe destacar las investigaciones de Alba (1995c, 1999, 2000a) en la República Dominicana, la comparación

entre los anglicismos presentes en los léxicos disponibles de Zamora y la República Dominicana de Bartol Hernández (1998), y el estudio cuantitativo que lleva a cabo López Morales (1999b) sobre los anglicismos léxicos presentes en el *Léxico disponible de Puerto Rico*.

En el estudio de Alba (1999) sobre anglicismos en el léxico disponible del español dominicano, la proporción de anglicismos alcanza el 6,06%. Los ámbitos que asimilan más voces de origen inglés son ‘Medios de transporte’, ‘Juegos y diversiones’, ‘El vestido’ y ‘Los muebles de la casa’; por el contrario, los anglicismos casi se hayan ausentes en ‘Animales’, ‘El cuerpo humano’ y ‘Trabajos del campo’. Según Alba (1999: 85), “los centros más influidos por el inglés constituyen áreas semánticas abiertas que se refieren a actividades sociales o de comunicación en las que el dominio y el prestigio anglosajón son indiscutibles”. Dado que los resultados indican que los informantes de nivel alto son los que más emplean los anglicismos, Alba “concluye que el uso de estos términos constituye un fenómeno prestigioso en la comunidad de habla dominicana” (Samper Padilla *et al.*, 2003: 120).

Bartol Hernández (1998) compara, cuantitativa y cualitativamente, los anglicismos que aparecen en los léxicos disponibles de los estudiantes preuniversitarios de la República Dominicana (Alba, 1995c) y Zamora (Gallos Camacho, 1998a). Concluye que “el léxico disponible de los alumnos de la República Dominicana presenta un número mucho mayor de anglicismos que el de los alumnos zamoranos”<sup>17</sup> (Bartol Hernández, 1998: 123). Cualitativamente, los anglicismos utilizados por los dominicanos son “crudos, chocantes, propios de una cultura con una influencia muy directa del inglés” (*baby doll, bartender, blower, car wash*, etc.), o adaptaciones de palabras inglesas (*escrín, greifrú, suiche, polocher*, etc.). A todo ello hay que unir que “muchos de los anglicismos dominicanos tienen un índice de disponibilidad más alto que sus correspondientes palabras españolas” (Bartol Hernández, 1998: 123). Los centros de interés más permeables a los anglicismos en las dos regiones son ‘Juegos y diversiones’, ‘Medios de transporte’ y ‘El vestido/La ropa’.

Por su parte, en el estudio que López Morales (1999b) lleva a cabo sobre los anglicismos en Puerto Rico, cifra en el 8,1% el porcentaje de anglicismos en el léxico disponible puertorriqueño. Esta cantidad es ciertamente superior a los porcentajes registrados en la República Dominicana y Zamora. Se aprecia, por lo tanto, una mayor influencia de la lengua inglesa en el léxico empleado en Puerto Rico. Como en la República Dominicana y Zamora, los centros de interés más influidos por la moda, la publicidad y la tecnología son los que presentan más anglicismos: ‘La ciudad’, ‘Alimentos’, ‘Medios de transporte’, ‘La ropa’ e ‘Iluminación

---

<sup>17</sup> Según los datos presentados por Bartol Hernández (1998: 101), el porcentajes medio en Zamora es 1,8% y en la República Dominicana 3,8%.

y Aire acondicionado'. Los centros de interés ligados a las facetas más tradicionales de la vida son en los que menos anglicismos aparecen: 'El campo', 'Trabajos del campo y del jardín', 'Objetos situados en la mesa para la comida', 'Animales' y 'La casa'.

Samper Padilla y Samper Hernández (2009: 871) clasifican los estudios realizados posteriormente en los siguientes grupos:

- Trabajos cuyo objeto de estudio es el conjunto de extranjerismos sin distinción de origen: Paredes García (2001b), Carrera y Bradley (2004), Bartol Hernández (2005) y Castañer Martín (2008);

- Trabajos que ciñen su objeto de estudio a los vocablos de origen anglosajón: González Martínez y Orellana (2006), Bellón Fernández (2004–2005 y 2007) y Samper Hernández (2012);

- Trabajos que abarcan la presencia de anglicismos (o extranjerismos en general) en todos los centros de interés (pues se persigue clasificar los diferentes ámbitos nocionales según su grado de permeabilidad): Paredes García (2001b), Bartol Hernández (2005), González Martínez y Orellana (2006) y Bellón Fernández (2007);

- Trabajos que estudian un campo léxico determinado: Bellón Fernández (2004–2005), Castañer Martín (2008) y Samper Hernández (2012), quienes seleccionan centros de interés representativos ('Medios de transporte' en un caso y 'La ropa' en los otros dos).

Según Samper Padilla y Samper Hernández (2009: 876), en todas las investigaciones realizadas los centros de interés que recogen mayor cantidad de anglicismos son 'Medios de transporte', 'Juegos y distracciones' y 'La ropa'. 'La ciudad' es otro centro de interés que se manifiesta como otro dominio que favorece la incorporación de voces de procedencia anglosajona. Frente a estos, otros dominios léxicos son poco propicios a la presencia de términos foráneos. Entre ellos sobresale 'Partes del cuerpo', que en muchas comunidades rechaza el anglicismo de modo absoluto. Otros dos campos léxicos con un escaso número de anglicismos son 'Trabajos del campo y del jardín' y 'El campo'.

El Cuadro 6.2, tomado de Samper Padilla y Samper Hernández (2009: 877), presenta los anglicismos más disponibles del campo léxico 'La ropa' en las comunidades de habla de República Dominicana, Puerto Rico, Zamora, Soria, Córdoba, Cádiz y Gran Canaria. En el listado se pueden advertir algunos anglicismos compartidos por todas las comunidades, con carácter panhispánico: *top*, *pijama*, *suéter*, *panti*, *biquini*, *bermudas*, *polo*, *minifalda* y *short*. Por su parte, vocablos (como *body* y *slip*) son exclusivos de los léxicos españoles, en tanto que sólo aparecen en el Caribe *brassiere*, *t-shirt*, *panty-hose*, *polo shirt* y *mediapanties*.

**Cuadro 6.2. Posición que ocupan los anglicismos más disponibles en distintas comunidades de habla**

Fuente: Samper Padilla y Samper Hernández (2009: 877)

	Rep. Domin.	Puerto Rico	Zamora	Soria	Córdoba	Cádiz	Gran Canaria
bermudas	22	791	30	43	34	36	55
biquini	71	44	53	58	53	41	47
blazer	64	50	–	–	302	125	–
body	–	–	29	59	39	34	87
brassiere	9	9	–	–	–	–	–
jacket	41	32	–	–	–	–	–
jeans	43	–	–	168	96	144	212
jersey	122	–	3	2	4	3	15
mediapanties	38	136	–	–	–	–	–
minifalda	33	56	27	65	58	59	77
niqui	–	–	–	33	99	204	205
panti	6	8	41	34	47	60	45
panty-hose	55	49	–	–	–	–	–
pijama	48	36	34	37	30	28	37
polo	170	69	24	31	32	47	56
polocher (polo shirt)	11	89	–	–	–	–	–
pulóver	–	82	–	–	257	111	7
rebeca	–	–	–	154	23	32	29
short	19	45	58	55	49	50	130
slip	–	–	61	53	52	40	129
sudadera	–	41	69	17	17	20	61
suéter	14	14	37	42	51	48	42
tenis	15	31	–	–	238	33	122
top	149	300	77	39	60	66	35
t-shirt	47	19	–	–	–	–	–

Los extranjerismos también son abundantes en el centro de interés ‘Medios de transporte’. En palabras de Castañer Martín (2008: 61), “se trata [...] de un campo léxico en el que la influencia extranjera ha sido considerable en todas las épocas de la lengua española y en el que algunos préstamos –incluso recientes– están tan arraigados y asimilados que se nos escapan con facilidad”. Los anglicismos que censa Bartol Hernández (1998: 110) en Zamora y la República Dominicana se recogen en el Cuadro 6.3.

**Cuadro 6.3. Anglicismos en el centro ‘Medios de transporte’**  
**Fuente: Bartol Hernández (1998: 110)**

Zamora			República Dominicana		
Posición	Vocablo	ID	Posición	Vocablo	ID
17	tranvía	0,207	24	yate	0,065
34	yate	0,069	25	jeepeta	0,064
45	sidecar	0,037	27	skateboard	0,059
47	tráiler	0,036	30	jet	0,058
57	ferry	0,023	32	jeep	0,056
85	trolebús	0,013	59	jumbo	0,013
86	auto stop	0,012	67	tranvía	0,011
90	tabla de surf	0,011	69	tráiler	0,010
113	tabla de windsurf	0,009	72	jetski	0,009
124	aircraft	0,005	77	subway	0,009
141	car	0,004	83	concord	0,007
			90	three Wheel	0,005
			95	four Wheel	0,005
			99	ferry	0,004
			107	go cart	0,003
			132	windsurf	0,002
			160	skylab	0,001
			167	surfing	0,001

Llama la atención la gran cantidad de extranjerismos que Castañer Martín (2008: 61–63) halla en el corpus de Aragón: *autobús, avión, motocicleta, taxi, furgoneta, helicóptero, avioneta, tractor, patinete, camioneta, sidecar, autocar, ala delta, tráiler, limusina, esquí, moto acuática, minibús, ferry, parapente, caravana, tándem, zepelín, autoestop, trolebús, teleférico, kart, concorde, hidroavión, autobús de línea, tanque, turismo, portaaviones, autobús urbano, scooter, microbús, ranchera, moto de nieve, hovercraft, moto de agua, jet, vespa, mountainbike, tabla de surf, vaporetto, buga, jumbo, land rover, campera, vespino, zodiac, quad, vuelo chárter, snowboard, dumper, kayak, reactor, isocarro, poni, monorraíl, avión de combate, jeep, helitaxi, mobilette, tabla de surf, Boeing, bici de cross, bus interurbano, avión de mercancías, pasquali, moto de competición, airbus, roulotte, TVG, roller, autobús de dos pisos, tractocamión, rickshaw, moto de carreras, camioneta de tres ruedas, telecabina, tirolina, minicar, portahelicópteros, tanqueta, catamarán, optimist, moto de cuatro ruedas, skateboard*. La presencia de términos extranjeros es superior a la de otras comunidades: un total de 89 unidades. Se cuentan, además, 36 extranjerismos entre

los primeros 100 vocablos del centro (y de ellos, 8 figuran entre los 20 más disponibles).

Finalmente, el tercer campo léxico que ofrece un elevado número de extranjerismos es ‘Juegos y distracciones’. Revisando los extranjerismos encontrados en los estudios<sup>18</sup> de Almería, Cádiz, Aragón, Zamora y República Dominicana, aparecen sobre todo anglicismos, referentes al mundo del deporte. Aunque algunos de ellos cuentan con un término correspondiente en español, como *fútbol/balompíe*, *voleibol/balonvolea*, *básquet(bol)/baloncesto*, la inmensa mayoría carece de una palabra castellana equivalente (*bádminton*, *críquet*, *footing*, *full-contact*, *golf*, *hockey*, *kick-boxing*, *kung fu*, *motocross*, *mountain bike*, *paddle*, *ping ball*, *rally*, *rugby*, *snowboard*, *squash*, *surf*, *surfing*, *taekwondo*, *tenis*, *trial*, *waterpolo*, *windsurf*). La mayoría de las 22 coincidencias habidas entre los anglicismos usados en Zamora y en la República Dominicana pertenecen a esta categoría.

Son frecuentes también los extranjerismos que designan diversas actividades de ocio (*aerobic*, *bar*, *bingo*, *boy scout*, *club*, *camping*, *cómico*, *disc jockey*, *fitness*, *folklore*, *hobby*, *jazz*, *peep-show*, *picnic*, *pub*, *puzle*, *sex shop*, *strip-poker*, *striptease*, *tobogán*). Algunas palabras aluden a juegos de mesa muy conocidos: *backgammon*, *póquer*, *puzzle*. A ellos se pueden sumar –si los aceptamos como plenamente lexicalizados– algunos juegos de mesa, electrónicos o de otro tipo (*monopoly*, *nintendo*, *play station*, *playmobil*, *scalextric*, *scrabble*, *tetris*, *trivial*)” Castañer Martín (2008: 55). Como casos curiosos, menciona Bartol Hernández (1998: 113) el doblete *footing-jogging*, la primera preferida en la Península Ibérica, la segunda en el Caribe; igualmente representa un caso especial para Mateo García (1998: 123) la forma *puenting*, un híbrido con sufijo inglés, ejemplo clásico de interferencia lingüística.

Los ejemplos aquí anotados vienen a dar la razón a González Martínez y Orellana Ramírez (2006: 19), cuando afirman que “cualitativamente, los anglicismos recogidos son, en su mayoría, designaciones que han llegado a nuestra lengua inseparablemente del objeto, de la realidad, por ellos designado (*type*, *tenis*, *fútbol*, *voleibol*, *disc jockey*, etc.), es decir, que han venido a llenar un hueco y se han popularizado, en la mayoría de los casos, junto con la realidad designada”. Conviene tener presente una perspectiva evolutiva, ya que, como afirma Paredes García (2001b: 567), “la consideración de un término como extranjerismo va ligado a una determinada sincronía: un vocablo extranjero se introduce en una lengua en un momento histórico y las reacciones se suscitan sólo durante un tiempo”. Así, muchos de los préstamos léxicos plenamente integrados en el sistema lingüístico español fueron una vez extranjerismos rechazados por académicos y personalidades de prestigio. Sin ir más lejos, la palabra *rugby*, a la que Mateo García (1998:

<sup>18</sup> Las investigaciones correspondientes son la de Almería (Mateo García, 1998: 122), Cádiz (González Martínez y Orellana Ramírez, 2006: 14–15), Aragón (Castañer Martín, 2008: 55), y Zamora y República Dominicana (Bartol Hernández, 1998: 111–113).

123) se refería como “deporte incompresiblemente no incluido por el *DRAE*” en la edición de 1992, figura ya en la 22.<sup>a</sup> edición, publicada en 2001 con la información “Voz inglesa, de *Rugby*, escuela pública de Warwickshire, en Inglaterra, donde se inventó”, antecediendo a su definición.

## 6.5. El léxico peculiar: los dialectalismos

En la introducción que Hernández Cabrera y Samper Padilla (2007: 287) hacen a su estudio sobre el “Léxico regional en los materiales de disponibilidad léxica” de Gran Canaria, se señala que los términos característicos del español de las Islas Canarias que figuran en las listas de léxico disponible de Gran Canaria son “un reflejo de la idiosincrasia del léxico canario, que, tanto por su situación geográfica como por las distintas contribuciones que han influido en su gestación y desarrollo, manifiesta una serie de características que le otorgan personalidad propia dentro del conjunto dialectal hispánico”. Si se generaliza esta afirmación a cualquier comunidad de habla, se colige que el estudio de los dialectalismos de un conjunto de léxico disponible mostrará el léxico peculiar y distintivo de esa comunidad de habla.

Los estudios sobre los dialectalismos que aparecen en el léxico disponible han ido ganando en interés de manera progresiva. Así, los trabajos publicados hasta 2014 cubren las siguientes áreas:

- Andalucía. En esta comunidad, González Martínez (1999, 2002) investiga los andalucismos en el léxico disponible de Cádiz; Mateo García (1998: 99–125) se ocupa desde el punto de vista cualitativo de lo que considera “léxico específico de la comunidad de habla” de Almería; Prado Aragonés (2009) estudia los dialectalismos en el léxico disponible de Huelva en relación con el centro de interés ‘Alimentos y bebidas’;

- Canarias. Todos los trabajos parten del léxico disponible de Gran Canaria y del equipo investigador de la Universidad de Las Palmas: se ocupan del nivel preuniversitario, Hernández Cabrera y Samper Padilla (2003), Samper Padilla y Hernández Cabrera (2006) y Hernández Cabrera y Samper Hernández (2005), en tanto que Samper Hernández (2005b) cuantifica los dialectalismos en el léxico disponible de escolares grancanarios;

- Castilla-La Mancha. Hernández Muñoz (2006: 365–402) que dedica un capítulo de su tesis al estudio de los dialectalismos en el léxico disponible de esta comunidad; Hernández Muñoz (2009b) analiza la variación léxica del léxico disponible en función de las zonas dialectales de Castilla-La Mancha;

- Aragón. Arnal Purroy (2008b) recoge los dialectalismos en el léxico disponible de los jóvenes aragoneses; poco después la misma Arnal Purroy (2009) estudia la vitalidad de los dialectalismos en las distintas áreas lingüísticas de Aragón a partir de los materiales de disponibilidad léxica recogidos;

- Navarra. Taberero Sala (2008b) clasifica los regionalismos que aparecen en el léxico disponible de esta comunidad;
- Chile. A. Valencia (2005) investiga los dialectalismos de este país sudamericano;
- Estudios comparativos. Llorente Pinto (2005) contrasta el léxico disponible y el léxico dialectal en la provincia de Ávila; Borrego Nieto y Fernández Juncal (2002) aplican la metodología de los estudios sobre disponibilidad léxica para determinar el grado de dialectalidad de la mitad oriental de Cantabria; Samper Padilla, Hernández Cabrera y Bellón Fernández (2003: 1092–1094) establecen una clasificación de las divergencias dialectales observadas al comparar los léxicos disponibles de Gran Canaria y Córdoba.

Es cierto que “las pruebas de disponibilidad no propician la activación de respuestas dialectales” (Arnal Purroy, 2008b: 18), ya que lo que se persigue en esas pruebas es recoger la norma léxica adulta, integrada, sobre todo, por voces propias del español general, compartidas por los léxicos disponibles de los distintos geolectos, “que justificarían la unidad esencial de una lengua tan ampliamente extendida como la española” (Samper Padilla, 1999: 551). Por ello, para recoger ese léxico estandarizado, o como señalan Borrego Nieto y Fernández Juncal (2002: 173), –altamente formalizado–, se realizan pruebas escritas en un centro de enseñanza a jóvenes de 17 ó 18 años que cursan 2º de Bachillerato, con algunas áreas temáticas que, *a priori*, no parecen ser muy rentables en cuestiones dialectológicas (‘La ciudad’ o ‘Medios de transporte’, por ejemplo).

Ello hace que la proporción de léxico dialectal en los correspondientes inventarios de disponibilidad léxica sea escasa: en Aragón, 280 dialectalismos suponen el 2,87% del total de vocablos disponibles (Arnal Purroy, 2008b: 29); en la provincia de Cádiz, 89 dialectalismos representan el 1,34% (González Martínez, 1999: 192); en Castilla-La Mancha, 162 dialectalismos equivalen al 1,27% del léxico disponible (Hernández Muñoz, 2006: 397–398). Tan solo en Gran Canaria la densidad del léxico dialectal es mayor, con 70 canarismos en el área temática de ‘Los alimentos’ (Hernández Cabrera y Samper Padilla, 2003: 345) y 106 en el centro de interés ‘El campo’ (Hernández Cabrera y Samper Hernández, 2005), así como 383 canarismos –un 5,35%– entre los alumnos de primaria y de ESO de Gran Canaria (Samper Hernández, 2005: 1070).

A pesar de lo exiguo de los resultados anteriores, como recuerda Alba (2000: 76, *apud* Hernández Cabrera y Samper Padilla, 2003: 340), “no nos encontramos con unos elementos léxicos buscados intencionadamente, sino con términos que afloran de forma espontánea en la encuesta, listos para ser empleados en el momento oportuno”. “Al tratarse de un léxico aportado por hablantes muy jóvenes, puede inferirse directamente el futuro inmediato de tales voces en

las distintas comunidades” (Samper Padilla *et al.*, 2003: 119)<sup>19</sup>, dependiendo el grado de vitalidad de cada término de su índice de disponibilidad.

De otra parte, según Borrego Nieto y Fernández Juncal (2002: 297), “la mayor o menor presencia de dialectalismos dentro del léxico disponible” sirve tanto para determinar el grado de dialectalidad de una comunidad como para identificar de forma rápida las zonas dialectales en un territorio, como así hacen en Aragón, Arnal Purroy (2009), en Castilla-La Mancha, Hernández Muñoz (2009b), y en Navarra, Taberero Sala (2008b). Todo ello hace coincidir en los estudios sobre léxico disponible el análisis de la variación social y la investigación dialectal, en una tendencia que Moreno Fernández (2004) denomina *dialectología social*.

Llevar a cabo un estudio acerca de dialectalismos o regionalismos en el léxico disponible no es tarea fácil, ya que en su etapa inicial el estudioso se topa con un problema teórico fundamental: la propia delimitación del objeto de estudio, por “la dificultad de determinar qué es exactamente un regionalismo” (Hernández Muñoz, 2006: 369), “de atribuir carácter dialectal a una determinada palabra” (Arnal Purroy, 2008b: 19). La mayor o menor restricción del concepto de regionalismo es determinante en el número de dialectalismos obtenidos para cada comunidad (Taberero Sala, 2008b: 812).

Un criterio más laxo es el que utilizan González Martínez (1999), que considera andalucismos los que se presentan como tales en el *DRAE* y los que aparecen en el vocabulario de Alcalá Venceslada y en otras fuentes bibliográficas; también es amplio el criterio de Hernández Cabrera y Samper Padilla (2003), quienes se basan en el *Diccionario diferencial del español de Canarias* y el *Tesoro lexicográfico del español de Canarias*. En este último se incluyen las voces propias y

aquellas del fondo patrimonial empleadas particularmente en las islas, los arcaísmos conservados, los vocablos que presentan cambio de significado y usos figurados, las locuciones y frases hechas, algunas unidades léxicas con una configuración sintagmática peculiar y también aquellas voces que, aun no siendo desconocidas en el español estándar peninsular, ofrecen una mayor frecuencia de uso en las islas (Hernández Cabrera y Samper Padilla, 2003: 341).

De otro lado, en Aragón, Arnal Purroy (2008b: 20) parte de un criterio más restrictivo, según el cual considera “como dialectalismos, regionalismos o, en concreto, aragonesismos las voces propias de Aragón que no existen o presentan alguna diferencia frente al español general o las variedades regionales situadas fuera de la zona oriental peninsular”. Como se ve, este planteamiento concuerda con criterios como la implantación territorial y el arraigo propuestos por Taberero Sala (2008b: 813) para la clasificación de los regionalismos<sup>20</sup>.

<sup>19</sup> Esta aseveración es igualmente válida para fijar la vitalidad de los extranjerismos.

<sup>20</sup> La autora sigue a Frago Gracia (1999: 143n) y Saralegui y Taberero (2002: 22–23).

De cualquier forma, todos los investigadores toman como base para la delimitación del corpus de dialectalismos la última edición del *Diccionario de la lengua española (DRAE)*, el *Diccionario de uso del español (DUE)* de María Moliner (1998) y el *Diccionario de español actual (DEA)* de Seco *et alii* (1999). La documentación de los dialectalismos es completada con obras de referencia regionales, estas con mayor o menor grado de rigor.

Gracias a esta labor de documentación se puede realizar una clasificación de los dialectalismos, según diferentes tipologías<sup>21</sup>. Así, Hernández Cabrera y Samper Padilla (2003), manejan la clasificación de los canarismos que Lorenzo, Morena y Ortega (1994: 12–13) contemplan en su *Diccionario de canarismos*:

a) Voces españolas comunes que presentan en Canarias acepciones características (*pooyo* ‘encimera’; *ropero*, con una significación más amplia en el español general);

b) Palabras derivadas de vocablos castellanos que dan lugar a nuevos términos mediante los procesos de sufijación, prefijación o composición (*habichuela* ‘judía verde’; *hierbahuerto* ‘hierbabuena’; *garbanza* ‘garbanzo mayor y de más calidad’);

c) Vocablos castellanos que han sufrido una evolución fonética peculiar en las islas (*jinojo* ‘hinojo’);

d) Voces que se consideran anticuadas o desusadas en la península —“arcaísmos”— (*bezo* ‘labio’; *damasco* ‘albaricoque’; *durazno* ‘melocotón’);

e) Voces procedentes de otras variedades regionales del español, sobre todo, andalucismos (*chocho* ‘altramuz’) y americanismos (*papa* ‘patata’; *maní* ‘caca-huete’; *guagua* ‘autobús’);

f) Palabras procedentes de la lengua aborigen hablada en las islas antes de la llegada de los castellanos (*gofio*);

g) Palabras tomadas de otras lenguas peninsulares, en especial, portuguesismos (*millo* ‘mijo’; *gaveta* ‘cajón corredizo’);

h) Préstamos de lenguas no peninsulares, como anglicismos (*queque* ‘cake’; *pulóver* ‘jersey’).

Hernández Cabrera y Samper Padilla (2003: 343) distinguen tres grandes grupos de centros de interés, según la presencia del léxico dialectal en los listados canarios:

a) Centros de alta densidad (por encima de 7,5%): 05 ‘Alimentos y bebidas’, 11 ‘El campo’ y 14 ‘Animales’;

---

<sup>21</sup> Taberero Sala (2008b: 814–817) en Navarra emplea una tipología con una división básica a partir del tradicional reparto lingüístico de Navarra en vascuence y romance, con términos de filiación vasca (*chacolí*, *chamarra*, *chapela*, *chistorra*, *cocochas*, *gabarra*, *pacharán*, *zulo*, etc.), y términos de filiación romance (*calderete* ‘guiso de conejo’, ‘caldero’; *canaleta* ‘canalón; *cuto* ‘cerdo’; *tajador* ‘sacapuntas’, etc.).

b) Centros de densidad media (entre 3% y 6%): 04 ‘Los muebles de la casa’, 06 ‘Objetos colocados en la mesa para la comida’, 07 ‘La cocina y sus utensilios’, 12 ‘Medios de transporte’, 13 ‘Trabajos del campo y del jardín’;

c) Centros de baja densidad (por debajo de 3%): 01 ‘El cuerpo humano’, 02 ‘La ropa’, 03 ‘Partes de la casa’, 08 ‘La escuela: muebles y materiales’, 09 ‘Iluminación, calefacción y medios de airear un recinto’, 10 ‘La ciudad’, 15 ‘Juegos y distracciones’, 16 ‘Profesiones y oficios’.

En el caso de Aragón, solo el centro de interés ‘Trabajos del campo y del jardín’ se sitúa entre los de alta densidad, con un 8,04% de aragonesismos, pertenecen al grupo de densidad media cinco centros de interés: ‘El campo’, ‘Partes de la casa’, ‘La cocina y sus utensilios’, ‘Alimentos y bebidas’ y ‘Los animales’, citados en orden decreciente; en tanto que los once restantes ofrecen una densidad muy baja de dialectalismos, que oscila entre el 2,26% del centro ‘Iluminación, calefacción y medios de airear un recinto’ y las áreas temáticas que, como ‘Partes del cuerpo’, ‘La ropa’, ‘La ciudad’, ‘Profesiones y oficios’ y ‘Los colores’, ni siquiera alcanzan el 1% de elementos dialectales (Arnal Purroy, 2008b: 31).

La misma autora compara los centros de interés con mayor densidad de dialectalismos en las sintopías de Aragón, Gran Canaria, Cádiz y Castilla-La Mancha (ver Cuadro 6.4). Como explican Hernández Cabrera y Samper (2003: 343), los centros de interés que arrojan la mayor proporción de léxico regional son los que “están más apegados al medio y a las tradiciones locales”, aunque la fuerza del medio y de la tradición difiere de unas comunidades a otras: en las regiones meridionales lo peculiar y lo específico de cada territorio gana fuerza en la gastronomía, en tanto que en Castilla y Aragón es el mundo agrícola el que cobra protagonismo en este léxico dialectal. Sea como fuere, en las cuatro sintopías comparadas, aparecen como elementos recurrentes ‘el campo’ en solitario o en relación con la agricultura –‘Trabajos del campo y del jardín’–, ‘La casa’ en relación con las partes de esta o con los muebles, y el mundo de la cocina por medio del espacio donde se cocina y los utensilios necesarios para ello, de los objetos que se colocan sobre la mesa a la hora de la comida, o de los mismos alimentos y bebidas que se toman.

**Cuadro 6.4. Comparación de los centros de interés con mayor densidad de dialectalismos en distintas regiones hispánicas. Fuente: Arnal Purroy (2008b: 31)**

Posición	Aragón	Gran Canaria	Cádiz	Castilla-La Mancha
1	2	3	4	5
1º	‘Trabajos del campo y del jardín’	‘Alimentos y bebidas’	‘Alimentos y bebidas’	‘Trabajos del campo y del jardín’
2º	‘El campo’	‘El campo’	‘La ropa’	‘La cocina y sus utensilios’

1	2	3	4	5
3º	‘Partes de la casa (sin los muebles)’	‘Los animales’	‘Los animales’	‘Los muebles de la casa’
4º	‘La cocina y sus utensilios’	‘Los muebles de la casa’	‘Partes de la casa (sin los muebles)’	‘Objetos colocados en la mesa...’

Por lo demás, hay que hacer notar que el campo 01 ‘Partes del cuerpo’ rechaza tanto el dialectalismo como el anglicismo, con lo que se confirma como el área léxica con una mayor proporción de léxico panhispánico (Samper Padilla *et al.*, 2003: 123).



## VII. LA COMPARACIÓN DE LOS LÉXICOS DISPONIBLES

### 7.1. La norma léxica del español peninsular

#### 7.1.1. La delimitación de la norma léxica según el léxico disponible

La publicación de los diccionarios de léxico disponible de cada comunidad de habla, con ser el objetivo principal del Proyecto Panhispánico de Disponibilidad Léxica, no es el objetivo final del mismo. El fin último del Proyecto es la elaboración de un diccionario del léxico disponible del español. Como pone de manifiesto Paredes García (2012: 79), gracias al vigor y empuje que han adquirido las investigaciones en España, con estudios en más de cincuenta provincias españolas, muchas de ellas con resultados ya concluidos, la publicación del *Diccionario del léxico disponible de España* es un objetivo al alcance de la mano.

Se tratará, en este caso, de un diccionario del léxico temático realmente utilizado por los hablantes españoles, un repertorio tanto de la norma léxica española como de los dialectalismos, extranjerismos, marcas lexicalizadas, tecnicismos, vulgarismos, etc. Con todo, el conjunto léxico más importante en este diccionario de disponibilidad léxica será el léxico más disponible, por ser el más usado a la vez que el más representativo y de referencia neutra. La norma léxica común de la lengua española estará conformada por este léxico disponible común a las diferentes comunidades de habla, y que se definirá a partir de la comparación de los léxicos disponibles de las diferentes comunidades hispánicas. Por lo tanto, la comparación, cuantitativa y cualitativa, de los listados correspondientes facilitará “el establecimiento y la confirmación del vocabulario general del español en las áreas semánticas que se analizan” (Samper Padilla *et al.*, 2003: 112).

A tal fin, una cuestión de suma importancia es la de cómo delimitar el conjunto léxico que debe compararse. No pueden utilizarse todos los datos recogidos en las encuestas ya que:

A partir de un determinado punto en los listados de disponibilidad aparece un conjunto de palabras que han sido aportadas por muy pocos encuestados, que, además, las han incluido en sus hojas de respuesta en posiciones poco relevantes. Podríamos encontrarlos entonces con que no estamos comparando hechos de norma, sino fenómenos particulares del habla de un grupo de informantes o incluso de un único informante (Samper Padilla, 1999: 554).

Como señala Bartol Hernández (2008: 207), “no tener en consideración estos hechos puede provocar una visión distorsionada de la realidad, como sucedió en el caso conocido de las comparaciones realizadas por el profesor mexicano López Chávez (1992, 1995a)”. Este profesor, que utilizó la totalidad de los datos de los léxicos disponibles de Madrid, Gran Canaria, República Dominicana y Puerto Rico, en las primeras comparaciones intermuestrales en el mundo hispánico, halló que los índices de compatibilidad entre las distintas zonas geográficas evaluadas eran «insospechadamente bajos», ya que en ningún caso se rebasaba el 25% de coincidencia en el vocabulario. Estos resultados eran absolutamente ilógicos, puesto que sugerían la errónea conclusión de la existencia de problemas de comprensión entre los miembros de los dos conjuntos comparados.

Habida cuenta de la necesidad de establecer una división entre los dos subconjuntos de vocablos – el que se compara y el que no –, diferentes han sido las soluciones adoptadas por los investigadores a la hora de fijar el punto de corte entre ambos subconjuntos. En su estudio sobre *El léxico disponible de la provincia de Córdoba*, Bellón Fernández (2011: 153–199) realiza un exhaustivo análisis de las diferentes propuestas implementadas. Los criterios de corte para las comparaciones de los listados de léxico disponible han sido:

- Una cantidad arbitraria y fija de vocablos de cada centro de interés (también denominado rango). Siguiendo el criterio fijado por Gougenheim *et al.* (1967: 165–181), investigadores como Carcedo González (1999c, 2000a, 2000b, 2000c, 2001), Mateo García (1995) en los análisis cualitativos atienden solo los 20 primeros vocablos de las listas de disponibilidad. Según los objetivos de la investigación, algunos autores han reducido la cifra a 10 vocablos (Carcedo González, 1999c, 2000b, 2000c; Mateo García, 1998; Hernández Muñoz, 2004), o la han incrementado a 25 vocablos (Alba, 1996), 50 vocablos (Alba, 1998; 2000b; Carcedo González, 2000c; Martínez Olmos, 2007; Trigo Ibáñez, 2011), o incluso hasta 100 vocablos (Alba, 1999, 2000a; Valencia, 1998b, 2000; Valencia y Echeverría, 1998).

- El 75% de la frecuencia acumulada de las listas de disponibilidad léxica. El establecimiento de este porcentaje permite apreciar las diferencias de comportamiento de los diferentes centros de interés y también proporciona un buen método para realizar comparaciones entre distintos trabajos. Este ha sido el método utilizado por Samper Padilla (1999) para confrontar los léxicos disponibles de Puerto Rico y Gran Canaria, González Martínez y Orellana Ramírez (2000) para el cotejo

de los léxicos disponibles de Cádiz y Zamora, y Samper Padilla, Hernández Cabrera y Bellón (2003) para comparar los léxicos disponibles de Gran Canaria y de la provincia de Córdoba.

– Un valor del índice de disponibilidad léxica. Como recoge Bellón Fernández (2011: 173), fue Ruiz Basto (1987: 97) quien propuso “establecer un criterio de corte que identificara los vocablos que constituyen la norma y las subnormas de los distintos grupos sociales”, considerando *vocabulario activo* el conjunto de vocablos cuyo índice de disponibilidad es superior a 0,1; y *vocabulario pasivo*, el conjunto de vocablos que están por debajo de este límite<sup>1</sup>. También han situado el punto de corte en un valor igual o superior a 0,1 Carcedo González (2000a) y Gómez Devís (2004). No obstante, no hay acuerdo con respecto a cuál es el valor de corte óptimo. Así, Bartol Hernández (2001), que en un principio era partidario de tener en cuenta para los cotejos cualitativos de léxico disponible solo aquellos vocablos cuyo índice de disponibilidad fuera, al menos, el 0,02, posteriormente (Bartol Hernández, 2008, 2011) se inclina por fijar el valor de corte en 0,01. Por su parte, tanto Bellón Fernández (2005, 2011) como Samper Hernández (2008) se decantan por el índice de disponibilidad 0,05.

De los tres criterios que se han propuesto –el rango (las 20, 50, 100 primeras palabras), la frecuencia acumulada, o el índice de disponibilidad–, Bartol Hernández (2001) entiende como el criterio más acertado el índice de disponibilidad (solo o relacionado con los otros dos), dado que es la medida propia de la disponibilidad. Apoya esta resolución Bellón Fernández (2011: 174), a quien le “parece razonable que, si estamos utilizando los listados de vocabulario disponible [...] en los que los vocablos se ordenan según su índice de disponibilidad, este sea, pues, el factor que hay que tomar en consideración a la hora de establecer un límite en la cantidad de términos extraídos de esas listas”.

### 7.1.2. Las comparaciones interdialectales y el léxico compartido

Como afirma López Morales (1999a: 20), “el análisis cuantitativo de la disponibilidad de una comunidad dialectal dada ofrece ya una buena descripción de esta parcela de su norma léxica. Sin embargo, donde estos estudios ofrecen su mejor y más valiosa contribución es en las comparaciones dialectales”.

En efecto, por medio de este procedimiento es posible conocer lo que de común y diverso tienen los léxicos disponibles de las diferentes zonas del ámbito hispanohablante. En este sentido, las comparaciones realizadas (Alba, 1998; Samper Padilla, 1999; González Martínez y Orellana Ramírez, 2000; Hernández

---

<sup>1</sup> Según Ruiz Basto (1987), para que un vocablo fuera considerado activo, debía cumplir una condición más: haber sido mencionado, al menos, por un 20% de los informantes.

Cabrera y Bellón Fernández, 2003) dentro del Proyecto Panhispánico “coinciden en el elevado grado de convergencia de las normas léxicas disponibles de los diferentes ámbitos hispánicos estudiados” (Carcedo González, 2000a: 153). Todos estos trabajos se caracterizan por comparar inventarios léxicos recogidos sobre criterios metodológicos comunes<sup>2</sup> y por delimitar<sup>3</sup> las unidades que se eligen para la comparación.

Como decimos, la relevancia de las convergencias en los estudios comparativos fue puesta de manifiesto en primer lugar por Orlando Alba (1998), quien llevó a cabo una comparación de las 50 primeras palabras de los léxicos disponibles de la República Dominicana, Puerto Rico, México, Madrid y Concepción en tres centros de interés: ‘El cuerpo humano’, ‘Medios de transporte’ y ‘Alimentos’. Alba halló que los índices de compatibilidad entre los distintos léxicos disponibles son superiores, en todos los casos, al 60%, a pesar de que algunos centros presentan mayor poder discriminador.

Por su parte, Samper Padilla (1999), limitando el vocabulario comparado a las palabras que suponen el 75% del índice acumulado en los listados de disponibilidad, establece que “el léxico coincidente [entre Gran Canaria y Puerto Rico] es, por tanto, abrumadoramente mayoritario [un 91%] y justifica la unidad de la lengua también en el terreno léxico” (Samper Padilla, 1999: 571). González Martínez y Orellana Ramírez (2000: 157), siguiendo las mismas pautas metodológicas, en su comparación de los léxicos disponibles de Cádiz y Zamora también concluyen que existe “un altísimo componente común a ambas sintopías”. Con igual metodología, Samper Padilla, Hernández Cabrera y Bellón Fernández (2003) obtienen un valor general de compatibilidad entre el vocabulario disponible de Gran Canaria y el de Córdoba del 56,07%, por lo que concluyen que el léxico compartido por los hablantes de Gran Canaria y Córdoba es abrumadoramente mayoritario.

### 7.1.3. Cotejo de cinco léxicos disponibles peninsulares

Teniendo en cuenta los principios teóricos metodológicos mencionados, nos disponemos a definir el léxico común a las normas léxicas disponibles de cinco comunidades de habla de España<sup>4</sup>. Los léxicos disponibles comparados y las in-

---

<sup>2</sup> Esta es la idea que subyace al Proyecto Panhispánico de Disponibilidad Léxica: mantener la homogeneidad de criterios que permita establecer comparaciones de tipo lingüístico, etnográfico y cultural.

<sup>3</sup> Recordemos que López Chávez (1992), por el contrario, se había servido de la totalidad de los listados, lo que le hacía obtener unos índices de compatibilidad entre Las Palmas y Puerto Rico de un exiguo 18,97%.

<sup>4</sup> Tomamos como modelo para la comparación la que realiza Carcedo González (2000: 153–211). De aquella se diferencia en los léxicos comparados. Los léxicos disponibles de hispanohablantes utilizados por este autor eran dos peninsulares –Cádiz (González Martínez, 1997) y Zamora

vestigaciones a las que estos pertenecen son: Asturias (Carcedo González, 2001), Castilla-La Mancha (Hernández Muñoz, 2006), Sevilla (Trigo Ibáñez, 2011), Valencia (Gómez Devís, 2004) y Zamora (Gallos Camacho, 1998a). Esta selección garantiza, por un lado, una muestra suficientemente representativa de la variación diatópica del español peninsular y, por otro, la comparabilidad, dada la uniformidad metodológica de estos estudios al hallarse encuadrados en el Proyecto Panhispánico de Disponibilidad Léxica.

El Cuadro 7.1 ofrece las frecuencias absolutas en vocablos para los dieciséis centros de interés tradicionales en los léxicos disponibles mencionados<sup>5</sup>. Las pequeñas diferencias de edición, así como el distinto tamaño de las muestras de esos trabajos, influyen en el volumen total de vocablos recogidos. La muestra más voluminosa en la de Castilla-La Mancha, con 680 informantes, hay tres muestras con valores dentro de la centena de 400 –Asturias (470), Valencia (465), Sevilla (400)–, en tanto que la muestra de Zamora cuenta solo con 100 informantes. De ahí las discrepancias numéricas que se advierten entre esos inventarios.

**Cuadro 7.1. Comparación por centros de interés del número de vocablos en cinco sintopías peninsulares**

Procedencia y nº de informantes → centro de interés↓	Asturias (470)	Castilla-La Mancha (680)	Sevilla (400)	Valencia (465)	Zamora (100)
1	2	3	4	5	6
01. Partes del cuerpo	336	366	354	339	178
02. La ropa	252	394	381	330	191
03. Partes de la casa	263	411	303	263	172
04. Muebles de la casa	252	451	361	308	208
05. Comidas y bebidas	589	781	653	561	355
06. Objetos sobre mesa para...	297	522	362	256	215
07. La cocina	383	718	551	372	319
08. La escuela	475	1131	653	473	302
09. Calefacción e iluminación	455	896	434	256	265
10. La ciudad	896	1272	1031	755	453
11. El campo	966	1510	1235	967	552

(Gallos Camacho, 1998a)– y dos americanos –Puerto Rico (López Morales, 1999a) y República Dominicana (Alba, 1995a)–. De todos estos, solo mantenemos el de Zamora, representativo de Castilla-León.

<sup>5</sup> Se comparan los vocablos correspondientes a los 16 centros de interés tradicionales. Tanto la investigación de Castilla-La Mancha como la de Valencia añaden un centro más. En Castilla-La Mancha ‘Inteligencia’, en Valencia ‘Colores’. Dado que estos centros no son objeto de la comparación, la cantidad de vocablos producidos en estos se ha descontado de la suma total.

Cuadro 7.1 (cont.)

1	2	3	4	5	6
12. Medios de transporte	234	470	420	280	208
13. Trabajos del campo...	469	1062	843	491	323
14. Animales	429	510	621	467	315
15. Juegos y distracciones	718	1097	1043	546	481
16. Profesiones y oficios	634	880	825	715	441
Total	7648	12471	10181	7379	4978

Ahora bien, al sustituir los valores absolutos por sus rangos correspondientes (ver Cuadro 7.2), se observa que ‘El campo’ aparece siempre como el centro de interés más productivo, ‘La ciudad’ se sitúa siempre en 2ª ó 3ª posición, y ‘Juegos y distracciones’ se halla entre las posiciones 2 y 5. Son estos los campos referenciales con más posibilidades evocadoras. A continuación, salvo en Castilla-La Mancha, ‘Profesiones y oficios’ se cuenta entre las cinco primeras posiciones. Una alta producción léxica se registra igualmente en ‘Comidas y bebidas’, entre las cinco primeras posiciones en tres de las investigaciones comparadas.

**Cuadro 7.2. Comparación de los rangos de los centros de interés según la producción de vocablos en los léxicos disponibles de cinco sintopías peninsulares**

Centro de interés	Asturias	Castilla-La Mancha	Sevilla	Valencia	Zamora
01. Partes del cuerpo	11	16	15	10	15
02. La ropa	14	15	12	11	14
03. Partes de la casa	13	14	16	14	16
04. Muebles de la casa	15	13	14	12	12
05. Comidas y bebidas	5	8	7	4	5
06. Objetos sobre mesa para...	12	10	13	16	11
07. La cocina	10	9	9	9	7
08. La escuela	6	3	6	7	9
09. Calefacción e iluminación	8	6	10	15	10
10. La ciudad	2	2	3	2	3
11. El campo	1	1	1	1	1
12. Medios de transporte	16	12	11	13	13
13. Trabajos del campo...	7	5	4	6	6
14. Animales	9	11	8	8	8
15. Juegos y distracciones	3	4	2	5	2
16. Profesiones y oficios	4	7	5	3	4

Por otra parte, los centros con un carácter mucho más concreto registran las producciones más bajas. De esta manera, con pequeñas variaciones entre los léxicos cotejados, por debajo de la posición de la posición 10 se encuentran ‘Objetos sobre la mesa para la comida’, ‘Medios de transporte’, ‘La ropa’, ‘Muebles de la casa’, ‘Partes del cuerpo humano’ y ‘Muebles de la casa’ y ‘Partes de la casa’<sup>6</sup>.

Como indicamos más arriba, es necesario delimitar las unidades léxicas que se comparan para dejar fuera de la comparación elementos atípicos y frecuentemente individuales (cf. Samper Padilla, 1999). Por consiguiente, efectuamos tal limitación, combinando los criterios del índice de disponibilidad y el rango. Así, de una parte, se seleccionan aquellas palabras que alcanzan un índice de disponibilidad igual o superior a 0,1 (i.d. +0,1). Éstas son las que, por alcanzar un mayor número de menciones entre los componentes de la muestra, así como por la posición más elevada que ocupan en esas listas, conforman el léxico altamente disponible: vocabulario activo. De otra parte, ampliamos el conjunto comparado a las 50 primeras unidades léxicas de cada centro de interés, pues, generalmente, se trata de lexemas básicos y con alta probabilidad de realización y, en cualquier caso, se trata de las unidades más representativas de cada conjunto.

**Cuadro 7.3. Número y proporción de vocablos de los léxicos disponibles comparados**

Léxicos comparados nº de informantes →	Asturias (470)	Castilla-La Mancha (680)	Sevilla (400)	Valencia (465)	Zamora (100)
Total de vocablos	7648	12471	10181	7379	4978
Vocablos con i.d. +0,1	469 (6,13%)	410 (3,29%)	422 (4,15%)	432 (5,85%)	417 (8,38%)
Vocablos en el rango 50	800 (10,46%)	800 (6,42%)	800 (7,86%)	800 (10,84%)	800 (16,01%)

El porcentaje comparado (véase Cuadro 7.3), aun siendo pequeño, es significativo, ya que se trata de hechos de norma. Aunque en las proporciones del léxico comparado se aprecia la influencia del distinto tamaño de las muestras, muy acusada en los extremos superior (Castilla-La Mancha) e inferior (Zamora), la cantidad de léxico (altamente disponible) que se incluye en el intervalo i.d. +0,1 es muy similar. Muestra de ello es la escasa diferencia habida entre Castilla-La Mancha –con 410 vocablos muy disponibles– y Zamora –con 417 vocablos.

<sup>6</sup> Compárese la similitud en la distribución de rangos con la descrita por Carcedo González (2000a: 155–156).

## 7.2. Análisis descriptivo de los centros de interés y su norma léxica

El léxico común a los conjuntos léxicos comparados viene delimitado, tanto por la composición léxica de los listados, que constituyen el objeto de estudio lingüístico, como por las soluciones particulares que los distintos investigadores toman al editar los materiales, esto es, por la intervención del investigador. Por consiguiente, a la hora de establecer el léxico compartido, hay que tomar en consideración ambos elementos: la naturaleza del objeto de estudio y la actuación del investigador.

En cuanto al objeto de estudio –el contenido de los listados léxicos– las palabras que conforman los centros de interés dependen de la propia estructura semántica del centro en cuestión, esto es, de los núcleos temáticos que se constituyen en su interior y, en consecuencia, del nivel de cohesión interna que muestra cada centro. Se trata, pues, de una cuestión de significado léxico, a la que también hay que sumar un amplio abanico de fenómenos lingüísticos de índole léxica, tales como la variación (diatópica, diastrática, diafásica), la importación léxica, la formación y creación de palabras, etc., que determinan igualmente el léxico disponible resultante.

En lo referente a la actuación del investigador, la tarea del lingüista de describir la realidad lingüística se ve influida por los criterios de edición. Y aquí no solo se trata de la aplicación de los criterios generales de edición del Proyecto Panhispánico, ya previstos y, en todo caso, uniformes y homogéneos en todos los trabajos. Lo que puede variar de una investigación a otra es la interpretación de cada investigador por las características de la comunidad estudiada pues, “es innegable que cada investigación presenta problemáticas distintas y, dado que se enmarcan en comunidades de habla diversas, deberán ofrecer soluciones particulares y coherentes con el entorno investigado” (Trigo Ibáñez, 2011: 42).

En consecuencia, considerando el efecto de ambos elementos, el cotejo de los léxicos disponibles es acompañado, *grosso modo*, del análisis descriptivo de las unidades léxicas que componen los centros de interés. Se da cuenta, así, por un lado, de la variedad temática dentro de los centros de interés por la relevancia de las distintas relaciones designativas, y por otro, de los fenómenos lingüísticos más destacados (categorías gramaticales, flexión gramatical, derivación, extranjerismos, marcas comerciales, etc.) que han supuesto la reflexión metodológica de los investigadores al llevar a cabo la tarea de lematización.

### 7.2.1. ‘Partes del cuerpo humano’

Este es un centro de interés que no presenta excesivas dificultades en el proceso de lematización, ya que se trata de una categoría semántica “bien definida”. En consecuencia, es un área temática de gran concreción, con un carácter muy

cerrado, por lo que presenta menor grado de dispersión o variedad de vocabulario, lo cual se corresponde con su índice de cohesión, el más alto de todo el conjunto de centros de interés (vid. 5.3.4). Estas características se manifiestan de manera muy clara en la coincidencia de los vocablos que ocupan las primeras posiciones en los léxicos comparados: son comunes las primeras diez unidades, y hasta 18 en la franja de disponibilidad +0,1 (ver Cuadro 7.4)<sup>7</sup>.

Dentro del campo léxico de las ‘partes del cuerpo’, pueden distinguirse en su estructura interna campos más restringidos, tales como las partes relacionadas con la cabeza (*barba, barbilla, bigote, boca, cabeza, cara, ceja, cerebro, diente, frente, labio, lengua, mandíbula, mejilla, muela, nariz, oído, ojo, oreja, pelo, pestaña, rostro...*), las extremidades (*antebrazo, brazo, codo, dedo, dedo del pie, hombro, mano, muñeca, muslo, pie, pierna, rodilla, tobillo, uña...*), el tronco<sup>8</sup> (*abdomen, barriga, cadera, cintura, culo, cuello, espalda, estómago, ingle, ombligo, panza, pecho, pubis, tórax, tripa, tronco, vientre...*), órganos y elementos internos (*arteria, corazón, hígado, hueso, intestino, músculo, órgano, páncreas, pulmón, riñón, sangre, vena...*), etc.

Este es un campo en el que aparecen numerosas formas en plural por la duplicidad y simetría del cuerpo humano o por la multiplicidad de elementos. En estos casos, se aplica la pauta general de neutralización de variantes meramente flexivas, reduciendo a singular vocablos como *brazos, músculos, ojos, pestañas*. De esta norma se excluyen los casos de distinción semántica, donde se consignan por separado las voces *pecho – pechos, tripa – tripas*, por constituir realidades distintas<sup>9</sup>. En cuanto a las formas derivadas, se mantienen los diminutivos lexicalizados como *barbilla, campanilla, espinilla, patillas, perilla*, que designan realidades distintas a sus formas simples.

Se unifican las lexías complejas documentadas en su forma compuesta y elidida, con ayuda del paréntesis, en casos como *columna (vertebral), vesícula (biliar), palma (de la mano), médula (espinal)*<sup>10</sup>, (*dedo*) *meñique*, etc.

Es este un centro con un gran número de tecnicismos (*esternocleidomastoideo, deltoides, líquido sinovial, metatarso, úvula...*), producto de su estrecha relación

<sup>7</sup> En este y en los demás cuadros comparativos, el fondo gris corresponde a las palabras con índice de disponibilidad inferior a 0,1.

<sup>8</sup> Véase el interesante trabajo de Llorente Pinto (2008), en el que la autora analiza los resultados de distintos léxicos disponibles (Aragón, Ávila, Burgos, Castilla-La Mancha, Ceuta, Huelva, Puerto Rico, República Dominicana, Salamanca, Soria, Valencia y Zamora) en el centro de interés ‘Partes del cuerpo humano’, referentes al tronco, para conocer cuáles son los vocablos más disponibles, y cuáles los más generales.

<sup>9</sup> No obstante, algunos investigadores unifican ambos términos (Alba, 1995a; Galloso Camacho, 1998a; López Morales, 1999a; Carcedo González, 2001; González Martínez, 2002; Bartol Hernández, 2004a).

<sup>10</sup> En este caso, Carcedo González (2001: 40) mantiene como entradas distintas *médula* y *médula espinal*, por considerar que es una combinación en la que “la designación de la lexía simple no es obvia”.

con el aprendizaje escolar (Samper Hernández, 2003a: 203). Igualmente aparece abundante léxico coloquial y familiar (*barriga, moflete, morro, panza, sobaco, teta, tripa...*), palabras vulgares o malsonantes (*cojón, conejo, chocho, ojete, pi-cha, verga...*), eufemismos (*partes, pompis, posaderas, trasero...*).

**Cuadro 7.4. Comparación del léxico más disponible en el centro de interés  
'Partes del cuerpo'**

No	Asturias	Cast. La Mancha	Sevilla	Valencia	Zamora
1	2	3	4	5	6
1	ojo	cabeza	cabeza	ojo	ojo
2	cabeza	brazo	brazo	cabeza	cabeza
3	brazo	pierna	pierna	nariz	nariz
4	pierna	mano	ojo	brazo	brazo
5	nariz	ojo	mano	oreja	pierna
6	dedo	pie	nariz	dedo	mano
7	mano	dedo	dedo	mano	dedo
8	pie	nariz	oreja	pierna	pie
9	oreja	oreja	pie	boca	oreja
10	boca	boca	boca	pie	boca
11	rodilla	rodilla	rodilla	rodilla	cuello
12	cuello	uña	uña	cuello	rodilla
13	corazón	cuello	pelo	codo	uña
14	uña	pelo	cuello	uña	codo
15	codo	codo	codo	hombro	pelo
16	pelo	hombro	hombro	pelo	corazón
17	diente	tobillo	diente	diente	estómago
18	estómago	muñeca	pecho	muñeca	muñeca
19	pulmón	diente	culo	tobillo	diente
20	hombro	espalda	espalda	ceja	hombro
21	pecho	pecho	tobillo	pecho	pulmón
22	hígado	tronco	muñeca	lengua	hígado
23	muñeca	corazón	ceja	espalda	tobillo
24	tobillo	lengua	lengua	estómago	espalda
25	lengua	ceja	pene	corazón	lengua
26	espalda	culo	pestaña	pestaña	antebrazo
27	tronco	estómago	labio	labio	tronco
28	pene	cadera	tronco	pulmón	labio
29	ceja	pulmón	corazón	muslo	ceja

1	2	3	4	5	6
30	riñón	labio	estómago	antebrazo	pecho
31	culo	pene	muslo	cadera	riñón
32	cerebro	muslo	barriga	hígado	intestino
33	pestaña	pestaña	cadera	tronco	pestaña
34	fémur	cara	pulmón	pene	páncreas
35	intestino	antebrazo	cara	cerebro	cadera
36	labio	ombligo	ombligo	cabello	culo
37	páncreas	hígado	antebrazo	intestino	muslo
38	cadera	hueso	hígado	cintura	cara
39	hueso	músculo	vagina	culo	tibia
40	cara	cerebro	frente	riñón	pene
41	tibia	riñón	cintura	mejilla	laringe
42	muslo	glúteo	cerebro	frente	frente
43	peroné	cintura	gemelo	abdomen	tórax
44	esófago	extremidad	hueso	cara	cerebro
45	oído	gemelo	glúteo	ombligo	peroné
46	músculo	abdomen	riñón	fémur	cúbito
47	antebrazo	intestino	músculo	tibia	cintura
48	barriga	frente	barbilla	barbilla	abdomen
49	vagina	pechos	testículo	peroné	vena
50	cintura	fémur	intestino	testículo	fémur

En cuanto al léxico altamente disponible (norma léxica) compartido por los cinco léxicos disponibles comparados, se resume de la siguiente manera<sup>11</sup>:

– Vocablos comunes en las 10 primeras posiciones: (10) *boca, brazo, cabeza, dedo, mano, nariz, ojo, oreja, pie, pierna*.

– Otros vocablos comunes en la franja de disponibilidad +0,1: (18) *ceja, codo, corazón, cuello, diente, espalda, estómago, hombro, labio, lengua, muñeca, pecho, pelo, pulmón, rodilla, tobillo, tronco, uña*.

– Otros vocablos comunes entre las 50 primeras posiciones: (11) *cadera, cara, cerebro, cintura, culo, hígado, intestino, muslo, pene, pestaña, riñón*.

– Vocablos comunes a 4 comunidades: (2) *fémur, frente*.

– Vocablos comunes a 3 comunidades: (7) *abdomen, antebrazo, hueso, músculo, ombligo, peroné, tibia*.

<sup>11</sup> Entre paréntesis aparece el cardinal del conjunto, esto es, el número de vocablos que lo integran.

### 7.2.2. 'La ropa'

En este centro de interés (ver Cuadro 7.5) aparecen no sólo unidades léxicas que designan prendas concretas de ropa (*calcetín, camisa, camiseta, chaqueta, falda, jersey, pantalón, vestido...*), sino también términos que se utilizan para designar calzado (*bota, botín, chancla, sandalia, zapato, zapatilla, zueco...*) y muchas otras palabras que reflejan asociaciones de muy diversa naturaleza en el marco amplio de la indumentaria.

Dentro de este campo pueden distinguirse campos más reducidos, como el de la ropa femenina (*biquini, blusa, camisón, falda, media, minifalda, rebeca, top, vestido...*), ropa interior (*bóxer, braga, calzoncillo, sostén, sujetador, tanga...*), prendas exteriores (*abrigo, americana, anorak, capa, cazadora, chaqueta, chaquetón, chubasquero, gabardina, impermeable...*), y otros campos léxicos más alejados del núcleo central, con elementos que pueden considerarse prendas y otros que no, entre los que se encuentran las diferentes partes de una prenda (*botón, cremallera, escote, dobladillo, tacón...*), materiales (*algodón, lana, licra, pana, piel, tejido, tela...*), adornos diversos (*lentejuela, encaje, puntilla...*), accesorios y complementos (*anillo, bolso, collar, colgante, diadema, guante, paraguas, pulsera, reloj...*), ropa de la casa (*manta, mantel, toalla...*), etc.

Abundan, además, las subespecificaciones del tipo *abrigo de visón, camiseta interior, cazadora vaquera, chaqueta de cuero, falda larga, pantalón corto, pantalón de deporte, ropa de baño*, etc. En este caso, estos grupos léxicos se mantienen separados de la correspondiente forma simple (*abrigo, camiseta, cazadora, chaqueta, falda, pantalón, ropa*). Por el contrario, se recurre al uso del paréntesis y una entrada única en los casos en que hay alternancia, como por ejemplo, *pantalón vaquero* y *vaquero* = (*pantalón*) *vaquero*.

Son abundantes los neologismos de diverso origen, sobre todo del francés (*corsé, chaqué, frac, maillot, mitón...*) y del inglés (*body, jeans, panty, pullover, slip, t-shirt...*).

Al igual que en el centro anterior, aparecen muchas prendas de vestir que se enuncian en plural por tener dos partes iguales. Se reducen a su forma singular *vaqueros* > *vaquero*, *pantalones* > *pantalón*; *bragas* > *bragas*. Los investigadores mantienen en plural, no obstante, las formas *bermudas, gafas, tirantes*.

**Cuadro 7.5. Comparación del léxico más disponible en el centro de interés 'La ropa'**

No	Asturias	Cast. La Mancha	Sevilla	Valencia	Zamora
1	2	3	4	5	6
1	pantalón	pantalón	pantalón	camisa	pantalón
2	camisa	camisa	camisa	pantalón	camisa
3	camiseta	camiseta	camiseta	calcetín	jersey

1	2	3	4	5	6
4	jersey	jersey	chaleco	zapato	calcetín
5	calcetín	calcetín	calcetín	camiseta	camiseta
6	falda	falda	falda	braga	falda
7	braga	braga	bragas	calzoncillo	braga
8	calzoncillo	calzoncillo	jersey	chaqueta	calzoncillo
9	chaqueta	abrigo	calzoncillo	falda	media
10	zapato	bufanda	chaqueta	zapatilla	abrigo
11	sujetador	zapato	abrigo	jersey	chaqueta
12	media	chaqueta	zapato	sujetador	zapato
13	abrigo	sujetador	bufanda	media	sujetador
14	cazadora	guante	sujetador	suéter	bufanda
15	bufanda	sudadera	chaquetón	bota	cazadora
16	guante	chándal	media	abrigo	blusa
17	sudadera	media	guante	bufanda	chaleco
18	playero	gorro	sudadera	guante	guante
19	vestido	zapatilla	chándal	cazadora	zapatilla
20	chándal	vestido	vestido	chaleco	vestido
21	gorro	(pantalón) vaquero	blusa	corbata	chándal
22	bota	cazadora	(pantalón) vaquero	sombrero	corbata
23	corbata	tanga	gorro	gorro	bota
24	chaleco	bota	botín	blusa	polo
25	blusa	chaleco	bota	cinturón	gabardina
26	bañador	blusa	tanga	chándal	americana
27	vaquero	suéter	calzona	vestido	minifalda
28	sombrero	corbata	rebeca	sudadera	gorro
29	polo	cinturón	pijama	pañuelo	body
30	chubasquero	(zapatilla) de- portiva	camiseta interior	(pantalón) vaquero	bermudas
31	bermudas	bañador	leotardo	camiseta interior	pañuelo
32	niki	pijama	sombrero	bermudas	bañador
33	cinturón	gorra	corbata	gorra	sombrero
34	chaquetón	sombrero	zapatilla	mallá	pijama
35	gorra	top	(zapato de) deporte	minifalda	chubasquero
36	traje	pluma	pañuelo	plumífero	chaquetón
37	pijama	pañuelo	cazadora	pijama	suéter

Cuadro 7.5 (cont.)

38	biquini	polo	camisón	bañador	niqui
39	americana	traje	panty	sandalia	faja
40	pañuelo	(forro) polar	gorra	(chaqueta) americana	manopla
41	anorak	camiseta interior	bañador	body	panty
42	sandalia	leotardo	malla	gabardina	gorra
43	zapatilla	pantalón corto	top	chaquetón	vaquero
44	top	malla	bermudas	anorak	cinturón
45	camisón	biquini	anorak	chancla	sandalia
46	gabardina	panti	traje	pajarita	pajarita
47	gayumbo	camisón	cinturón	leotardo	camisón
48	pajarita	minifalda	polo	(zapatilla) deportiva	anorak
49	slip	anorak	(forro) polar	bolso	plumífero
50	leotardo	faja	suéter	panty	traje

Aparecen como entradas diferentes la distinción basada en el morfema de género de las formas *bolso/bolsa* y *gorro/gorra* porque estas variantes morfológicas implican también diferencias semánticas. Por la misma razón –constituir realidades diferentes–, se mantienen los derivados de *bolso*: *bolsillo*; de *bota*: *botín*; de *blusa*: *blusón*; de *calza*: *calzón*, *calzoncillos*; de *calceta*: *calcetín*; de *camisa*: *camiseta*, *camisilla*, *camisola*, *camisón*; de *chancla*: *chancleta*; de *collar*: *collarín*; de *faldas*: *faldilla*, *faldón*; de *zapato*: *zapatilla*.

La norma léxica compartida por los cinco léxicos disponibles comparados resulta ser:

– Vocablos comunes en las 10 primeras posiciones: (10) *braga*, *calcetín*, *calzoncillo*, *camisa*, *camiseta*, *chaqueta*, *falda*, *jersey*, *pantalón*, *zapato*.

– Otros vocablos comunes en la franja de disponibilidad +0,1: (12) *abrigo*, *bota*, *blusa*, *bufanda*, *chaleco*, *chándal*, *gorro*, *guante*, *media*, *sudadera*, *sujetador*, *vestido*.

– Otros vocablos comunes entre las 50 primeras posiciones: (11) *anorak*, *bañador*, *cazadora*, *cinturón*, *corbata*, *gorra*, *pañuelo*, *pijama*, *sombrero*, *vaquero*, *zapatilla*.

– Vocablos comunes a 4 comunidades: (8) *bermudas*, *camisón*, *chaquetón*, *leotardo*, *panty*, *polo*, *suéter*, *traje*.

– Vocablos comunes a 3 comunidades: (8) *americana*, *camiseta interior*, *gabardina*, *malla*, *minifalda*, *pajarita*, *sandalia*, *top*.

### 7.2.3. 'Partes de la casa (sin muebles)'

Como indica Trigo Ibáñez (2011: 44), el centro de interés 'Partes de la casa' (Cuadro 7.6) es especialmente controvertido "puesto que la mayor parte de los informantes parece olvidar la consigna «sin los muebles» y, además de ello, ofrece respuestas que no encajan exactamente con el estímulo presentado". Dada la diversa naturaleza de las realidades evocadas por las lexías que integran las listas, Carcedo González (2001: 41) considera necesario establecer un criterio discriminatorio claro para decidir cuáles de ellas se incluyen y cuáles no. Como tal criterio, se atiende al carácter fijo/móvil de la realidad asociada<sup>12</sup>. Por ello desecha *cortina*, *cuadro*, *bombilla*, *espejo*. Sin embargo, otros autores (Prado Aragonés y Galloso Camacho, 2005; Hernández Muñoz, 2006; Trigo Ibáñez, 2011; etc.) mantienen todos los términos evocados al considerarlos un producto de la espontaneidad del informante y de determinadas asociaciones mentales que no pueden pasarse por alto.

En cualquier caso, todos los autores están de acuerdo en entender el estímulo 'partes de la casa' en un sentido amplio. Por ello, además de los nombres de las habitaciones (*cocina*, *comedor*, *cuarto de baño*, *despensa*, *dormitorio*, *habitación*, *sala de estar*, *salón...*), se incluyen elementos que, sin constituir propiamente una dependencia doméstica, se hallan estrechamente ligados con la vivienda. Así, en el vocabulario de este centro tienen cabida: componentes arquitectónicos (*arco*, *barandilla*, *capitel*, *columna*, *cimiento*, *pared*, *techo...*), elementos funcionales y de saneamiento (*ascensor*, *bañera*, *bidé*, *calentador*, *grifo*, *lavabo...*), elementos ornamentales (*columna*, *jardinera*, *moldura*, *cerámica...*), materiales de construcción (*azulejo*, *baldosa*, *cemento*, *ladrillo*, *mármol*, *pintura...*), útiles diversos (*bombilla*, *cable*, *cañería*, *enchufe*, *grifo...*), elementos que se encuentran fuera de la casa (*acera*, *caseta del perro*, *césped*, *jardín*, *piscina...*), tipos de vivienda (*adosado*, *ático*, *apartamento*, *chalé*, *chabola*, *piso...*) e incluso muebles (*armario empotrado*, *encimera*, *estantería*, etc.).

Tal y como señala Gómez Devís (2004: 84), "este es uno de los centros de interés en los que aparecen muchos sintagmas complejos, especialmente, con *cuarto* (*de costura*, *de invitados*, *de planchar...*) y *sala* (*de espera*, *de lectura*, *de juntas*, *de visitas*, *de proyección...*)". Según lo expuesto por Samper Padilla (1998a: 319), cuando se trata de variantes de un mismo núcleo lexemático, se mantienen con sus propias entradas *cuarto de estar* – *cuarto de estudio* – *cuarto de juegos* – *cuarto oscuro*. Pero se efectúa la reducción con paréntesis para las entradas sinónimas que aparecen con todos sus elementos o con reducción<sup>13</sup> de alguno de ellos: *aseo*

<sup>12</sup> Siguiendo un criterio similar, en Ceuta, Ayora Esteban (2006) descarta los enseres, adornos y electrodomésticos por no considerarlos 'partes de la casa'.

<sup>13</sup> Trigo Ibáñez (2011: 44) extiende al reducción a las formas sinonimicas *cuarto* y *sala*, por lo que en el léxico disponible de Sevilla aparecen formas como (*cuarto/sala*) *de estudio*, (*cuarto/sala de*) *estar*, etc.

y cuarto de aseo = (cuarto de) aseo; baño y cuarto de baño = (cuarto de) baño<sup>14</sup>; trastero y cuarto trastero = (cuarto) trastero.

En cuanto a los derivados, se conservan aquellos lexicalizados en el *DRAE* y que designan realidades distintas, como *barandilla* frente a *baranda*, *portalón* frente a *portal*, *salita*, y tanto en Sevilla como en Castilla-La Mancha *cuartillo*, que, aunque “no aparece en el *DRAE* con la acepción de ‘habitación de la casa’, está claramente lexicalizado en la región castellano-manchega por el número de menciones y por la aparición de *cuarto* y *cuartillo* en la misma encuesta” (Hernández Muñoz, 2006: 281).

**Cuadro 7.6. Comparación del léxico más disponible en el centro de interés  
‘Partes de la casa’**

No	Asturias	Cast. La Mancha	Sevilla	Valencia	Zamora
1	2	3	4	5	6
1	cocina	cocina	cocina	cocina	cocina
2	(cuarto de) baño	salón	salón	comedor	salón
3	salón	comedor	(cuarto de) baño	(cuarto de) baño	comedor
4	habitación	dormitorio	dormitorio	recibidor	(cuarto de) baño
5	comedor	habitación	salita	salón	dormitorio
6	despensa	baño	patio	salita	habitación
7	salita	terracea	comedor	terracea	pasillo
8	pasillo	pasillo	azotea	dormitorio	terracea
9	terracea	despensa	entrada	habitación	despensa
10	desván	patio	terracea	balcón	sala de estar
11	hall	garaje	pasillo	pasillo	(cuarto) trastero
12	garaje	cuarto de baño	jardín	(cuarto) trastero	hall
13	(cuarto) trastero	buhardilla	habitación	galería	garaje
14	dormitorio	trastero	(cuarto/sala de) estar	jardín	entrada
15	sótano	entrada	escalera	sala de estar	tejado
16	ventana	salita	sótano	garaje	salita (de estar)
17	pared	jardín	garaje	despacho	servicio
18	escalera	sótano	lavadero	entrada	buhardilla
19	tejado	sala de estar	cochera	buhardilla	sótano
20	entrada	escalera	(cuarto) trastero	despensa	desván
21	puerta	hall	balcón	(cuarto de) aseo	ventana
22	buhardilla	tejado	(cuarto de) aseo	escalera	escalera
23	jardín	balcón	despensa	sótano	patio

<sup>14</sup> Hernández Muñoz (2006), por el contrario, los mantiene como formas separadas.

1	2	3	4	5	6
24	recibidor	recibidor	recibidor	hall	balcón
25	sala de estar	cochera	hall	ventana	puerta
26	techo	ventana	ventana	desván	recibidor
27	suelo	aseo	porche	puerta	pared
28	(cuarto de) aseo	cuarto de estar	(cuarto/sala de) estudio	pared	chimenea
29	balcón	despacho	puerta	ático	despacho
30	ático	servicio	buhardilla	techo	sobrado
31	despacho	puerta	zaguán	suelo	cuarto de estar
32	chimenea	pared	pared	biblioteca	bodega
33	porche	desván	cuarto	bodega	techo
34	estudio	porche	techo	estudio	biblioteca
35	biblioteca	suelo	desván	servicio	jardín
36	servicio	techo	despacho	lavadero	(sala de) estudio
37	bodega	chimenea	suelo	chimenea	suelo
38	patio	corral	tejado	patio	corral
39	portal	ático	ático	fachada	cochera
40	piscina	estudio	cuartillo	tejado	porche
41	cochera	bodega	piscina	cuarto de estar	lavabo
42	lavabo	cámara	cuarto de invitados	dormitorio de matrimonio	corredor
43	teja	lavadero	(cuarto/sala de) juegos	pilar	vestíbulo
44	ladrillo	piscina	chimenea	vestíbulo	portal
45	persiana	biblioteca	biblioteca	porche	teja
46	descansillo	vestíbulo	cuarto de la plancha	cuarto de juegos	galería
47	azulejo	ladrillo	servicio	cuarto de invitados	cuarto
48	retrete	descansillo	entradita	lavabo	ladrillo
49	gimnasio	azotea	descansillo	corral	cimiento
50	sauna	cuarto de aseo	vestidor	sala de juegos	tabique

En cuanto a voces extranjeras, aparecen algunos extranjerismos muy usuales en el español actual los anglicismos *hall*, *jacuzzi*, *lobby*, *office* y el galicismo *toilette*.

La norma léxica común a los cinco léxicos disponibles comparados se resume en:

– Vocablos comunes en las 10 primeras posiciones: (7) *cocina*, *comedor*, *(cuarto de) baño*, *habitación*, *pasillo*, *salón*, *terraza*.

– Otros vocablos comunes en la franja de disponibilidad +0,1: (12) *balcón, buhardilla, (cuarto) trastero, despensa, dormitorio, entrada, escalera, garaje, recibidor, sala de estar, salita, sótano*.

– Otros vocablos comunes entre las 50 primeras posiciones: (16) *biblioteca, chimenea, despacho, desván, estudio, hall, jardín, pared, patio, porche, puerta, servicio, suelo, techo, tejado, ventana*.

– Vocablos comunes a 4 comunidades: (4) *ático, cochera, (cuarto de) aseo, (cuarto de) estar*.

– Vocablos comunes a 3 comunidades: (7) *bodega, corral, descansillo, ladrillo, lavabo, piscina, vestíbulo*.

#### 7.2.4. ‘Los muebles de la casa’

Este centro despierta una serie de asociaciones entre los informantes que excede a lo que entendemos por mobiliario en sentido estricto (ver Cuadro 7.7). Además de muebles típicos (*armario, butaca, cama, estantería, escritorio, mesa, silla, sillón, sofá...*), se mencionan con profusión electrodomésticos (*aspiradora, lavadora, lavavajillas, microondas, nevera, televisión...*), objetos de decoración (*adorno, cortina, cuadro, espejo, jarrón, lámpara, reloj...*), útiles de la casa diversos (*botiquín, impresora, máquina de coser, ordenador, plancha, teléfono...*), diferentes partes de los muebles (*cabecero, encimera, reposacabezas, reposabrazos, respaldar*), etc. Se incluyen también vocablos como *dormitorio, despacho, recibidor, entrada*, porque además de ser específicos del centro de interés anterior, como indica J. A. Samper Padilla (1998a: 320) designan metonímicamente el mobiliario de esas habitaciones.

A propósito de la aceptación o no de algunos vocablos incluidos en las respuestas de los informantes, aparecen pequeñas discrepancias entre los investigadores. De este modo, en la subcategoría del mobiliario relacionado con la cocina, mientras que autores como Trigo Ibáñez (2011) y Hernández Muñoz (2006) no prescinden de ninguna voz para poder hacer estudios asociativos, o Ayora Esteban (2006) acepta *sartén* como uno de los enseres del mobiliario, Gómez Devís (2004: 84, n. 87) explica que “se han suprimido algunos utensilios como *sartén*”. Este diferente criterio también se aplica a otros vocablos. Así, Ayora Esteban (2006: 145), al contrario que otros autores (cf. Carcedo González, 2001: 43), no incluye *puerta* y *ventana* al no interpretarlos como elementos estéticos. Sobre este particular, Carcedo González (2000a: 77) argumenta acertadamente que “si se actuase en sentido muy estricto, serían muchas más las asociaciones que deberían desecharse” y que, por lo tanto, “no podrían admitirse ni en ‘Muebles’ ni en ‘Partes de la casa’, lo que no parece muy razonable”.

En cuanto a los derivados, según el criterio de Samper Padilla (1998a: 320), se mantienen las tres formas *mesa, mesilla* y *mesita* como entradas separadas, ya

que *mesita* no equivale a mesa pequeña sino a *mesilla*. Además, el *DRAE* también recoge los dos diminutivos lexicalizados. Sus respectivas lexías compuestas también aparecen diferenciadas<sup>15</sup>: *mesilla de noche*, *mesita de noche*, *mesa de noche*, entre otras. Igualmente representan distinción de significado los diminutivos (*balacín*, *barandilla*, *botiquín*, *comodín*, *hornilla*, *lamparilla*, *lamparita*, *sillín*) y los aumentativos (*cajón*, *sillón*, *taquillón*).

Como señala Hernández Muñoz (2006: 281), “en este centro de interés hay numerosas variaciones de género en la actualización de determinadas piezas del mobiliario que probablemente responden a preferencias regionales o comarcales”. Esta autora adopta la decisión de mantenerlas todas como entradas diferentes. Otros autores mantienen separadas solo las que el *DRAE* reconoce como vocablos diferenciados o los que funcionan como variantes dialectales. Entre el grupo de lexemas que conllevan una diferencia semántica se encuentran los pares *hornillo/hornilla* y *secador/secadora*; entre las formas que indican preferencia dialectal se incluyen las parejas *fregador/fregadora* y *fregadero/fregadera*<sup>16</sup>.

Por lo demás, se usa el paréntesis para indicar la presencia/ausencia de un elemento en los pares (*armario*) *ropero*, *cadena* (*musical*), (*cama*) *litera*, (*mesa*) *camilla*, (*mueble*) *bar*, etc., y en los acortamientos *foto*(*grafía*), *frigo*(*rífico*), *tele*(*visión*), etc. Destaca también la presencia de siglas (*DVD*, *TV*, *WC*).

**Cuadro 7.7. Comparación del léxico más disponible en el centro de interés  
‘Los muebles de la casa’**

Nº	Asturias	Cast. La Mancha	Sevilla	Valencia	Zamora
1	2	3	4	5	6
1	mesa	mesa	silla	silla	silla
2	armario	silla	mesa	mesa	mesa
3	silla	armario	cama	armario	cama
4	cama	cama	sofá	cama	(armario) ropero
5	sofá	sillón	armario	sillón	sillón
6	mesita (de noche)	sofá	sillón	sofá	sofá
7	sillón	estantería	mesita(-illa) de noche	estantería	mesilla (de noche)

<sup>15</sup> Como excepción a este proceder se presenta el léxico disponible de Sevilla, donde Trigo Ibáñez (2011) unifica ambas formas mediante el uso del paréntesis.

<sup>16</sup> La variante normativa *fregadero* suele ser la forma a la que se reduce este par (Galoso Camacho, 2001; Carcedo González, 2001; Hernández Muñoz, 2002; González Martínez, 2002; entre otros). La forma *fregadera* se presenta –y es mantenida– como variación dialectal en los léxicos disponibles de Aragón (Arnal Purroy *et al.*, 2004), Soria (Bartol Hernández, 2004a) y Castilla-La Mancha (Hernández Muñoz, 2006).

Cuadro 7.7 (cont.)

1	2	3	4	5	6
8	estantería	mesita	estantería	mesita de noche	estantería
9	escritorio	escritorio	ropero	escritorio	cocina
10	lámpara	cómoda	mueble bar	cómoda	mesita (de noche)
11	cómoda	mesita de noche	escritorio	aparador	(mesa) camilla
12	lavadora	encimera	cómoda	lámpara	lámpara
13	nevera	taburete	butaca	espejo	espejo
14	televisión	lámpara	lámpara	taburete	tele(visión)
15	tocador	mueble bar	repisa	mesita	mueble bar
16	espejo	armario empotrado	encimera	mecedora	cómoda
17	mueble bar	tresillo	frigorífico	tresillo	lavadora
18	bañera	televisión	tele(visor)	(mueble) bar	tresillo
19	butaca	espejo	espejo	nevera	escritorio
20	microondas	butaca	lavadora	lavadora	encimera
21	mesilla (de noche)	mecedora	mecedora	(armario) ropero	taquillón
22	zapatero	cajón	(mesa) camilla	librería	fregadero
23	lavabo	tocador	vitrina	cuadro	frigorífico
24	banqueta	puerta	mueble de cocina	butaca	nevera
25	horno	baúl	peinadora	cajón	tocador
26	cocina	mueble de cocina	cuadro	tele(visión)	mesa de estudio
27	mecedora	cocina	tresillo	zapatero	librería
28	cuadro	mesa camilla	lavabo	litera	taburete
29	lavavajillas	perchero	taburete	mesa de comedor	cajón
30	ropero	aparador	microondas	sinfonier	lavavajillas
31	puerta	banqueta	váter	mesa camilla	horno
32	aparador	lavabo	aparador	horno	butaca
33	taburete	sinfonier	tocador	(horno) microondas	armario empotrado
34	cajón	vitrina	bañera	puerta	mesa de tele(visión)
35	tresillo	coqueta	despensa	tocador	cuadro
36	armario empotrado	zapatero	lavavajillas	vitrina	bañera
37	bidé	taquillón	mesa del tele(visor)	perchero	lavabo

1	2	3	4	5	6
38	alfombra	lavadora	puerta	rinconera	coqueta
39	taquillón	frigo(rífico)	baúl	baúl	litera
40	perchero	librería	zapatero	cocina	comodín
41	ordenador	ropero	armario empotrado	armario empotrado	televisor
42	sinfonier	cuadro	mesa del comedor	estante	vídeo
43	comodín	comodín	cajón	banco	aparador
44	baúl	fregadero	librería	alfombra	sinfonier
45	alacena	cajonera	Romi	banqueta	mesa de salón
46	video	microondas	(cama) litera	bañera	recibidor
47	banco	horno	mesa de estudio	sofá cama	puerta
48	(mesa) camilla	sofá cama	perchero	mesa de cocina	mueble de tele (visión)
49	secadora	litera	secadora	lavavajillas	microondas
50	vitrina	bañera	horno	mesa de estudio	radio

La norma léxica común a los cinco léxicos disponibles comparados se compone de:

– Vocablos comunes en las 10 primeras posiciones: (8) *armario, cama, estantería, mesa, mesita/-illa (de noche), silla, sillón, sofá*.

– Otros vocablos comunes en la franja de disponibilidad +0,1: (3) *cómoda, escritorio, lámpara*.

– Otros vocablos comunes entre las 50 primeras posiciones: (16) *aparador, armario empotrado, bañera, cajón, cuadro, espejo, horno, lavadora, (mesa) camilla, microondas, mueble bar, puerta, ropero, taburete, tocador, tresillo*.

– Vocablos comunes a 4 comunidades: (14) *baúl, butaca, (cama) litera, cocina, lavabo, lavavajillas, librería, mecedora, nevera, perchero, sinfonier, tele(visión), vitrina, zapatero*.

– Vocablos comunes a 3 comunidades: (6) *banqueta, comodín, encimera, frigo(rífico), mesa de estudio, taquillón*.

### 7.2.5. ‘Comidas y bebidas’

El centro de interés ‘Comidas y bebidas’ (Cuadro 7.8) es un área bastante delimitada, muy concreta, aunque de gran variabilidad lingüística. En su contenido incluye nombres genéricos de alimentos en su estado básico (*aceite, agua, carne, huevo, leche, pescado...*), alimentos elaborados y platos cocinados (*albóndiga, arroz con leche, carne en salsa, empanada, gazpacho, pescado adobado...*), nombres

genéricos de las comidas (*almuerzo, cena, comida, desayuno, entrantes, merienda, postre...*) o los diferentes platos nacionales o regionales (*chili con carne, cocido madrileño, fabada, mojo agridulce, paella, tortilla española...*). Además, en todos los léxicos ocupa un lugar destacado el campo léxico de las bebidas, de muy diversos tipos (*agua, cerveza, coca cola, leche, ron, vino, vodka, zumo...*). Se incluyen, además, algunos términos que no designan directamente un tipo de comida o bebida, como *tapa, ración o copa, chupito*.

En este centro queda reflejada como entradas diversas la variación diatópica del español (*papa/patata, alubia/habichuela/judía, etc.*) para aludir a la misma realidad alimenticia. Además, en las regiones bilingües aparecen préstamos léxicos de la lengua o variedad autóctona al español: *fabes* del dialecto asturiano, *garrofó* de la lengua catalana.

Este centro de interés es uno de los que más marcas comerciales recogen. Algunos léxicos disponibles, como el de Salamanca (Galoso Camacho, 2002b), Valencia (Gómez Devís, 2004), Huelva (Prado Aragonés y Galoso Camacho, 2005), Castilla-La Mancha (Hernández Muñoz, 2006), etc., las conservan todas para conocer los diferentes grados de lexicalización. Otras investigaciones, como las de Asturias (Carcedo González, 2001), Ceuta (Ayora Esteban, 2006), Alicante (Martínez Olmos, 2007), Sevilla (Trigo Ibáñez, 2011), etc., incluyen solo aquellas marcas que consideran propiamente lexicalizadas. A tal fin, se sigue el criterio establecido por Samper Padilla (1998a: 317), que considera lexicalizadas todas aquellas marcas que en una situación comunicativa dada pueden sustituir al nombre común del producto, aunque la correspondencia no sea del todo exacta. Igualmente se admiten todas aquellas marcas que presenten un índice de aparición significativamente superior al genérico, pues ello indica que, aunque no estén lexicalizadas, están en vías de lexicalizarse (Ayora Esteban, 2006: 151). Esta decisión lleva a considerar claramente lexicalizadas voces como: *Aquarius, Coca-Cola, Cola Cao, Danone, Donut, Fanta, Ketchup, Martini, Nocilla, etc.*

Este campo nocional, es, a su vez, uno de los que mayor número de diminutivos lexicalizados agrupa. De esta manera, en las investigaciones se consignan como entradas los diminutivos *cebolleta, cochinito, empanadilla, ensaladilla, granadina, manzanilla, natillas, paletilla, palomitas y palometa, pastelito(-illo), pepinillo, pescadilla, quesito, tortita y tortilla*, separadas de su respectiva forma primitiva *cebolla, cochino, empanada, ensalada, granada, manzana, nata, paleta, paloma, pastel, pepino, pescado, queso, torta*. En cuanto al género, en los léxicos disponibles constan en entradas separadas por aludir a realidades diferentes los pares *hueva/huevo, magra/magro, morcillo/morcilla y pimienta/pimienta*.

Otra característica de este centro es la presencia de muchas lexías complejas que sirven como subespecificaciones de los núcleos: *agua mineral, arroz a la cubana, arroz blanco, carne asada, queso manchego, tortilla de patata, zumo de limón, etc.* Como de costumbre, se usa el paréntesis agrupando bajo

la misma entrada aquellas palabras que presentan en ocasiones solo uno de los componentes, como en (*carne de*) *cerdo*, *ensaladilla (rusa)*, (*producto*) *lácteo*, (*vino*) *tinto*, etc.

**Cuadro 7.8. Comparación del léxico más disponible en el centro de interés ‘Comidas y bebidas’**

Nº	Asturias	Cast. La Mancha	Sevilla	Valencia	Zamora
1	2	3	4	5	6
1	agua	agua	coca cola	agua	agua
2	coca cola	coca cola	agua	vino	vino
3	vino	whisky	güisqui	tomate	coca cola
4	lenteja	vino	fanta	lechuga	güisqui
5	patata	naranja	patata	güisqui	pan
6	garbanzo	pan	carne	coca cola	naranja
7	güisqui	zumo	lenteja	leche	lenteja
8	espagueti	lenteja	pan	arroz	alubia
9	arroz	macarrón	cerveza	naranja	lechuga
10	leche	patata	zumo	manzana	cerveza
11	macarrón	leche	pescado	macarrón	tomate
12	carne	tomate	tomate	pera	garbanzo
13	fabada	carne	garbanzo	refresco	carne
14	vodka	pescado	ron	zumo	manzana
15	pescado	cerveza	lechuga	pan	leche
16	lechuga	arroz	leche	carne	patata
17	pan	fanta	vino	cerveza	zumo
18	zumo	vodka	arroz	patata	pescado
19	tomate	lechuga	macarrón	queso	pera
20	fanta	limón	jamón	espagueti	vodka
21	filete	pollo	espagueti	plátano	macarrón
22	sopa	manzana	queso	vodka	arroz
23	chorizo	judía	vodka	(carne de) pollo	chorizo
24	huevo	pera	huevo	pescado	ginebra
25	cerveza	espagueti	pizza	huevo	filete
26	(carne de) pollo	plátano	hamburguesa	lenteja	gaseosa
27	fabes	chorizo	naranja	chorizo	ron
28	manzana	garbanzo	refresco	(carne de) ternera	queso
29	jamón	ron	verdura	limón	Fanta

Cuadro 7.8 (cont.)

1	2	3	4	5	6
30	tortilla	jamón	filete	yogur	pollo
31	ron	queso	chorizo	jamón	limón
32	hamburguesa	fanta naranja	(carne de) pollo	(bebida) gaseosa	zanahoria
33	ensalada	ginebra	plátano	zanahoria	plátano
34	pizza	fanta limón	zanahoria	cebolla	melocotón
35	sidra	huevo	ginebra	fideo	coñac
36	martini	verdura	casera	limonada	sandía
37	queso	martini	pera	melocotón	melón
38	naranja	chocolate	chícharo	ginebra	salchichón
39	licor	zanahoria	café	naranjada	huevo
40	refresco	hamburguesa	puchero	chocolate	quivi
41	ginebra	pizza	manzana	ron	sopa
42	calimocho	cocido	salchicha	judía	jamón
43	pera	café	pasta	fresa	espagueti
44	plátano	yogur	pimiento	licor	café
45	café	fruta	fruta	garbanzo	cebolla
46	kas	ponche	yogur	melón	champán
47	chocolate	refresco	chocolate	fanta	pimiento
48	merluza	pasta	(vino) tinto	alcachofa	salchicha
49	paella	legumbre	salchichón	(carne de) cordero	verdura
50	yogur	sopa	fideo	quivi	hamburguesa

La norma léxica común a los cinco léxicos disponibles comparados incluye:

– Vocablos comunes en las 10 primeras posiciones: (3) *agua, coca cola, güisqui*.

– Otros vocablos comunes en la franja de disponibilidad +0,1: (19) *arroz, carne, cerveza, (carne de) pollo, chorizo, leche, lechuga, lenteja, macarrón, naranja, pan, patata, pescado, queso, ron, tomate, vino, vodka, zumo*.

– Otros vocablos comunes entre las 50 primeras posiciones: (9) *espagueti, fanta, garbanzo, ginebra, huevo, jamón, manzana, pera, plátano*.

– Vocablos comunes a 4 comunidades: (6) *café, chocolate, hamburguesa, refresco, yogur, zanahoria*.

– Vocablos comunes a 3 comunidades: (5) *filete, limón, pizza, sopa, verdura*.

### 7.2.6. 'Objetos sobre la mesa para la comida'

El centro de interés 'Objetos colocados encima de la mesa para la comida' (Cuadro 7.9) es una categoría no convencional del tipo *goal-derived category* o categoría *ad hoc*<sup>17</sup> (Barsalou, 1983), cuyos miembros se definen por ser elementos para la consecución de un objetivo concreto. Hernández Muñoz (2006: 146–147) explica la especificidad de este campo del siguiente modo:

la fuerza aglutinadora de este campo nocional es el objetivo de aportar todos aquellos elementos que se necesitan para una comida, no sólo los estrictamente necesarios, sino todos aquellos que hagan una comida más completa y placentera. Sólo de esta manera se entiende la presencia de términos tan dispares como *pan, platos, mantel* o *florero*. En este centro de interés no existe un prototipo identificable, es difícil encontrar los rasgos que encadenan a todos los miembros del grupo. ¿Cómo relacionamos subgrupos tan dispares como *vajilla, cubiertos, comida* o *elementos decorativos*? Lo que hace que un elemento sea más típico en esta categoría es el cumplimiento de la finalidad: aquello más imprescindible para poder comer en la mesa.

Ante este hecho, que “las respuestas ofrecidas por los informantes no se limitan al estímulo dado, sino que ofrecen también variedad de alimentos, personas, utensilios de cocina, etc.” (Trigo Ibáñez, 2011: 46), la actitud de los investigadores varía de unos a otros. Así, hay soluciones más restrictivas, como la de Gómez Devís (2004: 86), que excluye “aquellos que se refieren claramente a alimentos y bebidas y que ya que han aparecido previamente en el centro 05 (*postre, café, refresco, ajo, limón*)”; soluciones intermedias, como la de Carcedo González (2001: 45), que mantiene las denominaciones de alimentos “que acompañan habitualmente a las distintas comidas (*pan, café, etc.*)”; o soluciones totalmente permisivas, como la de Trigo Ibáñez (2011: 46), que decide “incluir las realidades evocadas para respetar la espontaneidad de cada encuestado y establecer comparaciones con otras investigaciones que se rigen por criterios de edición similares”.

Por lo demás, sí hay acuerdo en aceptar en la composición de este centro las denominaciones genéricas de platos (*bebida, comida, entremés, fruta, postre...*), los nombres genéricos de algunas bebidas frecuentes en las comidas (*agua, cerveza, vino, zumo...*) y diversos condimentos (*aceite, azúcar, especia, ketchup, mostaza, pimienta, sal, salsa, vinagre...*). Sobre este último aspecto, Samper Padilla (1998a: 323) justifica el mantenimiento de algunos de estos términos (*agua, sal, vinagre, pimienta* y *aceite*) “como ejemplo de una metonimia frecuente en el habla”, empleando el nombre del contenido por el continente. En consecuencia, en los listados definitivos aparecen tanto estos condimentos como sus recipientes:

<sup>17</sup> Frente a la opinión de Hernández Muñoz (2006: 146), que considera este centro como una categoría *ad hoc*, Tomé Cornejo (2011: 27–28) entiende que “los informantes comparten el hábito de poner la mesa y que por ello se trata de una categoría altamente familiar y fijada en la memoria a largo plazo”, por lo que se trata, en su opinión, de una *categoría derivada por la finalidad*.

*aceitera, azucarero, bote, botella, especiero, jarra, pimentero, salero, salsero, vinagrera, etc.*

Respecto a las variantes flexivas de género, en este centro aparece un elevado número de parejas donde se manifiesta la vacilación masculino/femenino con alternancia de o/a en el mismo vocablo. Las soluciones de los investigadores ante este fenómeno también ofrecen variedad. Así, entre otras soluciones, Hernández Muñoz (2006: 284) mantiene todas las formas según han sido evocadas por los informantes; Arnal Purroy *et al.* (2004: 31) las neutraliza en el término normativo; Gómez Devís (2004: 86) según la mayor frecuencia; Bartol Hernández (2004b: 1082) las unifica marcando la variación de género o/a. En algunos casos la diferencia de significado entre las dos formas es evidente: *bote/bota, cuchillo/cuchilla*. En otros, la diferenciación semántica no es tan clara, aunque el *DRAE* presenta dos entradas distintas, que también aparecen en los léxicos disponibles: *azucarero/azucarera, canastillo/canastilla, canasto/canasta, cesto/cesta, jarro/jarra, panero/panera, peroll/perola, salero/salera, etc.*

**Cuadro 7.9. Comparación del léxico más disponible en el centro de interés ‘Objetos sobre la mesa para la comida’**

Nº	Asturias	Cast. La Mancha	Sevilla	Valencia	Zamora
1	2	3	4	5	6
1	tenedor	tenedor	tenedor	tenedor	tenedor
2	cuchara	cuchara	cuchillo	cuchara	cuchara
3	cuchillo	cuchillo	cuchara	cuchillo	cuchillo
4	vaso	vaso	vaso	vaso	servilleta
5	servilleta	servilleta	servilleta	servilleta	vaso
6	plato	plato	plato	mantel	plato
7	mantel	mantel	mantel	plato	mantel
8	jarra	pan	jarra	jarra	jarra
9	plato llano	jarra	ensaladera	botella	cucharilla
10	cucharilla	botella	copa	copa	cazuela
11	plato hondo	plato hondo	cucharilla(-ita)	cucharilla	plato hondo
12	fuelle	cucharilla	pan	plato llano	servilletero
13	botella	agua	plato hondo	plato hondo	copa
14	copa	plato llano	plato llano	salero	botella
15	ensaladera	panera	botella	servilletero	plato llano
16	plato de postre	copa	fuelle	bandeja	pan
17	taza	vino	servilletero	fuelle	fuelle
18	bandeja	fuelle	panera	ensaladera	ensaladera
19	servilletero	servilletero	salero	tenedor de pescado	sopera

1	2	3	4	5	6
20	salero	ensaladera	bandeja	tenedor de carne	taza
21	pota	frutero	cubierto	frutero	salero
22	sopera	salero	agua	cuchillo de pescado	bandeja
23	tenedor de pescado	comida	sopera	sopera	cazo
24	pan	salvamanteles	frutero	vela	plato de postre
25	cuchillo de pescado	jarra de agua	taza	cuchillo de carne	panera
26	tenedor de carne	bandeja	cazo	copa de vino	frutero
27	copa de vino	cacerola	salvamantel	panera	cucharón
28	azucarero	taza	jarra de agua	taza	tenedor de pescado
29	garcilla	vinagrera	cuenco	cuchara de postre	cuchara de postre
30	frutero	olla	salsera	copa de champán	copa de vino
31	cuchillo de carne	sal	comida	aceitera	agua
32	espumadera	botella de vino	cesta de pan	posavasos	jarra de agua
33	copa de agua	botella de agua	olla	plato de postre	botella de vino
34	salvamanteles	sartén	posavasos	cucharón	tenedor de carne
35	palillo	fruta	vela	copa de agua	vaso de agua
36	panera	coca cola	plato de postre	vinagrera	copa de agua
37	cesta de pan	aceite	bol	azucarera	cuchara sobpera
38	vinagrera	gaseosa	hule	candelabro	plato plano
39	bol	vela	bebida	salvamanteles	cuchillo de pescado
40	salsera	palillo	pico	vaso de vino	vino
41	posavasos	vinagre	vinagrera	agua	sartén
42	jarra de agua	hule	tenedor (de/para) pescado	palillo	azucarero
43	olla	plato de postre	vino	vaso de agua	cesta de pan
44	cazuela	posavasos	aceitera	pan	vinagrera
45	cuchara de postre	bebida	cucharón	vinagreras	cuchillo de carne
46	tenedor de postre	cubierto	cerveza	cuchara de café	aceitera
47	cucharón	ensalada	cuchillo (de/para) pescado	tenedor de postre	plato sobpero
48	botella de vino	cucharón	aceite	cubierto	vaso de vino
49	comida	sopera	cenicero	cazuela	posavasos
50	agua	cazo	cuchillo (de/para) carne	centro de flores	copa de champán

En los léxicos disponibles se incluyen los diminutivos que aportan un significado diferente al de su base léxica. Así hay entradas separadas para *botellín*, *cañita*, *cubito* (*de hielo*), *pajita*, *paleta*, *palillo*, *toallita húmeda* frente a *botella*, *caña*, *cubo*, *paja*, *pala*, *palo*, *toalla*. Por esta misma razón, según el criterio de Samper Padilla (1998a: 323), no se unifican las variantes *plato*, *platito*, *platillo* ni *cuchara*, *cucharilla*, *cucharita*, *cucharón*.

En este centro son frecuentes los compuestos sintagmáticos en torno a los núcleos: *bandeja* (*de carne*, *de ensalada*, *de pescado...*), *copa* (*de agua*, *de vino*, *de champán...*), *cuchara* (*grande*, *pequeña*, *de helado*, *de sopa...*), *cuchillo* (*de carne*, *de pescado...*), *plato* (*hondo*, *llano*, *de postre...*), *tenedor* (*de carne*, *de pescado*, *de postre...*) y *vaso* (*alto*, *grande*, *pequeño*, *de vino...*).

Se incluyen, como formas ya lexicalizadas, los nombres comerciales *tetra brik* ('recipiente de cartón para envasar líquidos') y *tupperware* ('envase de plástico para conservar los alimentos').

La norma léxica común a los cinco léxicos disponibles comparados incluye:

– Vocablos comunes en las 10 primeras posiciones: (8) *cuchara*, *cuchillo*, *jarra*, *mantel*, *plato*, *servilleta*, *tenedor*, *vaso*.

– Otros vocablos comunes en la franja de disponibilidad +0,1: (8) *cucharilla*, *botella*, *copa*, *ensaladera*, *fuelle*, *plato hondo*, *plato llano*, *servilletero*.

– Otros vocablos comunes entre las 50 primeras posiciones: (12) *agua*, *bandeja*, *cucharón*, *frutero*, *pan*, *panera*, *plato de postre*, *posavasos*, *salero*, *sopera*, *taza*, *vinagrera*.

– Vocablos comunes a 4 comunidades: (5) *cuchillo de carne*, *cuchillo de pescado*, *jarra de agua*, *salvamanteles*, *tenedor de pescado*.

– Vocablos comunes a 3 comunidades: (13) *aceitera*, *azucarero*, *botella de vino*, *cazo*, *cazuela*, *cesta de pan*, *copa de agua*, *copa de vino*, *cuchara de postre*, *olla*, *palillo*, *tenedor de carne*, *vino*.

### 7.2.7. 'La cocina y sus utensilios'

'La cocina' (Cuadro 7.10) es un centro de interés heterogéneo, que presenta puntos de intersección con los centros de interés 03 'Partes de la casa', 04 'Los muebles de la casa', 05 'Comidas y bebidas' y 06 'Objetos sobre la mesa para la comida'. Como resultado de ello, evoca un porcentaje considerable de vocabulario que aparece en estos centros. Ello ha propiciado que investigadores como Carcedo González (2001: 45) y Gómez Devís (2004: 86) hayan recurrido a un cierto grado de restricción en los estímulos asociativos para excluir aquellos términos que, en su opinión, están escasamente relacionados con el mismo (*caja del gato*, *cesta del perro*, *conserva*, *cocinera*, *mostrador...*) "y otros que, además de escasa relación, ya han sido citados en otros centros de interés y presentan allí índices de disponibilidad más altos" (Gómez Devís, 2004: 86)<sup>18</sup>.

<sup>18</sup> Entre los vocablos descartados, Gómez Devís (2004: 86) menciona *ventana*, *puerta*, *azulejo*, *baldosa*, *poyo*, *perchero* y *botijo*. No obstante, muchos de estos términos son absolutamente

Así pues, el vocabulario de este centro incluye las denominaciones de los cubiertos (*cubierto, cuchara, cucharilla, cucharón, cuchillo, tenedor...*), las partes de la vajilla (*bol, bandeja, cuenco, ensaladera, fuente, plato, plato hondo, plato llano, taza, tazón...*), recipientes para líquidos (*botella, copa, jarra, vaso...*), los recipientes para cocinar y calentar (*cacerola, cafetera, cazo, olla, olla exprés, paellera, sartén, tapa, tapadera, tetera...*), diversos utensilios de cocina (*abrelatas, colador, espumadera, sacacorchos, tabla, tijeras...*), numerosos electrodomésticos (*batidora, friegaplatos, frigo(rífico), freidora, horno, lavadora, lavavajillas, microondas, nevera, secadora, tostador...*), y otros aparatos electrónicos que hoy día es habitual tener en la cocina (*radio, radiocasete, teléfono, televisión y televisor*), muebles (*estantería, fregadero, mesa, silla...*), partes de la cocina (*despensa, lavadero, pared, puerta, suelo...*), algunos textiles (*delantal, mantel, servilleta...*), algunos alimentos básicos (*carne, cebolla, fruta, harina, hortaliza, huevo, leche, legumbre, mantequilla, pan, pasta, patata, tomate...*), especias y condimentos (*aceite, ajo, azafrán, azúcar, curry, especia, laurel, orégano, perejil, sal, vinagre...*), etc.

Hernández Muñoz (2006: 285) constata que “este centro de interés junto con el anterior, *Objetos colocados encima de la mesa*, son los núcleos temáticos que producen más pares de sustantivos donde podemos observar oscilaciones en el género gramatical”. En algunos casos, el diccionario académico agrupa en un mismo lema la forma masculina y la femenina (*asador/a, aspirador/a, batidor/a*). En otros casos, claramente se trata de dos palabras que designan diferentes referentes: *bote/bota de vino, caldero/caldera, campano/campana, palo de madera/pala, rodillo/rodilla*. Hay también otros casos en los que el referente de ambos sustantivos parece ser el mismo, pero la norma léxica distingue dos palabras. Así, están recogidas en el *DRAE* (*azucarero/azucarera, cesto del pan/cesta del pan, bolso del pan/bolsa del pan, huevero/huevera, jarro/jarra, mazo/maza, panero/panera, tostador/tostadora, tostadero/tostadera*). Por lo demás, en las respuestas de los informantes hay bastantes parejas en las que sólo uno de sus componentes (o ninguno de ellos) aparece en el diccionario académico. En estos casos, autores como Bartol Hernández (2004a) o Hernández Muñoz (2006) conservan todas las formas para efectuar estudios sobre variantes regionales, usos idiosincrásicos de un individuo o de un reducido grupo de individuos o improvisaciones en el momento de realizar la encuesta. No obstante, la consigna general dentro del Proyecto Panhispánico de Disponibilidad Léxica es la unificar la flexión de género en torno a la forma normativa, como así hace Arnal Purroy *et al.* (2004: 33), o a la forma más usual, opción preferida por Ayora Esteban (2006: 155). En Castilla-La Mancha los términos a los que afecta el doble género (Hernández Muñoz, 2006: 286) son:

- femenino normativo: *aceitera, espumadera, freidora, licuadora, paellera, sopera, trituradora, vinagrera.*
- masculino normativo: *cuchillero, escurridor, escurridero, exprimidor, fregadero, fregador, picador, rallador, raspador, servilletero.*

---

necesarios para la descripción de una cocina, por lo que su exclusión de este centro de interés no nos parece fácilmente defendible.

Al igual que en los centros anteriores, se mantienen como entradas diferentes los diminutivos y aumentativos lexicalizados que en la oposición base léxica/forma derivada aportan significados diferentes, como en *cocina/cocinilla*, *cuchara/cucharilla/cucharón*, *horno/hornillo*, *pincho/pinchito*, *tenaza/tenacillas*, *tenedor/tenedorcillo*.

Como en el centro de interés anterior, abundan los compuestos sintagmáticos alrededor de los núcleos: *cuchara (de madera, de palo, sopera...)*, *cuchillo (eléctrico, jamonero, de cocina...)*, *paleta (de plástico, de madera, de hojalata...)*, *plato (sopero, pequeño, grande...)* y *tenedor (de madera, de pescado...)*, con predominio, en cualquier caso, de la composición preposicional (*aceite de oliva, cación de cubiertos, sartén con mango, tabla de cortar*, etc.)

El uso de los paréntesis se extiende a los acortamientos o a las formas que presentan algún componente suprimido: *bombona (de gas)*, *campana (extractora)*, *extractor (de humos)*, *frigo(rífico)*, *(horno) microondas*, *tele(visión)*, *robot (de cocina)*, *(placa) vitro(cerámica)*, etc.

Aparecen como marcas comerciales lexicalizadas *butano*, *fairy*, *minipimer*, *mistol*, *papel albal*, *tupperware*, *túrmix*, entre otras.

**Cuadro 7.10. Comparación del léxico más disponible en el centro de interés  
'La cocina y sus utensilios'**

Nº	Asturias	Cast. La Mancha	Sevilla	Valencia	Zamora
1	2	3	4	5	6
1	nevera	cuchillo	(horno) microondas	nevera	cazuela
2	microondas	tenedor	horno	(horno) microondas	sartén
3	mesa	cuchara	frigo(rífico)	horno	cuchillo
4	horno	sartén	sartén	cuchillo	horno
5	silla	microondas	olla	tenedor	microondas
6	armario	horno	cuchillo	cuchara	tenedor
7	sartén	plato	fregadero	lavadora	cuchara
8	cuchillo	frigo(rífico)	cuchara	lavavajillas	lavadora
9	lavavajillas	vaso	lavavajillas	sartén	batidora
10	fregadero	lavavajillas	tenedor	plato	cazo
11	tenedor	olla	(placa) vitro(cerámica)	vaso	olla
12	cuchara	cacerola	cazo	batidora	plato
13	lavadora	fregadero	espumadera	olla	nevera
14	vitrocerámica	cazo	plato	pila	lavavajillas
15	pota	vitro(cerámica)	lavadora	congelador	cocina
16	batidora	lavadora	cacerola	cazo	frigorífico

1	2	3	4	5	6
17	plato	nevera	batidora	cazuela	espumadera
18	espumadera	batidora	freidora	grifo	fregadero
19	vaso	freidora	(placa de) hornilla	frigorífico	mesa
20	cocina	mesa	vaso	mesa	vaso
21	olla	paleta	mesa	encimera	armario
22	cazo	espumadera	tostador	exprimidor	freidora
23	congelador	encimera	encimera	armario	lavaplatos
24	cafetera	servilleta	congelador	freidora	cafetera
25	frigorífico	cazuela	nevera	cocina	aceite
26	exprimidor	cafetera	despensa	tostadora	cucharón
27	grifo	cucharilla	silla	secadora	fuelle
28	cocina de gas	cucharón	exprimidor (de zumo)	sandwichera	campana
29	freidora	grifo	campana	silla	silla
30	cacerola	silla	sandwichera	cafetera	bandeja
31	tostador	taza	cazuela	lavaplatos	cocina de gas
32	televisión	fuelle	cucharón	extractor	exprimidor
33	fuelle	congelador	mueble	cucharón	escurr platos
34	campana (extractora)	sandwichera	cafetera	servilleta	congelador
35	rodillo	jarra	secadora	licuadora	sal
36	garcilla	tijera	grifo	fogón	picadora
37	licuadora	armario	escurridor	fregadero	licuadora
38	encimera	exprimidor	colador	salero	tabla
39	calentador	cocina	servilleta	banco	puchero
40	despensa	tostadora	estropajo	bandeja	colador
41	fogón	extractor	bandeja	cubierto	escurridor
42	cajón	fogón	cubierto	abrelatas	extractor
43	lavaplatos	estropajo	olla exprés	paleta	fogón
44	cucharón	escurridor	tijera	copa	cubierto
45	estantería	mantel	extractor (de humo)	jarra	vinagre
46	cazuela	fuego	perol	trapo	fuegos
47	cocina de leña	colador	fuelle	cacerola	cacerola
48	banqueta	bandeja	lavaplatos	fuelle	ensaladera
49	cubierto	tostador	salero	frutero	jarra
50	taza	copa	armario	cajón	bol

La norma léxica común a los cinco léxicos disponibles comparados se compone de:

– Vocablos comunes en las 10 primeras posiciones: (4) *cuchillo, horno, (horno) microondas, sartén*.

– Otros vocablos comunes en la franja de disponibilidad +0,1: (13) *batidora, cazo, cuchara, freidora, frigo(rífico), lavadora, lavavajillas, mesa, nevera, olla, plato, tenedor, vaso*.

– Otros vocablos comunes entre las 50 primeras posiciones: (9) *armario, cacerola, cafetera, cazuela, congelador, cucharón, exprimidor, fregadero, silla*.

– Vocablos comunes a 4 comunidades: (11) *bandeja, cocina, cubierto, encimera, espumadera, extractor, fogón, fuente, grifo, lavaplatos, tostador*.

– Vocablos comunes a 3 comunidades: (6) *colador, escurridor, jarra, licuadora, sandwichera, vitro(cerámica)*.

### 7.2.8. ‘La escuela (muebles y materiales)’

La principal divergencia entre los investigadores a la hora de editar las respuestas de los informantes en este centro estriba en si aceptar o no las designaciones de personas, ya que la etiqueta del centro hace alusión específicamente a ‘muebles y materiales’. La pauta establecida por Samper Padilla (1998a: 324), es la de eliminar estas designaciones de personas (*profesor, alumno...*). Autores como Samper Padilla y Hernández Cabrera (1997), Carcedo González (2000a) o Ayora Esteban (2006) actúan en este sentido. Sin embargo, la mayoría de autores sí han aceptado las designaciones personales del contexto escolar y las relaciones asociativas de contigüidad que en él se dan.

Por consiguiente, al centro de interés ‘La escuela’ (Cuadro 7.11) pertenecen, además de los nombres del material escolar (*bolígrafo, carpeta, cuaderno, diccionario, estuche, folio, fotocopia, goma, hoja, lápiz, libro, mochila, papel, pizarra, pluma, ordenador, regla, sacapuntas, tijeras, tiza...*) y los muebles (*banca, escritorio, mesa, pupitre, silla...*), vocablos que designan personas (*alumno, profesor, director, conserje, coordinador, bedel...*), dependencias de los centros, sean éstas del interior (*aula, bar, biblioteca, clase, comedor, conserjería, despacho, departamento, dirección, gimnasio, laboratorio, sala de profesores, salón de actos, secretaría...*) o del exterior (*aparcamiento, fuente, jardín, campo de fútbol, patio...*), e incluso fuera del recinto escolar (*árbol, autobús*). Las posibilidades asociativas se extienden también a los materiales de las dependencias escolares (*cafetera, colchoneta, cisterna, espaldera, fuente, manguera, mapa, pelota, plinto...*), la indumentaria (*cazadora, chándal, chaqueta, chubasquero, jersey, uniforme, zapatilla...*) y la actividad escolar (*advertencia, aprobado, apuntes, bocadillo, desayuno, estudiar, ejercicios, examen, expediente, horario, lista, nota, recreo, suspenso...*).

Como señala Hernández Muñoz (2006: 287), “lo más destacable de este campo nocional es que, junto a algunos términos normativos, figura una versión apocopada que pertenece a la jerga juvenil y/o estudiantil”: *bocadillo/bocata, compañeros/compis, matemáticas/matracas – mates*. El paréntesis se usa para los apócope que alternan con el término completo en *bolí(grafo), profe(sor), tele(-visión)*. Además se usa el paréntesis según indican las normas básicas de edición: *consejo (escolar), goma (de borrar), (lápiz de) color, pluma (estilográfica), órgano (musical), tablón (de anuncios)*, etc.

En cuanto a las formas derivadas, se mantienen como entrada separada aquellas que tienen pertinencia semántica: *barandilla, capuchón, carboncillo, carrito, cartilla, casilla, colorín, cuadernillo, cuchilla, falsilla, lapicero, plantilla, platillo, rejilla, sillón, tablón, zapatilla*.

Aparecen diferentes marcas comerciales que, aunque no están consignadas en los diccionarios de la lengua más usuales, en gran medida se encuentran lexicalizadas ya que se repiten de unas comunidades a otras. Estas marcas que dan nombre en su mayoría a utensilios habituales en el colegio son: *cutter, dyn A4, fixo, kleenex, plastidcor, pilot, rottring, tippex*.

La norma léxica común a los cinco léxicos disponibles comparados se compone de:

– Vocablos comunes en las 10 primeras posiciones: (7) *bolí(grafo), lápiz, libro, mesa, pizarra, silla, tiza*.

– Otros vocablos comunes en la franja de disponibilidad +0,1: (4) *borrador, estuche, goma (de borrar), pupitre*.

– Otros vocablos comunes entre las 50 primeras posiciones: (7) *carpeta, cuaderno, folio, pluma, puerta, regla, ventana*.

– Vocablos comunes a 4 comunidades: (17) *armario, cartabón, compás, es-cuadra, estantería, libreta, mapa, mochila, ordenador, papel, papelera, percha, perchero, persiana, portaminas, rotulador, video*.

– Vocablos comunes a 3 comunidades: (6) *calculadora, cartera, diccionario, fotocopiadora, tablón (de anuncios), tippex*.

**Cuadro 7.11. Comparación del léxico más disponible en el centro de interés ‘La escuela’**

Nº	Asturias	Cast. La Mancha	Sevilla	Valencia	Zamora
1	2	3	4	5	6
1	silla	profesor	silla	mesa	silla
2	mesa	pizarra	mesa	silla	mesa
3	bolí(grafo)	silla	pizarra	pizarra	bolí(grafo)
4	lápiz	mesa	lápiz	bolí(grafo)	tiza
5	libro	libro	bolí(grafo)	lápiz	libro
6	tiza	alumno	tiza	tiza	encerado

Cuadro 7.11 (cont.)

1	2	3	4	5	6
7	goma	bolí(grafo)	libro	libro	cuaderno
8	pizarra	tiza	cuaderno	goma (de borrar)	goma
9	encerado	pupitre	goma de borrar	libreta	pizarra
10	libreta	lápiz	borrador	pupitre	lápiz
11	borrador	clase	folio	estuche	pupitre
12	armario	cuaderno	estuche	armario	borrador
13	pupitre	estuche	carpeta	borrador	carpeta
14	carpeta	borrador	pupitre	carpeta	rotulador
15	video	aula	armario	cartera	folio
16	estuche	recreo	puerta	rotulador	lapicero
17	televisión	examen	ventana	folio	pluma
18	folio	patio	mochila	ventana	papel
19	mochila	director	sacapuntas	sacapuntas	papelera
20	regla	goma	rotulador	puerta	estuche
21	ventana	cartera	ordenador	percha	pintura
22	papelera	biblioteca	regla	pluma	mesa de profesor
23	perchero	mochila	perchero	regla	regla
24	rotulador	gimnasio	papelera	tablón de anuncios	sacapuntas
25	ordenador	compañero	libreta	papelera	tarima
26	papel	folio	mesa (del) profesor	compás	puerta
27	compás	carpeta	tippex	papel	hoja
28	puerta	niño	percha	tarima	ventana
29	percha	conserje	estantería	perchero	perchero
30	mapa	ventana	papel	carpesano	cartera
31	escuadra	mates(máticas)	fotocopiadora	tippex	armario
32	cartabón	cafetería	maleta	ordenador	tablón de anuncios
33	estantería	jefe de estudios	pluma	portaminas	mapa
34	tippex	puerta	compás	hoja	calendario
35	portaminas	pasillo	portaminas	video	compás
36	corcho	asignatura	escuadra	fotocopiadora	diccionario
37	hoja	amigo	goma	escuadra	percha
38	fotocopiadora	ordenador	cartabón	estantería	persiana
39	diccionario	suspensio	archivador	cartabón	portaminas
40	pluma	maestro	(lápiz de) color	mochila	radiador

1	2	3	4	5	6
41	diapositiva	secretaría	diccionario	escalera	escuadra
42	calculadora	nota	persiana	tele(visión)	clasificador
43	cortina	lapicero	agenda	cajón	cartabón
44	tajalápiz	estudiar	mapa	mapa	sillón
45	radiador	escalera	video	corcho	crucifijo
46	proyector	sacapuntas	tablón	cuaderno	video
47	color	regla	profesor	calculadora	lapicera
48	cuaderno	aprobado	tele(visor)	sillón	libreta
49	balón	conserjería	luz	persiana	estantería
50	persiana	secretario	bloc	banco	calculadora

### 7.2.9. 'Iluminación, calefacción y medios de airear un recinto'

La primera cuestión que influye en el contenido léxico de este centro es el título del mismo, que no ha sido siempre el mismo en todas las investigaciones del Proyecto Panhispánico, ya que los distintos investigadores lo han adaptado a las circunstancias climáticas específicas de la zona geográfica en cuestión. Ello, obviamente, ha llevado aparejado una modificación en el campo referencial de este centro de interés. Tres son las soluciones que se manejan respecto a la etiqueta de este centro de interés:

– ‘Calefacción e iluminación’. Algunos investigadores se mantienen fieles a la denominación original empleada por Gougenheim y sus colaboradores (‘le chauffage et l’éclairage’). Entre otras, siguen esta línea las investigaciones del País Vasco (Etxebarria Arostegui, 1996, 1999), Asturias (Carcedo González, 2001), Ávila, Salamanca y Zamora (Galloso Camacho, 2003a), Soria (Bartol Hernández, 2004a), Cuenca (Hernández Muñoz, 2004), Castilla-La Mancha (Hernández Muñoz, 2006), Alicante (Martínez Olmos, 2007), las zonas limítrofes del sur de España y Portugal (Prado Aragonés *et al.*, 2009), Galicia (López Meirama, 2011), Cantabria (Fernández Juncal, 2013), etc.

– ‘Iluminación y aire acondicionado’. Otros investigadores sustituyen el campo referencial de la ‘calefacción’ por el de ‘aire acondicionado’, dado el nulo uso de la calefacción en su comunidad. Este es el caso de las investigaciones en el Caribe, en la República Dominicana (Alba, 1995a) y Puerto Rico (López Morales, 1999a).

– ‘Iluminación, calefacción y medios de airear un recinto’. La tercera opción es la de unir los tres estímulos mencionados en las dos opciones anteriores bajo la misma etiqueta. Esta es la solución adoptada en las investigaciones de Gran Canaria (Samper Padilla, 1998b), Almería (Mateo García, 1998), Cádiz (González

Martínez, 2002), Aragón (Arnal Purroy *et al.*, 2004), Valencia (Gómez Devís, 2004)<sup>19</sup>, Ceuta (Ayora Esteban, 2006), Melilla (Fernández Smith *et al.*, 2008), Navarra (Areta Lara, 2009), Sevilla (Trigo Ibáñez, 2011), etc.

A pesar de las modificaciones en el nombre del centro de interés, “los resultados finalmente obtenidos no distan mucho de los de otras investigaciones” (Gómez Devís, 2004: 89), lo que hace posible las comparaciones interdialectales.

Sea como fuere, este es el área léxica que presenta menor productividad en casi todas las investigaciones que lo incluyen. Se trata de un centro de interés (ver Cuadro 7.12) que ofrece respuestas muy diversas y se dispersa en asociaciones secundarias. Los listados de este centro incluyen medios artificiales<sup>20</sup> para airear, iluminar y calentar un recinto (*aparato de calefacción, antorcha, bengala, bombilla, brasero, calefactor, extractor, generador, lámpara, radiador...*), medios naturales (*aire, brisa, claridad, corriente de aire, fiebre, luna, luz solar, sol, sombra, viento...*), los que hacen referencia a un recinto (*aire acondicionado, chimenea, climatizador, estufa, puerta, ventana...*), y también los asociados a la ropa (*abrigo, bufanda, calcetín, jersey, manta, sábana...*). Las posibilidades asociativas se extienden también a utensilios, aparatos y materiales que se emplean o intervienen en estos procesos de iluminación, calefacción o refrigeración de un recinto o persona (*abanico, carbón, cartón, cerilla, enchufe, gas, hielo, leña, paipay...*) y menciones precisas de distintas partes de los utensilios y aparatos (*casquillo, cable, filamento, fusible...*), verbos (*abanicarse, abrir, abrigarse, cerrar, soplar, ventilar...*) y adjetivos (*cálido, caliente, claro, eléctrico, halógeno, hidráulico, natural, solar...*).

Aparecen bastantes lexías complejas como *cocina de leña, estufa eléctrica, lámpara de pie, luz de emergencia, tubo fluorescente*, etc., con una presencia importante de verbos en construcción sintagmática con un complemento: *abrir/cerrar* (la puerta, la ventana), *encender* (la luz, una cerilla, una vela, un mechero), *frotarse las manos, hacer una hoguera, quitarse la ropa*, etc. Como de costumbre, se usa el paréntesis en los casos en los que alterna la presencia/ausencia de un elemento en estas lexías: *cable (de la luz), comunidad (de vecinos) corriente (eléctrica), extractor (de humo), (gas) butano, (abrir la) ventana*, etc.

Se mantienen las formas con diminutivo *bombilla, carbonilla, cerilla, cerillo, farolillo, hornillo, infernillo, lamparita, lamparilla, ventanilla*, separadas del lexema primitivo *bomba, carbón, cera, farol, horno, infierno, lámpara, ventana*, por encontrarse lexicalizadas. Se opera del mismo modo por la misma razón con las variantes flexivas *farol, farola y ventanal*.

<sup>19</sup> El título exacto que Gómez Devís (2004) da al centro 09 es ‘Iluminación, calefacción y ventilación’.

<sup>20</sup> Samper Padilla (1998a: 325) limitó a los informantes canarios el tipo de asociaciones permisibles a solo medios artificiales de iluminar, calentar o ventilar’.

**Cuadro 7.12. Comparación del léxico más disponible en el centro de interés  
'Iluminación y calefacción y medios de airear un recinto'**

Nº	Asturias	Cast. La Mancha	Sevilla	Valencia	Zamora
1	2	3	4	5	6
1	bombilla	bombilla	(aparato de) aire acondicionado	estufa	radiador
2	radiador	radiador	bombilla	aire acondicionado	bombilla
3	lámpara	lámpara	ventilador	ventilador	lámpara
4	estufa	luz	lámpara	bombilla	fluorescente
5	fluorescente	calor	ventana	lámpara	estufa
6	linterna	estufa	estufa	ventana	brasero
7	caldera	caldera	calefactor	radiador	caldera
8	calor	cable	flexo	linterna	flexo
9	luz	flexo	tubo fluorescente	chimenea	linterna
10	foco	fluorescente	linterna	flexo	vela
11	flexo	interruptor	radiador	abanico	foco
12	mechero	calefactor	vela	vela	cable
13	cable	brasero	calefacción	sol	calefactor
14	vela	electricidad	calentador	calefactor	interruptor
15	enchufe	foco	chimenea	tubo fluorescente	luz
16	calentador	enchufe	abanico	foco	carbón
17	interruptor	tubo	puerta	calefacción	chimenea
18	sol	gasoil	(foco) halógeno	fuego	tubería
19	electricidad	linterna	brasero	luz de neón	enchufe
20	farola	gas	luz	puerta	electricidad
21	gas	tubería	foco	calentador	sol
22	carbón	frío	mechero	hoguera	fuego
23	fuego	agua	extractor (de humo)	mechero	acumulador
24	cerilla	vela	fuego	lámpara halógena	calor
25	tubería	halógeno	sol	estufa de gas	mechero
26	chimenea	chimenea	(abrir/cerrar la) ventana	luz	gasóleo
27	calefactor	termostato	copa	estufa eléctrica	calentador
28	tubo	sol	candela	brasero	gas
29	agua	farola	cerilla	caldera	gasoil
30	gasoil	calentador	farola	extractor	leña

Cuadro 7.12 (cont.)

1	2	3	4	5	6
31	halógeno	gas natural	persiana	luz natural	petróleo
32	hoguera	mechero	rejilla	aire	cerilla
33	termostato	tubo fluorescente	caldera	cerilla	frío
34	cocina de carbón	calefacción	respiradero	manta	agua
35	faro	fuego	climatizador	estufa de leña	butano
36	candil	aire acondicionado	neón	viento	lumbre
37	madera	gasóleo	abrir la puerta	halógeno	fueloil
38	gasóleo	butano	hoguera	lamparilla	hoguera
39	horno	leña	lámpara fluorescente	placa solar	madera
40	manta eléctrica	carbón	interruptor	soplar	farola
41	resistencia	lumbre	candelabro	farola	termostato
42	energía	ventana	lámpara de mesa	abrir ventana	tubo
43	bombona	invierno	electricidad	refrigerador	contador
44	frío	bienestar	lámpara de pie	antorcha	bombona
45	brasero	fusible	luz de emergencia	balcón	resistencia
46	acumulador	energía	(copa de) cisco	gas	farol
47	cocina de leña	casquillo	cable	ojo de buey	portalámparas
48	generador	caliente	gas	leña	candelabro
49	leña	oscuridad	gas natural	persiana	araña
50	ventana	cerilla	termo	lumogas	quinqué

La norma léxica común a los cinco léxicos disponibles comparados queda como sigue:

– Vocablos comunes en las 10 primeras posiciones: (4) *bombilla, estufa, lámpara, radiador*.

– Otros vocablos comunes en la franja de disponibilidad +0,1: (3) *flexo, (tubo) fluorescente, linterna*.

– Otros vocablos comunes entre las 50 primeras posiciones: (14) *brasero, caldera, calefactor, calentador, cerilla, chimenea, farola, foco, fuego, gas, luz, mechero, sol, vela*.

– Vocablos comunes a 4 comunidades: (7) *cable, electricidad, (foco) halógeno, hoguera, interruptor, leña, ventana*.

– Vocablos comunes a 3 comunidades: (12) *agua, (aparato de) aire acondicionado, calefacción, calor, carbón, enchufe, frío, gasoil, gasóleo, termostato, tubería, tubo*.

### 7.2.10. 'La ciudad'

El centro de interés de 'La ciudad' (Cuadro 7.13) es uno de los más productivos en cuanto a número de vocablos, con un amplísimo abanico de posibilidades asociativas que incluye términos de distintas categorías gramaticales, con valores semánticos muy diversos. Así, junto a los sustantivos, categoría gramatical característica de los vocabularios disponibles (*alcalde, bachillerato, concierto, droga, gato, encrucijada...*), se incluye un gran número de adjetivos (*urbano, ancho, sucio...*), verbos en infinitivo (*bailar, consumir, vender...*) o lexías complejas (*bloque de viviendas, salir de noche, ir de compras, hacer cola, etc.*).

En las listas de vocabulario abundan los vocablos que nombran distintos tipos de vías (*acera, avenida, calle, camino, carretera, plaza, paseo, rambla, rotonda...*), partes de la ciudad (*afueras, barrio, casco antiguo, centro, jardín, periferia, parque, puerto, urbanización...*), medios de transporte (*autobús, autocar, bicicleta, moto, taxi, tranvía...*), así como otros sustantivos relacionados con la circulación (*atasco, embotellamiento, estación, parada, semáforo, tráfico, zona peatonal...*). Igualmente se encuentran lexías que evocan establecimientos de restauración y hotelería (*bar, cafetería, hostel, hotel, pensión, restaurante...*), de alimentación (*carnicería, centro comercial, frutería, mercado, panadería, pescadería, pastelería, supermercado, verdulería...*), de ocio (*cine, discoteca, gimnasio, museo, pub, teatro...*), establecimientos comerciales y mercantiles (*banco, estanco, farmacia, ferretería, floristería, joyería, librería, óptica, peluquería, quiosco, tienda...*). Asimismo, aparecen nombres de edificios y monumentos de interés público o turístico (*apartamento, casa, catedral, estadio, hospital, iglesia, monumento, museo, palacio, piso, rascacielos...*), así como de servicios públicos (*biblioteca, bomberos, comisaría, correos, hospital, oficina de turismo, policía...*), centros educativos (*academia, biblioteca, colegio, conservatorio, escuela, facultad, instituto, universidad...*), entidades de la administración (*ayuntamiento, juzgado, ministerio...*), oficinas (*correos, hacienda, gestoría, telefónica, seguros...*), personas (*amigo, anciano, gente, hombre, peatón...*), elementos del mundo natural (*árbol, césped, colina, fuente, jardín, pájaro, paloma, rata...*). También tienen importancia en este centro una elevada cantidad de vocablos abstractos que expresan todo tipo de percepciones subjetivas (*actividad, agobio, agotamiento, empujón, iluminación, nerviosismo, oportunidad, peligro, pobreza, prisa, trabajo, etc.*).

De los vocablos aportados por los encuestados, los investigadores tan sólo prescinden de los nombres propios de empresas o establecimientos no lexicalizados (*El Corte Inglés, Mac Donalds, Mercadona, Telepizza, Mango, Zara, etc.*)<sup>21</sup>,

<sup>21</sup> Hernández Muñoz (2006: 289) explica que elimina los nombres de supermercados, grandes almacenes, o cadenas comerciales porque no están presentes en todas las localidades donde se

nombres de marcas de automóviles (*BMW, Ford, Mercedes...*) y los nombres propios de ciudades y regiones (*Madrid, Barcelona, Castilla...*).

Atendiendo a la neutralización de las variantes flexivas se mantienen en plural *grandes almacenes, (oficina de) correos, recreativos, y ultramarinos*. En cuanto a las formas derivadas, adquieren la entidad de vocablos las voces con sufijos apreciativos se mantiene la distinción entre *calle/callejuela/callejón, carro/carrito, faroll/farolillo y mercado/mercadillo, plaza/plazoleta*, además de la oposición de género *faroll/farola*.

Respecto a la aparición de voces extranjeras, este campo es relativamente productivo: *bar, boutique, burger, container, graffiti, grunge, graffiti, heavy, parking, pub, sex shop, skin head, yonqui*, muchos de ellos asociados a actitudes y comportamientos de los jóvenes.

Por lo demás, se uniforman con paréntesis denominaciones completas e incompletas como *bloque (de casas), cajero (automático), centro (recreativo)*, etc., y las formas plenas y apocopadas en *(auto)bús, bici(cleta), cole(gio), disco(teca), parada de (auto)bús, super(mercado)*, etc.

**Cuadro 7.13. Comparación del léxico más disponible en el centro de interés ‘La ciudad’**

Nº	Asturias	Cast. La Mancha	Sevilla	Valencia	Zamora
1	2	3	4	5	6
1	coche	coche	coche	coche	coche
2	edificio	tienda	casa	moto(cicleta)	tienda
3	parque	parque	tienda	autobús	calle
4	tienda	calle	edificio	semáforo	parque
5	semáforo	edificio	calle	árbol	semáforo
6	calle	contaminación	semáforo	tienda	edificio
7	casa	semáforo	parque	edificio	acera
8	contaminación	casa	autobús	acera	bar
9	bar	gente	moto	parque	casa
10	autobús	(auto)bús	gente	calle	carretera
11	acera	moto	carretera	farola	jardín
12	gente	piso	contaminación	bici(cleta)	autobús
13	ruido	bar	piso	casa	contaminación
14	carretera	ruido	ruido	finca	gente
15	fuelle	carretera	bar	bar	discoteca
16	farola	discoteca	centro comercial	camión	avenida
17	banco	cine	acera	perro	piso

realizó la encuesta, sino sólo en los núcleos urbanos, y por lo tanto, [...] no pertenecen al universo general de todos los hablantes.

1	2	3	4	5	6
18	moto	plaza	cine	banco	plaza
19	colegio	metro	plaza	persona	supermercado
20	discoteca	super(mercado)	disco(teca)	jardín	colegio
21	rascacielos	tráfico	tráfico	señal de tráfico	moto
22	supermercado	acera	farola	papelera	instituto
23	cine	colegio	árbol	gente	banco
24	tráfico	árbol	bici(cleta)	colegio	farola
25	piso	farola	colegio	carretera	ayuntamiento
26	árbol	banco	humo	policía	árbol
27	paso de cebra	instituto	banco	paso de cebra	camión
28	restaurante	ayuntamiento	hospital	supermercado	rascacielos
29	instituto	hospital	monumento	taxi	ruido
30	ayuntamiento	atasco	ayuntamiento	gato	iglesia
31	universidad	jardín	puente	fuelle	bici(cleta)
32	plaza	restaurante	restaurante	quiosco	peatón
33	hospital	museo	iglesia	ayuntamiento	cine
34	camión	universidad	camión	avenida	hospital
35	centro comercial	fuelle	super(mercado)	restaurante	asfalto
36	humo	rascacielos	avenida	plaza	policía
37	rotonda	iglesia	instituto	discoteca	humo
38	jardín	paso de cebra	taxi	monumento	comercio
39	cafetería	avenida	persona	metro	biblioteca
40	policía	centro comercial	universidad	hospital	señal
41	iglesia	bici(cleta)	museo	contaminación	pub
42	museo	monumento	rascacielos	puente	catedral
43	perro	persona	comercio	cine	paso de cebra
44	bicicleta	camión	jardín	instituto	quiosco
45	metro	taxi	teatro	grandes almacenes	escuela
46	avenida	comercio	paso de peatones	tren	trafico
47	teatro	tren	atasco	humo	piscina
48	biblioteca	teatro	metro	iglesia	museo
49	pub	pub	polución	asfalto	teatro
50	persona	policía	río	niño	fuelle

La norma léxica común a los cinco léxicos disponibles comparados queda como sigue:

– Vocablos comunes en las 10 primeras posiciones: (6) *calle, coche, edificio, parque, semáforo, tienda*.

– Otros vocablos comunes en la franja de disponibilidad +0,1: (11) *acera, árbol, (auto)bús, banco, bar, carretera, casa, colegio, farola, gente, moto(cicleta)*.

– Otros vocablos comunes entre las 50 primeras posiciones: (13) *avenida, ayuntamiento, bici(cleta), camión, cine, contaminación, discoteca, hospital, iglesia, instituto, jardín, plaza, super(mercado)*.

– Vocablos comunes a 4 comunidades: (13) *fuelle, humo, metro, museo, paso de cebra, persona, piso, policía, rascacielos, restaurante, ruido, teatro, tráfico*.

– Vocablos comunes a 3 comunidades: (6) *centro comercial, comercio, monumento, pub, taxi, universidad*.

### 7.2.11. ‘El campo’

Este núcleo temático, de carácter abierto, es uno de los más prolíferos, caracterizado por un amplio elenco de asociaciones léxicas y múltiples percepciones subjetivas. Precisamente el nivel de subjetividad que caracteriza a las asociaciones de este centro en lo que respecta al grado de relación entre las distintas unidades y el tema nuclear, lleva a Samper Hernández, (2002: 38) a replantear el concepto de ‘campo’ como ‘lo no urbano’.

Entre el variado espectro de realidades que evoca ‘El campo’ (ver Cuadro 7.14), aparecen designaciones relacionadas con las tareas agrícolas y ganaderas (*agricultor, campesino, cosechadora, cultivo, granja, labrador, pastor, segador, regadío, tractor...*), los instrumentos del campo (*arado, azada, azadón, guadaña, hacha, hoz, rastrillo...*) y las construcciones rurales (*cabaña, casa rural, chalé, choza, cortijo, villa...*). Aparecen igualmente nombres de animales, tanto de granja (*buey, burro, caballo, cabra, cerdo, conejo, gallina, gallo, oveja, pavo, toro, vaca...*) y domésticos (*gato, perro*), como de aves (*cigüeña, pájaro, paloma...*), insectos (*abeja, bicho, cucaracha, escarabajo, hormiga, mariposa, mariquita, mosca, mosquito, saltamontes...*) y otros animales salvajes (*ardilla, ciervo, jabalí, lince, lobo, oso, zorro...*), así como nombres relacionados con lo vegetal y los cultivos (*árbol, arbusto, flor, hierba, matorral, planta, prado, pradera, viña...*) y nombres de árboles (*abeto, almendro, chopo, ciprés, higuera, manzano, olivo, pino, roble...*), de diversos frutos (*aceituna, castaña, champiñón, chumbo, nuez, piñón...*) y vocablos relacionados con la naturaleza y el paisaje (*aire, agua, arroyo, brisa, cerro, cielo, cordillera, cueva, lago, montaña, río...*).

Como indica Sánchez-Saus (2011: 416), “aparte de estas áreas, el centro ‘El campo’ recoge una gran cantidad de sustantivos asociados de manera indirecta con él, fundamentalmente debido a la vinculación de este con el ocio”, como son *aire fresco, barbacoa, bicicleta, caza, descanso, mochila, picnic, sol, vacaciones...* En relación con lo anterior, se actualizan, además, sustantivos abstractos (en su mayoría de carácter positivo), como *alegría, calma, felicidad, libertad, no*

*contaminación, no ruido*<sup>22</sup>, *paz, salud, silencio, tranquilidad*, etc., aunque también aparecen otros de carácter negativo, como *suciedad, aburrimiento, ruido*, etc., pero con índices inferiores a los obtenidos en ‘La ciudad’.

Esta diversidad léxica y asociativa propicia que las listas de léxico contengan elementos de distinta índole gramatical; aparte de los sustantivos (*casa, río, tractor...*), los adjetivos ganan peso en la muestra, apareciendo en solitario (*claro, fresco, limpio, rural, verde...*) o en compuestos sintagmáticos (*aire puro, roca caliza, vida familiar, zona quemada...*).

En cuanto a la neutralización de la variante género, debido a la diferencia semántica, se mantienen como entradas diferentes *caracol/caracola; charco/charca, fruto/fruta, gorro/gorra, huerto/huerta, manzano/manzana, segador/segadora, trillo/trilla*. Como indica Hernández Muñoz (2006: 290), un grupo de ellos, indica tanto a las personas que desempeñan una acción como a la máquina utilizada para el mismo fin: *cosechador/cosechadora, segador/segadora, cultivador/cultivadora, podador/podadora, sembrador/sembradora*. Otro grupo lo forman los nombres de los árboles y los frutos: *cerezo/cereza, ciruelo/ciruela, manzano/manzana, naranjo/naranja*, etc.

Asimismo, en el caso de las formas derivadas, es pertinente la diferenciación semántica de: *casa, caserón, caseta, casona; cigarro, cigarrillo; hierba, hierbajo; manzana, manzanilla; río, riachuelo; tomate, tomatera; trigo, trival*.

**Cuadro 7.14. Comparación del léxico más disponible en el centro de interés ‘El campo’**

Nº	Asturias	Cast. La Mancha	Sevilla	Valencia	Zamora
1	2	3	4	5	6
1	árbol	árbol	árbol	árbol	árbol
2	vaca	tractor	hierba	pino	flor
3	flor	flor	flor	hierba	tractor
4	hierba	animal	animal	naranjo	hierba
5	tractor	hierba	pájaro	piedra	vaca
6	animal	río	río	flor	río
7	prado	tierra	tranquilidad	tierra	tierra
8	río	montaña	tierra	montaña	trigo
9	casa	piedra	vaca	río	animal

<sup>22</sup> La léxias complejas del tipo *no contaminación, no polución, no ruido*, o *sin ruido, sin tráfico* designan en el centro 11 ‘El campo’ la ausencia de aspectos negativos característicos de los núcleos urbanos. Al respecto, Trigo Ibáñez (2011: 51) se suma a la idea expresada por Galloso Camacho (2003a) y corroborada por Ávila Muñoz (2006) de que, “si bien no son vocablos en sentido estricto, sí encierran una intención de expresar sentimientos o sensaciones que los informantes no son capaces de etiquetar lingüísticamente”.

Cuadro 7.14 (cont.)

1	2	3	4	5	6
10	caballo	pájaro	tractor	césped	pájaro
11	oveja	tranquilidad	pedra	manzano	oveja
12	tranquilidad	oveja	casa	arbusto	cebada
13	perro	conejo	montaña	pájaro	perro
14	pájaro	camino	perro	abeto	pedra
15	montaña	viña	caballo	limonero	arado
16	gallina	vaca	insecto	almendro	camino
17	bosque	casa	césped	margarita	montaña
18	cerdo	trigo	verde	tractor	verde
19	huerto	perro	conejo	peral	caballo
20	gato	arado	cabra	rosa	conejo
21	verde	planta	lago	roble	casa
22	naturaleza	pino	cerdo	casa	monte
23	tierra	olivo	camino	vaca	pino
24	granja	arbusto	cultivo	amapola	pradera
25	arado	naturaleza	oveja	encina	arbusto
26	toro	insecto	planta	camino	hormiga
27	camino	liebre	olivo	chopo	lago
28	cabra	cabra	agua	planta	sol
29	conejo	oliva	sol	conejo	agricultor
30	planta	cebada	aire	animal	insecto
31	cuadra	remolque	naturaleza	olivo	huerta
32	arbusto	lago	arbusto	acequia	aire
33	hormiga	agua	margarita	lago	bosque
34	insecto	aceituna	cortijo	matorral	remolque
35	campesino	césped	aire puro	romero	cerdo
36	monte	aire	toro	ciprés	cosechadora
37	pradera	encina	amapola	hormiga	roble
38	margarita	jabalí	pino	agua	encina
39	pedra	verde	gallina	insecto	tranquilidad
40	agricultor	caballo	naranja	tomillo	mosquito
41	sol	agricultor	granja	oveja	avena
42	campo	amapola	hormiga	ardilla	remolacha
43	segadora	huerta	bicho	hoja	labrador
44	silencio	pastor	arena	higuera	agua
45	mariposa	sol	bosque	sauce	prado

1	2	3	4	5	6
46	aire	cepa	paz	granja	gallina
47	maíz	margarita	silencio	roca	granja
48	aire puro	granja	chalé	fuelle	sembrado
49	cultivo	ciervo	piscina	mosca	azada
50	guadaña	bosque	gato	chalé	chopo

La norma léxica común a los cinco léxicos disponibles comparados es la siguiente:

- Vocablos comunes en las 10 primeras posiciones: (4) *árbol, flor, hierba, río*.
- Otros vocablos comunes en la franja de disponibilidad +0,1: (3) *montaña, pájaro, tractor*.
- Otros vocablos comunes entre las 50 primeras posiciones: (11) *animal, arbusto, camino, casa, conejo, granja, insecto, oveja, piedra, tierra, vaca*.
- Vocablos comunes a 4 comunidades: (13) *agua, aire, bosque, caballo, hormiga, lago, margarita, perro, pino, planta, sol, tranquilidad, verde*.
- Vocablos comunes a 3 comunidades: (10) *agricultor, amapola, arado, cabra, cerdo, césped, encina, gallina, naturaleza, olivo*.

### 7.2.12. ‘Medios de transporte’

El centro de interés ‘Medios de transporte’ es uno de los más concretos, con una alta cohesión temática. A su vocabulario (ver Cuadro 7.15) pertenecen los nombres de medios de transporte de muy diverso tipo; así, los medios terrestres son los más numerosos, tanto de dos ruedas (*ciclomotor, motocicleta, scooter, vespa...*), como de cuatro o más, tanto de uso privado como público (*autobús, autocar, camión, camioneta, coche, furgoneta, microbús, minibús, taxi, todoterreno...*) y los que se desplazan sobre vías (*AVE, metro, talgo, tranvía, tren...*). Están presentes también las designaciones de aeronaves (*aeroplano, avión, avioneta, globo, helicóptero...*), así como las de embarcaciones (*barca, barco, barco de vela, bote, canoa, catamarán, ferry, lancha, piragua, submarino, trasatlántico, velero, yate...*).

En este centro se actualizan, además, cuantiosos nombres de animales (*asno, burro, buey, caballo, camello, mulo...*) y de vehículos y juguetes sin motor (*bicicleta, carro, carroza, coche de caballos, monopatín, patín, patinete, tándem, silla de ruedas, trineo...*), e incluso los medios de transporte atípicos – “artificiales”, según la terminología de Samper Padilla (1998a: 329) – (*avión de combate, bombardero, cohete, OVNI, satélite, zeppelin...*). El centro también incluye menciones de las diferentes partes de estos medios (*batería, bujía, caja de cambio, capota, chapa, embrague, hélice, rueda, tuerca...*), productos para su funcionamiento

(*anticongelante, combustible, diésel, gasolina, fuel...*), efectos producidos por los medios (*accidente, atasco, explosión, multa, ruido, transbordo, viaje...*), lugares relacionados con los medios (*agencia, aeropuerto, concesionario, estación, parada, puerto...*), etc.

En general, se eliminan los nombres de marcas comerciales (BMW, *Mercedes, SEAT, Toyota*, etc.), al igual que se desestiman los nombres de compañías<sup>23</sup> aéreas y navieras (*Iberia, Transmediterránea, Vueling...*), aunque se mantiene RENFE, que asume el valor genérico de *tren*, y las marcas lexicalizadas *jeep, scooter, seiscientos, vespa, vespino, zodiac*. También se conservan en las listas los nombres de modelos de un medio de transporte: *jet, boeing, jumbo, concorde*, en avión; *berlina, cupé, monovolumen*, en automóvil; *AVE, interciti, talgo*, en tren.

En cuanto a los derivados, se mantiene la diferenciación de los que contienen evocaciones distintas a sus formas simples: *avión/avioneta, barco/barca, coche/cochecito, camión/camioneta, carreta/carretilla, carro/carrillo, patín/patinete, vagón/vagoneta*.

En este centro aparecen varias respuestas que empiezan por preposición (*a, en*): *a caballo, en burro*, etc.). Todas se lematizan sin la preposición, salvo *a pie, a pata, a cuestras y a paso ligero*.

La lematización hace que se reduzcan a una sola entrada *vapor y barco de vapor* → *barco (de vapor)*, *velero y barco velero* → *barco (velero)*, *cercanías y tren de cercanías* → *(tren de) cercanías*, y así en casos similares. La lematización también hace que las variantes morfológicas infinitivo/gerundio (*andar/andando*) se reduzcan a una sola entrada. Autores como Bartol Hernández (2004b: 1087) las resuelven a favor de la más frecuente, generalmente el gerundio: *andando, volando, corriendo*. Por su parte, Gómez Devís (2004: 92) y Hernández Muñoz (2006: 292) las unifican usando el paréntesis: *anda(ndo), corr(iendo), camin(ando), nada(ando)*.

Se usa también el paréntesis en los acortamientos (*auto*)bús y *bici*(cleta). Un caso especial lo representan las formas *moto* y *motocicleta*. A diferencia de González Martínez (2002) o Gómez Devís (2004), la mayoría de autores suelen consignarlas como dos entradas separadas, de acuerdo con la propuesta de Samper Padilla:

Puede llamar la atención entonces que hayamos mantenido la diferencia entre *motocicleta* y *moto*. La razón es que en nuestra comunidad de habla estos últimos vocablos (al menos mayoritariamente) designan realidades diferentes (la primera se emplea para referirse a vehículos de menor cilindrada). Que este es un uso extendido entre los sujetos encuestados se puede comprobar cuando el mismo alumno aporta las dos formas en el listado correspondiente

<sup>23</sup> Carcedo González (2001: 47) incluye, sin embargo, el nombre de la empresa de transporte por carretera ALSA, en la creencia de que “cumple entre los asturianos la función de los genéricos *autobús* o *autocar*”.

a este centro de interés mientras que eso no ocurre con *bicicleta/bici* (aquí se elige una de las dos; nunca ambas en la misma lista)<sup>24</sup> (Samper Padilla, 1998a: 328).

Como ya hemos indicado anteriormente, se mantienen las siglas y acrónimos *AVE*, *OVNI*, *RENFE*, *TALGO*, como claros ejemplos de lexicalización.

Por otra parte, La presencia de voces extranjeras es muy notable en anglicismos: *autostop*, *ferry*, *footing*, *hovercraft*, *jeep*, *jet*, *kart*, *mountain bike*, *skate-board*, *snowboard*...

En cuanto a la norma léxica común a los cinco léxicos disponibles comparados, esta queda como sigue:

– Vocablos comunes en las 10 primeras posiciones: (8) *avión*, *(auto)bús*, *barco*, *bici(cleta)*, *camión*, *coche*, *moto*, *tren*.

– Otros vocablos comunes en la franja de disponibilidad +0,1: (9) *avioneta*, *caballo*, *furgoneta*, *helicóptero*, *metro*, *monopatín*, *patín*, *tranvía*, *triciclo*.

– Otros vocablos comunes entre las 50 primeras posiciones: (20) *aeroplano*, *ala delta*, *a pie*, *barca*, *(barco) velero*, *burro*, *camioneta*, *canoa*, *carro*, *ciclomotor*, *globo*, *lancha*, *motocicleta*, *patinete*, *sidecar*, *submarino*, *taxi*, *tractor*, *transatlántico*, *yate*.

– Vocablos comunes a 4 comunidades: (5) *autocar*, *automóvil*, *AVE*, *camello*, *esquí*.

– Vocablos comunes a 3 comunidades: (4) *anda(ndo)*, *cohete*, *piragua*, *tráiler*.

**Cuadro 7.15. Comparación del léxico más disponible en el centro de interés ‘Medios de transporte’**

Nº	Asturias	Cast. La Mancha	Sevilla	Valencia	Zamora
1	2	3	4	5	6
1	coche	coche	coche	coche	coche
2	avión	(auto)bús	bici(cleta)	avión	bici(cleta)
3	(auto)bús	avión	autobús	moto(cicleta)	avión
4	bici(cleta)	bici(cleta)	avión	bici(cleta)	autobús
5	tren	moto	moto	(auto)bús	tren
6	moto	tren	tren	barco	moto
7	barco	barco	barco	tren	barco
8	camión	metro	camión	camión	camión
9	metro	camión	patinete	helicóptero	metro
10	patín	patín	monopatín	avioneta	patín

<sup>24</sup> Sobre esta diferenciación, Carcedo González (2001: 49) aporta un razonamiento similar: “se mantienen separados *moto* y *motocicleta*, pues, a pesar de que la primera se considera habitualmente forma abreviada (DRAE, 1992, s.v.), los informantes asturianos parecen entender que designan realidades distintas: una buena parte de ellos proporciona las dos”.

Cuadro 7.15 (cont.)

1	2	3	4	5	6
11	caballo	taxi	patines	patín	caballo
12	taxi	monopatín	metro	metro	triciclo
13	helicóptero	helicóptero	taxi	monopatín	helicóptero
14	furgoneta	furgoneta	helicóptero	barca	motocicleta
15	monopatín	patinete	furgoneta	tranvía	monopatín
16	avioneta	triciclo	motocicleta	furgoneta	burro
17	tranvía	motocicleta	caballo	lancha	tranvía
18	burro	caballo	tranvía	taxi	furgoneta
19	triciclo	tranvía	avioneta	triciclo	tractor
20	a pie	tractor	barca	globo	carro
21	lancha	avioneta	tractor	transatlántico	barca
22	tractor	carro	triciclo	submarino	avioneta
23	motocicleta	a pie	lancha	(barco) velero	patinete
24	patinete	ciclomotor	ciclomotor	yate	autocar
25	carro	lancha	globo (aerostático)	ala delta	submarino
26	barca	barca	(ir a) pie	caballo	taxi
27	globo	burro	AVE	ciclomotor	ciclomotor
28	autocar	globo	burro	aeroplano	globo
29	transatlántico	anda(ndo)	patera	carro	lancha
30	AVE	AVE	carro	patinete	aeroplano
31	submarino	patera	transatlántico	camioneta	camello
32	camioneta	trasatlántico	yate	canoa	yate
33	ciclomotor	submarino	patín	paracaídas	mulo
34	ferry	yate	camioneta	moto acuática	trasatlántico
35	sidecar	autocar	coche de caballos	tractor	a pie
36	aeroplano	tráiler	submarino	a pie	ala delta
37	yate	ala delta	aeroplano	zeppelin	andando
38	ALSA	camioneta	sidecar	cohete	canoa
39	(barco) velero	mulo	todoterreno	parapente	cohete
40	piragua	todoterreno	canoa	burro	trineo
41	ala delta	aeroplano	microbús	bote	balsa
42	canoa	monovolumen	andar	esquí	piragua
43	camello	teleférico	automóvil	portaaviones	velero
44	trineo	esquí	cohete	sidecar	elefante
45	teleférico	minibús	(barco) velero	balsa	esquí

1	2	3	4	5	6
46	esquí	remolque	ala delta	automóvil	sidecar
47	automóvil	camello	andando	piragua	AVE
48	minibús	canoa	autocar	ferry	tráiler
49	microbús	velero	monovolumen	buque	camioneta
50	moto acuática	automóvil	camello	tráiler	pie

### 7.2.13. ‘Trabajos del campo y del jardín’

En palabras de Bartol Hernández (2004b: 1088), “la característica más importante de este centro de interés es la presencia en las respuestas de abundantes verbos en infinitivo solos o acompañados de un sintagma nominal que especifica su significado” (ver Cuadro 7.16):

– *cortar, cortar árboles, cortar el césped, cortar flores, cortar hierba, cortar ramas...*

– *cuidar, cuidar animales, cuidar el huerto, cuidar el rebaño, cuidar flores...*

De esta forma, los verbos que designan las actividades habitualmente realizadas en el campo o en el jardín se presentan en infinitivo (*abonar, fumigar, plantar, podar, recoger, recolectar, regar, segar, sembrar, trasplantar...*), o en combinación con un complemento directo –CD– (*cortar flores, dar de comer a los animales, echar abono, pintar la valla, recoger patatas, sacar patatas...*). Según Hernández Muñoz (2006: 293), los núcleos verbales que documentan mayor número de combinaciones con un CD son: *arrancar, arreglar, atar, coger, cortar, cuidar, cultivar, echar, hacer, limpiar, plantar, podar, poner, quemar, quitar, recoger, regar, sacar, segar, sembrar*. No obstante, en la mayoría de los casos se documenta también el verbo sin el complemento (generalmente con alto índice de disponibilidad). La excepción a esta regla la constituyen los verbos *poner, echar, hacer* y *coger*, que no aparecen sin el complemento, debido con toda probabilidad “a que se trata de verbos de alcance general (“atemáticos”) que no se circunscriben al campo temático de los trabajos del campo y del jardín y por esa razón los informantes siempre hacen explícito el complemento que concreta su significado”.

En cuanto a la lematización, si bien Hernández Muñoz (2006) mantiene todos los casos de combinaciones, dadas las múltiples variantes de las construcciones verbales complementadas, Trigo Ibáñez (2011: 53) ve realiza la unificación en la expresión más genérica: *cortar hierba/cortar las hierbas/cortar hierbas/cortar la hierba* → *cortar la hierba*; *ordeñar vaca/ordeñar la vaca/ordeñar las vacas* → *ordeñar las vacas*, etc. Sin embargo, anteriormente Bartol Hernández (2004b: 1088) razonaba que “parece lógico pensar que en el cálculo de la disponibilidad de *cuidar* no solo deben tenerse en cuenta las apariciones del infinitivo

aislado. También habría que tener en cuenta las otras ocurrencias”. Para ello, este estudioso, apoyándose en el criterio de Carcedo González (2001) y Hernández Muñoz (2002, 2004), unifica mediante el uso del paréntesis el infinitivo aislado y complementado (por ejemplo, *alimentar/alimentar a los animales* → *alimentar (a los animales)*; *arar/arar la tierra* → *arar (la tierra)*, etc.). De acuerdo con este proceder, Bartol Hernández (2004b: 1088) lleva a cabo la edición de las encuestas sorianas, “procurando en la medida de lo posible unificar las variantes para que los índices de disponibilidad se ajusten lo más posible a la realidad”, obteniendo las entradas *abonar (la tierra)*, *arar (campo)*, *arrancar (hojas)*, *cavar (tierra)*, *criar (animales)*, *empacar (hierba)*, *matar (animales)*, *oxigenar (tierra)*, *rastrillar (tierra)*, *remover (tierra)*, *talar (árboles)*, *vender (productos)*.

Pero, además de los verbos con o sin complemento, los sustantivos, también abundantes. Estos aparecen relacionados con:

a) los medios técnicos, las herramientas y maquinaria (*arado, aspersor, azada, carreta, carro, cubo, bomba, escalera, espantapájaros, fertilizante, guante, hoz, manguera, martillo, noria, pala, pico, rastrillo, sierra, tijera, tractor, tridente...*),

b) las diferentes actividades agrícolas y ganaderas (*abonado, agricultura, cultivo, fumigación, labranza, poda, recolección, riego, siega, siembra, tala, trasplante, vendimia...*),

c) las personas que ejecutan los diferentes trabajos del área (*agricultor, apicultor, bracero, campesino, florista, ganadero, granjero, jardinero, jornalero, labrador, pescador, podador, vaquero, veterinario...*),

d) los animales utilizados en las labores del campo (*burro, buey, caballo, mulo*).

En cuanto a la consideración del léxico pertinente de este – y otros centros –, los trabajos donde se relaciona disponibilidad léxica y psicolingüística adoptan un criterio laxo y aceptan la mayoría de respuestas, respetando las asociaciones mentales de los informantes. Otros autores se ciñen más a la etiqueta categorial, y en este centro desechan los términos que dan nombre a los productos que se cultivan (*patata, lechuga, manzana, fresa*), así como los de árboles y flores, excepto cuando estos nombres aparecen como complemento del verbo que designa la acción (*recoger fresas, recoger maíz, recoger uvas, vender huevos, vender leche*), o excepto cuando se trata de la denominación genérica: *árbol, flor, fruto, legumbre*. Así se hace en los léxicos disponibles de Cádiz (González Martínez, 1998), Las Palmas (Samper Padilla, 1998b), Asturias (Carcedo González, 2001), Valencia (Gómez Devís, 2004), Ceuta (Ayora Esteban, 2006), etc.

En lo que atañe al género, la tarea de lematización debe respetar la oposición genérica – *or/-ora*, dado que la forma masculina evoca profesiones y la femenina maquinaria: *cosecador/cosecadora, podador/podadora, segador/segadora, sulfatador/sulfatadora*. Asimismo, el valor de pertinencia semántica requiere mantener las variantes *cubo/cuba, jardinero/jardinera, regadero/regadera, trillo/trilla*.

**Cuadro 7.16. Comparación del léxico más disponible en el centro de interés  
‘Trabajos del campo y del jardín’**

Nº	Asturias	Cast. La Mancha	Sevilla	Valencia	Zamora
1	2	3	4	5	6
1	segar	podar	regar	regar	sembrar
2	plantar	regar	sembrar	podar	arar
3	regar	arar	arar	plantar	podar
4	podar	sembrar	podar	arar	regar
5	sembrar	jardinero	recolectar	sembrar	abonar
6	arar	plantar	jardinero	recolectar	plantar
7	abonar	agricultor	plantar	abonar	cosechar
8	recolectar	abonar	agricultor	segar	segar
9	jardinero	vendimiar	segar	cortar	recolectar
10	agricultor	segar	abonar	recoger	recoger
11	cultivar	recolectar	cortar	cultivar	cavar
12	cortar	ganadero	recoger	labrar	cultivar
13	recoger	cosechar	cortar (el) césped	cavar	cortar
14	ganadero	cultivar	cultivar	fumigar	jardinero
15	cuchar	cavar	ganadero	jardinero	fumigar
16	cavar	cortar	fumigar	trasplantar	trillar
17	trasplantar	pastor	talar	agricultor	injertar
18	ordeñar	recoger	cuidar	cortar el césped	agricultor
19	fumigar	labrar	cosechar	injertar	sulfatar
20	injertar	sarmentar	limpiar	vendimiar	escardar
21	limpiar	fumigar	trasplantar	cosechar	trasplantar
22	cosechar	trasplantar	recolector	limpiar	vendimiar
23	labrar	cortar el césped	ordeñar	talar	ganadero
24	pastor	podador	labrar	azada	talar
25	sulfatar	vendimia	escardar	sulfatar	siembra
26	cuidar	recolector	jornalero	trillar	cosecha
27	empacar	injertar	jardinería	labrador	aricar
28	ensillar	tractorista	sembrador	tractor	limpiar
29	talar	coger aceitunas	podador	arrancar	apicultor
30	apicultor	sulfatar	agricultura	tijera	cuidar
31	sallar	cuidar	labrador	cuidar	empacar
32	agricultura	labrador	injertar	pala	entresacar
33	labrador	sembrador	siembra	pulverizar	arrancar

Cuadro 7.15 (cont.)

1	2	3	4	5	6
34	sachar	limpiar	aceitunero	sarmentar	podador
35	florista	agricultura	granjero	quitar malas hierbas	labrar
36	cortar el césped	talar	cavar	rastrillo	agricultura
37	ganadería	jardinería	sulfatar	cortacésped	pastor
38	granjero	segador	pastor	manguera	recolección
39	segador	vendimiador	tractorista	remover	recolector
40	tijeras	leñador	arador	excavar	rastrillar
41	segadora	cazador	ganadería	arado	sembrador
42	siembra	granjero	trillar	expurgar	arado
43	plantar flores	trillar	leñador	abono	vendimia
44	recolección	apicultor	recogida	recolector	hortelano
45	podadera	poda	cortar (las) flores	quemar	cortar el césped
46	arado	cosechador	segador	arreglar	envenenar
47	pala	ordeñar	echar abono	podador	remover
48	campesino	forestal	recolección	hoz	leñador
49	jardinería	varear	desinfectar	<i>empeltar</i>	horticultor
50	catar	aceituna	dar de comer a (los) animales	<i>collir</i>	jardinería

En este centro, la norma léxica común a los cinco léxicos disponibles comparados es:

– Vocablos comunes en las 10 primeras posiciones: (7) *abonar, arar, plantar, podar, regar, segar, sembrar*.

– Otros vocablos comunes en la franja de disponibilidad +0,1: (4) *cortar, cultivar, jardinero, recolectar*.

– Otros vocablos comunes entre las 50 primeras posiciones: (13) *agricultor, cavar, cortar el césped, cosechar, cuidar, fumigar, injertar, labrar, limpiar, recoger, sulfatar, talar, trasplantar*.

– Vocablos comunes a 4 comunidades: (8) *agricultura, ganadero, jardinería, labrador, pastor, podador, recolector, trillar*.

– Vocablos comunes a 3 comunidades: (9) *apicultor, arado, granjero, ordeñar, recolección, segador, sembrador, siembra, vendimiar*.

#### 7.2.14. ‘Los animales’

El centro de interés ‘Los animales’ (Cuadro 7.17) es un campo referencial con un carácter muy concreto, de bastante precisión semántica, lo que facilita la labor

de edición. El vocabulario de este centro recoge, en su mayoría, sustantivos que nombran todo tipo de animales. Así, aparte de las denominaciones genéricas que subclasifican a los animales de acuerdo a rasgos diversos (*anfibio, arácnido, artrópodo, bacteria, carroñero, crustáceo, felino, herbívoro, insecto, invertebrado, mamífero, pájaro, parásito, pez, simio...*), la nomenclatura referida a las especies animales incluye: los animales domésticos (*canario, hámster, gato, perro*<sup>25</sup>...), los animales de granja (*burro, caballo, cabra, cerdo, conejo, gallina, gallo, mulo, oveja, pato, pavo, toro, vaca...*), los mamíferos salvajes (*ardilla, cebra, ciervo, elefante, gorila, hipopótamo, jabalí, jirafa, león, lobo, mono, oso, pantera, rata, ratón, zorro...*), las aves (*águila, avestruz, buitres, cigüeña, cisne, flamenco, gorrion, gaviota, loro, pájaro, paloma, perdiz...*), los reptiles (*camaleón, cobra, cocodrilo, culebra, lagarto, serpiente, tortuga...*) y los insectos y otros invertebrados (*abeja, araña, caracol, cucaracha, escarabajo, escorpión, gusano, hormiga, mariposa, mariquita, mosca, mosquito, saltamontes...*), y los peces y otros animales marinos (*atún, bacalao, ballena, delfín, foca, gamba, medusa, pescado, pulpo, salmón, sardina, tiburón...*), así como las subespecificaciones de determinados animales (*boxer, bulldog, caniche, cocker, dálmata, galgo, gato montés, labrador, macaco, oso panda, oso polar, pitbull, tiburón blanco, títi...*), y animales míticos y fantásticos (*dragón, minotauro, unicornio...*).

También integran este conjunto léxico los nombres de las partes del animal (*antena, colmillo, pata, piel, pluma, trompa*), los nombres de objetos que pueden llevar algunos animales (*bozal, cascabel...*), de lugares donde se crían algunos de ellos (*granja, perrera, zoológico...*) y de sus alimentos (*pienso...*).

La neutralización de las variantes meramente flexivas adopta la forma no marcada, salvo cuando macho y hembra poseen lexemas distintos o sufijos específicos para el femenino, como en los sustantivos heterónimos (*caballo/yegua; vaca/toro*), los que presentan un sufijo específico en el femenino (*gallo/gallina*) y en los casos en los que el femenino remite a significados totalmente distintos al masculino (*bicho/bicha; caballo/caballa; caracol/caracola; lagarto/lagartija*). Se conservan asimismo los diminutivos lexicalizados *cabrito, cochinito, mariquita, periquito, perrillo de la pradera, salmonete*.

Se recurre al paréntesis para las variantes completas/incompletas, faltando en esta última el primer término, de carácter genérico. Esto da entradas del tipo: (*perro*) *pequinés*, (*oso*) *panda*, (*gato*) *persa*, (*serpiente*) *pitón*, (*pez*) *raya*.

La norma léxica común a los cinco léxicos disponibles comparados consta de:

– Vocablos comunes en las 10 primeras posiciones: (7) *caballo, elefante, gato, león, perro, tigre, vaca*.

– Otros vocablos comunes en la franja de disponibilidad +0,1: (23) *águila, cabra, canario, cerdo, cocodrilo, ballena, conejo, gallina, jirafa, lobo, loro,*

---

<sup>25</sup> *Perro y gato* –por este orden– son las dos palabras más disponibles en los léxicos disponibles de todas las comunidades españolas.

*mono, oso, oveja, pájaro, pantera, pato, rata, ratón, rinoceronte, serpiente, tiburón, toro.*

– Otros vocablos comunes entre las 50 primeras posiciones: (8) *avestruz, burro, cebra, hipopótamo, hormiga, jabalí, pez, zorro.*

– Vocablos comunes a 4 comunidades: (6) *ciervo, delfín, gallo, leopardo, mosca, tortuga.*

– Vocablos comunes a 3 comunidades: (5) *gorila, gorrión, hámster, paloma, periquito.*

**Cuadro 7.17. Comparación del léxico más disponible en el centro de interés ‘Animales’**

Nº	Asturias	Cast. La Mancha	Sevilla	Valencia	Zamora
1	2	3	4	5	6
1	perro	perro	perro	perro	perro
2	gato	gato	gato	gato	gato
3	vaca	león	león	león	león
4	caballo	caballo	caballo	tigre	caballo
5	león	tigre	vaca	elefante	tigre
6	tigre	elefante	tigre	caballo	vaca
7	elefante	vaca	elefante	serpiente	elefante
8	gallina	conejo	serpiente	mono	burro
9	cerdo	ratón	pájaro	águila	conejo
10	ratón	gallina	jirafa	jirafa	gallina
11	toro	jirafa	ratón	canario	cerdo
12	burro	serpiente	cerdo	periquito	lobo
13	conejo	oveja	toro	ballena	ratón
14	oveja	canario	mono	vaca	oveja
15	serpiente	cerdo	conejo	pantera	jirafa
16	oso	pez	oso	tiburón	toro
17	mono	mono	ballena	loro	pantera
18	ballena	oso	gallina	conejo	canario
19	lobo	pájaro	canario	pájaro	cabra
20	jirafa	ballena	cabra	leopardo	serpiente
21	cabra	águila	águila	oso	zorro
22	tiburón	cabra	tiburón	cerdo	mulo
23	águila	delfín	oveja	ratón	pato
24	pez	burro	pez	toro	águila
25	pájaro	tiburón	rata	cocodrilo	rinoceronte
26	leopardo	toro	delfín	gallina	gallo
27	canario	periquito	burro	rata	ballena

1	2	3	4	5	6
28	cocodrilo	liebre	gorrión	hipopótamo	buey
29	pato	jabalí	lobo	delfín	leopardo
30	zorro	rata	loro	mosca	mono
31	gallo	loro	cocodrilo	rinoceronte	oso
32	rinoceronte	tortuga	leopardo	pez	rata
33	jabalí	cocodrilo	ciervo	hormiga	loro
34	pantera	pato	pantera	avestruz	pájaro
35	loro	zorro	pato	cabra	yegua
36	rata	paloma	cebra	cebra	cocodrilo
37	delfín	cebra	avestruz	oveja	tiburón
38	cebra	ciervo	hormiga	tortuga	periquito
39	yegua	pantera	pollo	pato	liebre
40	gorrión	lobo	tortuga	lobo	avestruz
41	hormiga	araña	rinoceronte	gorila	jabalí
42	periquito	hipopótamo	hipopótamo	hiena	hormiga
43	tortuga	rinoceronte	gallo	gorrión	paloma
44	mosca	avestruz	ardilla	mosquito	pez
45	hámster	perdiz	lince	zorro	culebra
46	avestruz	gallo	gorila	burro	camello
47	paloma	mosca	mosca	ciervo	cebra
48	camello	hormiga	jabalí	jabalí	hipopótamo
49	hipopótamo	gorila	araña	chimpancé	perdiz
50	ciervo	hámster	zorro	buitre	hámster

### 7.2.15. 'Juegos y distracciones'<sup>26</sup>

El centro de interés 'Juegos y distracciones' (Cuadro 7.18) es un área temática que posee un carácter heterogéneo y difuso, con muchas posibilidades en la naturaleza de las asociaciones, lo que lo convierte en un centro difícil de lematizar. El problema afecta no solo a las respuestas que aparentemente no aluden a elementos del campo semántico (asociaciones periféricas), a nombres de juegos desconocidos y a una gran cantidad de marcas comerciales, sino también a las abundantes combinaciones de infinitivo + sustantivo (precedido o no de preposición), que aparecen junto a respuestas en las que solo aparece el sustantivo. Así se presentan los pares: *contar chistes/chistes*; *hacer deporte/deporte/s*; *hacer footing/footing*;

<sup>26</sup> Entre los léxicos disponibles que comparamos, Hernández Muñoz (2006) lo denomina 'Juegos y diversiones'. Bartol Hernández (2004) también utiliza esta etiqueta.

*ir al campo/campo; ir al cine/cine; ir al teatro/teatro; ir en moto/moto/s; jugar a las cartas/cartas; jugar al escondite/escondite; jugar al ordenador/ordenador; salir de fiesta/fiesta/s; saltar a la comba/comba; tocar la guitarra/guitarra; ver la tele(visión)/tele(visión);* y muchas otras. Incluso hay ocasiones en que la variación afecta al verbo y en las que se cuenta con una terna de opciones (*ir al fútbol/jugar al fútbol/fútbol; ir en bicicleta/montar en bicicleta/bicicleta; oír música/escuchar música/música; quedar con amigos/salir con amigos/amigos*).

Nuevamente entre los léxicos que comparamos –y en general entre los investigadores de la disponibilidad léxica– hay dos posturas diferentes respecto a la edición de los materiales:

a) De una parte, Samper Padilla (1998a), Carcedo González (2001), González Martínez (2002), Gómez Devís (2004), Ayora Esteban (2006), etc., siguiendo las pautas establecidas por Mackey (1971a) fusionan las combinaciones verbo + complemento nominal (*montar a caballo, ir al cine, ir/salir con los amigos, comer/cenar en el restaurante, hablar por teléfono, jugar a las cartas, leer el periódico, hacer footing, tocar la guitarra, tomar el sol, escuchar música*) con el núcleo nominal: *caballo, cine, amigos, restaurante, teléfono, cartas, periódico, footing, guitarra, sol, música*. Y lo mismo ocurre cuando junto al nombre de la distracción aparece el genérico *juego* (*juego de cartas, juego de pelota*, etc. se reducen a *cartas* y *pelota*, respectivamente, *juego de oca, jugar a la oca* y *oca en oca*). Estos autores solo mantienen entradas diferentes para aquellas combinaciones en las que aparece un verbo o sustantivo que realiza una asociación sintagmática no esperable, ante la necesidad de deshacer la ambigüedad: *componer música, escribir cartas, limpiar la casa, juego de ordenador, salir a tomar algo, salir de juerga*, etc. (Carcedo González, 2000a: 83; Gómez Devís, 2004: 94).

b) De otra parte, López Morales (1999a), Galloso Camacho (2002b), Hernández Muñoz (2002, 2004, 2006), Trigo Ibáñez (2011), etc. editan como entradas diferentes todas y cada una de las construcciones presentes en las encuestas. Hernández Muñoz (2006: 295) argumenta que entradas como *flauta* y *tocar la flauta* pueden referirse a acciones distintas, y que no consignar todas las opciones por separado supone una manipulación de los materiales, además de que el criterio empleado por los autores arriba mencionados “reestructura completamente los rangos de disponibilidad de todas las palabras del centro de interés”.

c) Bartol Hernández (2004b: 1090) ofrece una solución de compromiso en el *Léxico disponible de Soria*, basada en dos reglas:

– Reducir al sustantivo todas las construcciones en las que aparecen o el verbo *jugar* o el sustantivo *juego*<sup>27</sup> (*jugar a la consola > consola; juego/s de cartas*

<sup>27</sup> En este aspecto, Hernández Muñoz (2006: 296) sí sigue este principio, reduciendo al sustantivo la combinación *jugar a* y *juego de*, “siempre y cuando aparezca también el sustantivo de manera aislada y mantenga estrictamente el mismo significado (*juego de palmas > palmas/s, juego de las sillas > sillas, juego de muñecas > muñeca/s, jugar a la diana > diana, jugar a la pelota > pelota, jugar videojuegos > videojuego/s*, etc.”

> *cartas*), excepto los genéricos *juegos de lucha*, *juegos de mesa* y *juegos de confianza* (que además no tienen sus correspondientes soluciones *lucha*, *mesa*, *confianza*) y *jugar a algo*.

– En el resto de los casos, unificar las respuestas por medio del paréntesis siempre y cuando la variación solo presente dos opciones: (*contar chistes-chistes* > (*contar*) *chistes*; *hacer footing-footing* > (*hacer*) *footing*; *ir al parque de atracciones-parque de atracciones* > (*ir al*) *parque de atracciones*. Ello hace que mantenga entradas diferentes para *ir al cine*, *ver cine*, *cine*; *ir al fútbol*, *jugar al fútbol*, *fútbol*; *ir en bici*, *montar en bicicleta*, *bicicleta*, al ser imposible la unificación en una única entrada.

Por lo que se refiere a los campos léxicos que conforman el centro de interés ‘Juegos y distracciones’, este reúne palabras de ámbitos muy diversos. Los sustantivos son de nuevo la categoría lingüística mayoritaria y cubren temas como la cultura (*arte*, *biblioteca*, *cine*, *concierto*, *danza*, *lectura*, *libro*, *museo*, *música*, *novela*, *película*, *ópera*, *teatro...*), el deporte (*atletismo*, *baloncesto*, *balonmano*, *ciclismo*, *esquí*, *fútbol*, *hockey*, *natación*, *rugby*, *tenis*, *voleibol...*), los juegos infantiles y de naipes (*brisca*, *comba*, *corro*, *escoba*, *escondite*, *muñeca*, *videojuego...*), los juegos de mesa (*ajedrez*, *damas*, *dominó*, *oca*, *parchís*, *tres en raya*, *trivial...*), así como menciones específicas de utensilios que son necesarios para determinados juegos o deportes (*aro*, *baraja*, *bicicleta*, *bolas*, *canicas*, *cartas*, *dados*, *goma*, *patín*, *pelota*, *portería...*), la comida y la bebida (*bebida*, *café*, *cerveza*, *cena*, *chocolate*, *churro*, *helado*, *sangría*, *té...*), los medios de comunicación (*diario*, *internet*, *periódico*, *radio*, *revista*, *televisión*, *TV...*), las relaciones sociales (*amigo*, *baile*, *botellón*, *copa*, *discoteca*, *familia*, *fiesta*, *hobby*, *marcha*, *sexo...*), los lugares de ocio (*bar*, *cafetería*, *casino*, *discoteca*, *pub*, *restaurante...*), las vacaciones (*campo*, *camping*, *excursión*, *foto*, *mar*, *playa*, *piscina*, *siesta*, *sol*, *turismo*, *viaje...*), la música (*disco*, *flauta*, *guitarra*, *instrumento*, *música*, *piano*, *trompeta...*), las aficiones (*baile*, *buceo*, *bricolaje*, *paseo...*) y el estudio y el trabajo (*ejercicio*, *deberes*, *ordenador*, *PC*, *tarea...*).

Como adelantamos anteriormente, los verbos son más frecuentes en este centro que en los anteriores. Los temas a los que se refieren son acciones habituales (*andar*, *beber*, *charlar*, *cocinar*, *comer*, *conducir*, *dormir*, *escribir*, *fumar*, *hablar*, *reír...*), acciones relacionadas con el deporte (*correr*, *entrenar*, *escalar*, *esquiar*, *nadar*, *saltar...*), con el ocio (*bailar*, *cantar*, *dibujar*, *ir de compras*, *jugar*, *montar a caballo*, *pasear*, *pescar*, *salir*, *tomar el sol...*), con las relaciones sociales (*hablar por teléfono*, *hacer el amor*, *quedar*, *salir con amigos*, *salir de marcha*, *visitar...*) y con el estudio y el trabajo (*aprender*, *estudiar*, *pensar*, *trabajar...*).

En cuanto a las marcas comerciales, estas aparecen con un índice de disponibilidad léxica elevado. Estas marcas pertenecen a juegos de mesa (*hundir la flota*, *monopoly*, *scatergeries*, *tabú*, *trivial...*), diversos juguetes y muñecas (*barbie*, *lego*, *nancy*, *playmobil*, *scalextric...*) y juegos informáticos (*gameboy*, *nintendo*,

*PC fútbol, playstation, tetris...*). Figuran, además, numerosos anglicismos asimilados por el español, como *basketball, footing, fútbol, ping-pong, rugby, voleibol, footing, windsurf*, etc., el híbrido *puenting*, palabras en proceso de lexicalización (*bicicross, email, hulahop, jogging, pádel, rap...*) y otros sin lexicalizar (*full contact, motorbike, rafting, skate-board, softball, training...*).

Otra peculiaridad de este campo es la aparición de elementos que se lematizan en su forma plural. Su valor, según Gómez Devís (2004: 94) es el de “globalizador de una actividad habitual y por tanto reiterada”, aludiendo generalmente a diversos juegos, como *canicas, cartas, casitas, chapas, chinos, dados, médicos, muñecas*, etc.

**Cuadro 7.18. Comparación del léxico más disponible en el centro de interés ‘Juegos y distracciones’**

Nº	Asturias	Cast. La Mancha	Sevilla	Valencia	Zamora
1	2	3	4	5	6
1	fútbol	fútbol	parchís	fútbol	parchís
2	parchís	parchís	fútbol	baloncesto	fútbol
3	baloncesto	baloncesto	oca	tenis	baloncesto
4	tele(visión)	escondite	cartas	parchís	ajedrez
5	tenis	cartas	baloncesto	balonmano	cartas
6	ajedrez	juego de la oca	leer	natación	oca
7	escondite	tenis	ajedrez	voleibol	tute
8	oca	comba	escondite	cartas	mus
9	cartas	ajedrez	trivial	ajedrez	damas
10	trivial	trivial	tenis	correr	trivial
11	tute	ordenador	ordenador	atletismo	balonmano
12	damas	voleibol	comba	tele(visión)	tenis
13	voleibol	bailar	tele(visión)	escondite	brisca
14	música	monopoly	voleibol	damas	tele(visión)
15	comba	damas	monopoly	leer	escondite
16	leer	balonmano	scatergeries	oca	bomba
17	natación	mus	video juego	dominó	voleibol
18	balonmano	videoconsola	cine	frontón	cine
19	póquer	discoteca	tabú	bailar	póquer
20	mus	natación	elástico	música	leer
21	brisca	muñeca	dominó	cine	chinchón
22	monopoly	tute	correr	trivial	música
23	ordenador	goma	ver tele(visión)	rugby	natación
24	bailar	tres en raya	bailar	ordenador	bailar

1	2	3	4	5	6
25	tres en raya	cine	coger	pillar	dominó
26	cine	dominó	escuchar música	pasear	videojuego
27	videojuego	videojuego	balonmano	pimpón	discoteca
28	goma	canicas	damas	nadar	pasear
29	correr	pillado	música	ciclismo	fútbol
30	escoba	cantar	play station	póquer	monopoly
31	ciclismo	tabú	cantar	hockey	cantar
32	scatergeries	ping pong	tres en raya	béisbol	tres en raya
33	consola	televisión	dados	esquí	teatro
34	billar	correr	(video)consola	discoteca	correr
35	cascayu	scattergeries	pasear	squash	escoba
36	chinchón	música	salir	cantar	goma
37	atletismo	consola	natación	golf	dados
38	pasear	salir	internet	monopoly	lectura
39	golf	póquer	teje	saltar	bingo
40	nadar	cinquillo	escribir	dados	julepe
41	béisbol	corro de la patata	póquer	karate	gilé
42	canicas	fiesta	billar	<i>truc</i>	canicas
43	dominó	béisbol	deporte	tres en raya	saltar
44	rugby	brisca	muñecas	patinar	cinquillo
45	bici(cleta)	internet	discoteca	videojuegos	ciclismo
46	hockey	leer	juego (de mesa)	bingo	tetris
47	queda	botellón	bici(cleta)	consola	béisbol
48	radio	coche	películas	futbito	atletismo
49	duro	fútbol	teto	yudo	ver la tele(visión)
50	ping pong	petanca	radio	rol	futbito

La norma léxica común a los cinco léxicos disponibles comparados es la siguiente:

- Vocablos comunes en las 10 primeras posiciones: (5) *ajedrez, baloncesto, cartas, fútbol, parchís*.
- Otros vocablos comunes en la franja de disponibilidad +0,1: (5) *escondite, oca, tenis, trivial, voleibol*.
- Otros vocablos comunes entre las 50 primeras posiciones: (14) *bailar, balonmano, cine, correr, damas, dominó, leer, monopoly, música, natación, póquer, tele(visión), tres en raya, videojuego*.

– Vocablos comunes a 4 comunidades: (6) *béisbol, cantar, (video)consola, discoteca, ordenador, pasear*.

– Vocablos comunes a 3 comunidades: (10) *brisca, canicas, ciclismo, comba, dados, goma, mus, ping pong, scatergeries, tute*.

### 7.2.16. ‘Profesiones y oficios’

El centro de interés ‘Profesiones y oficios’ (Cuadro 7.19) posee un carácter muy concreto y compacto. Como indica Sánchez-Saus (2011: 424), este centro está compuesto mayoritariamente por sustantivos derivados mediante los prefijos *-ero, -tor/-dor/-or, -ista, -logo* y *-al*. El primero de ellos, *-ero*, tiene el significado de ‘oficio, ocupación, profesión o cargo’ y está presente en *banquero, barrenadero, bombero, camarero, camionero, carnicero, carpintero, cartero, cocinero, enfermero, fontanero, ganadero, ingeniero, jardinero, minero, obrero, panadero, peluquero, pescadero, tendero, zapatero*, etc. Las variantes del sufijo ‘agente’ *-tor/-dor/-or* forman los nombres *actor, agricultor, conductor, constructor, director, diseñador, entrenador, escritor, escultor, jugador, lector, limpiador, pescador, pintor, profesor, trabajador, traductor, vendedor*, etc. Por su parte, *-ista*, ‘persona que tiene determinada ocupación, profesión u oficio’, es el morfema derivativo de *artista, dentista, deportista, economista, electricista, futbolista, oficinista, periodista, recepcionista, taxista*, etc. Asimismo son numerosos los participios activos en *-nte*, ‘que ejecuta la acción expresada por la base’, sustantivados y lexicalizados como *ayudante, cantante, comerciante, delineante, dibujante, docente, estudiante, fabricante, presidente, representante, vigilante*, etc.<sup>28</sup> El morfema *-logo*, que tiene el significado de ‘persona versada o especialista’, forma nombres de profesiones relacionadas con distintas ciencias: *arqueólogo, biólogo, filólogo, ginecólogo, oftalmólogo psicólogo*, etc.

Por último, con los morfemas *-al* e *-ino*, ‘pertenencia o relación’, se forman adjetivos que, sustantivados, nombran asimismo profesiones: *comercial, concejal, fiscal, general, industrial, oficial, profesional, y campesino* (combinado con el interfijo *-es-*) (Sánchez-Saus, 2011: 424).

Aparecen, además, otros nombres de profesiones, trabajos y oficios, no derivados mediante estos morfemas, tales como *abogado, albañil, ama de casa, arquitecto, bailarín, empleado, empresario, farmacéutico, guardia civil, informático, juez, maestro, mecánico, médico, militar, piloto, policía, político, secretario, soldado, técnico, veterinario*, etc.

<sup>28</sup> Otros nombres de oficios aparecen con la misma terminación, aunque su significado, sin embargo, ha sufrido un cambio: *dependiente* (‘empleado que tiene a su cargo atender a los clientes en tiendas’), *comandante* (‘jefe militar de categoría comprendida entre las de capitán y teniente coronel’) y *practicante* (‘persona que posee título para el ejercicio de la cirugía menor’) (Sánchez-Saus, 2011: 424).

Además de las profesiones y oficios tradicionales y reconocidos, este centro de interés incluye actividades que pueden reconocerse como profesiones de hecho (*camello, canguro, chapero, chulo, famoso, matón, mendigo, parado, playboy, top model...*), nombres de profesionales del deporte (*atleta, boxeador, entrenador, futbolista, tenista...*) e incluso de algunos profesionales de prácticas delictivas (*asesino, carterista, proxeneta, terrorista, timador, traficante...*). Del mismo modo, se incluyen los nombres de actividades profesionales y de las carreras universitarias y estudios en general que desembocan en una profesión: *enfermería, estadística, hostelería, informática, ingeniería, magisterio, odontología, peluquería, periodismo, psicología* (cf. Samper Padilla, 1998a: 332).

En la neutralización de las variantes flexivas se conservan aquellas formas que presentan un sufijo específico (*actor/actriz*) y cuando el *DRAE* registra entradas diferentes para cada género (*modista* y *modisto*, *costurera* y *costurero*) y, por supuesto, en el caso en que designen competencias distintas (*asistent/a*/*asistente*; *monja/monje*; *sirvient/a*/*sirviente*). Y se mantienen en femenino los vocablos que el *DRAE* proporciona solo en femenino, como es el caso de *ama de casa, ama de llaves, azafata, chacha, costurera, doncella, chacha, empleada de hogar, matrona, niñera, puta*.

**Cuadro 7.19. Comparación del léxico más disponible en el centro de interés ‘Profesiones y oficios’**

Nº	Asturias	Cast. La Mancha	Sevilla	Valencia	Zamora
1	2	3	4	5	6
1	profesor	profesor	profesor	profesor	profesor
2	médico	albañil	médico	médico	médico
3	abogado	médico	abogado	abogado	abogado
4	albañil	abogado	albañil	arquitecto	ingeniero
5	camarero	agricultor	maestro	ingeniero	arquitecto
6	carpintero	enfermero	arquitecto	albañil	albañil
7	enfermero	camarero	agricultor	carpintero	maestro
8	fontanero	ingeniero	empresario	fontanero	enfermero/a
9	ingeniero	maestro	ama de casa	maestro	agricultor
10	policía	arquitecto	electricista	agricultor	carpintero
11	electricista	policía	ingeniero	electricista	barrendero
12	arquitecto	carpintero	carpintero	policía	empresario
13	empresario	empresario	camarero	pintor	policía
14	agricultor	electricista	fontanero	enfermero	fontanero
15	barrendero	fontanero	enfermero	barrendero	periodista
16	maestro	ama de casa	policía	empresario	camarero

Cuadro 7.19 (cont.)

1	2	3	4	5	6
17	ganadero	ganadero	psicólogo	panadero	electricista
18	pintor	secretario	panadero	camarero	estudiante
19	mecánico	bombero	pintor	secretario	economista
20	secretario	banquero	periodista	jardinero	pintor
21	camionero	mecánico	carnicero	vendedor	físico
22	estudiante	jardinero	cocinero	bombero	juez
23	ama de casa	camionero	jardinero	psicólogo	banquero
24	panadero	pintor	dependiente	basurero	dependiente/a
25	jardinero	panadero	secretario	ama de casa	químico
26	conductor	futbolista	peluquero	banquero	tendero
27	juez	psicólogo	biólogo	juez	ganadero
28	bombero	tendero	informático	dependiente	carnicero/a
29	cocinero	dependiente	administrativo	mecánico	secretario/a
30	taxista	cocinero	mecánico	camionero	bombero
31	banquero	periodista	juez	periodista	mecánico
32	futbolista	barrendero	ganadero	carnicero	cartero
33	psicólogo	carnicero	barrendero	conductor	conductor
34	periodista	juez	taxista	economista	oficinista
35	carnicero	informático	vendedor	taxista	farmacéutico/a
36	minero	cantante	futbolista	farmacéutico	jardinero
37	tendero	taxista	bombero	futbolista	panadero
38	economista	conductor	camionero	escritor	camionero
39	pescador	vendedor	limpiadora	administrativo	aparejador
40	peluquero	prostituta	director	actor	peluquero/a
41	dependiente	peluquero	farmacéutico	cartero	dentista
42	vendedor	administrativo	actor	cocinero	vendedor
43	director	estudiante	cantante	director	taxista
44	biólogo	actor	frutero	ATS	matemático
45	actor	economista	dentista	escultor	psicólogo
46	piloto	pescadero	químico	dentista	delineante
47	veterinario	dentista	economista	político	militar
48	pescadero	comerciante	basurero	zapatero	director
49	obrero	biólogo	veterinario	prostituta	cantante
50	guardia civil	veterinario	ATS	piloto	actor

La norma léxica común a los cinco léxicos disponibles comparados queda como sigue:

– Vocablos comunes en las 10 primeras posiciones: (4) *abogado, albañil, médico, profesor*.

– Otros vocablos comunes en la franja de disponibilidad +0,1: (13) *agricultor, arquitecto, camarero, carpintero, electricista, empresario, enfermero, fontanero, ingeniero, maestro, periodista, pintor, policía*.

– Otros vocablos comunes entre las 50 primeras posiciones: (15) *actor, barrero, bombero, camionero, carnicero, dependiente, economista, jardinero, juez, mecánico, panadero, psicólogo, secretario, taxista, vendedor*.

– Vocablos comunes a 4 comunidades: (9) *ama de casa, banquero, cocinero, conductor, dentista, director, futbolista, ganadero, peluquero*.

– Vocablos comunes a 3 comunidades: (6) *biólogo, cantante, estudiante, farmacéutico, tendero, veterinario*.



## CONCLUSIONES

Las páginas de este libro han servido para mostrar los logros y las posibilidades de la investigación en el campo de la disponibilidad léxica en español como lengua materna<sup>1</sup>. El estudio de la disponibilidad léxica, es decir, de las palabras que vienen con prontitud a la mente como una respuesta a un estímulo lingüístico asociado a los dominios léxicos de uso corriente en la vida diaria ('El cuerpo humano', 'Los alimentos', 'La casa', etc.) nos permite –ahora sí– “contar con una fotografía muy precisa de una parcela léxica tan relevante para conocer muchos aspectos del español general y de las diferenciaciones dialectales en una capa de la población hasta ahora poco estudiada” (Samper Padilla *et al.*, 2003: 123). En tan solo once años –los que van de 2003 a 2014–, este objetivo que planteaban los estudiosos del equipo canario se ha hecho realidad, gracias al amplio número de estudios realizados y a la diversidad de los temas abordados. De ello da testimonio la abundante bibliografía consultada que acompaña este estudio, que es, a su vez, no obstante, solo una parte del total<sup>2</sup>.

La disponibilidad léxica nació ligada a la didáctica de la lengua –en concreto, de la lengua francesa como segunda lengua, con motivo de los trabajos preparatorios para la confección de *Le Française Élémentaire* (Gougenheim *et al.*, 1954)–, y arrancó en el caso de la lengua española también en conexión con la enseñanza de la lengua materna –López Morales (1973, 1978) dedica sus primeros trabajos en este campo a la planificación de la adquisición del vocabulario y la evaluación del desarrollo de la competencia léxica. No ha sido, sin embargo, el ámbito didáctico el objeto de estudio de esta monografía, si bien queremos dejar aquí constancia de los beneficios de su aplicación didáctica. Si se usaran los resultados de los estudios de disponibilidad léxica en la enseñanza del español, se podría garantizar que las personas al final de su etapa escolar dispusieran de las palabras que “tienen mayor rentabilidad en el uso cotidiano” en “las situaciones comunicativas de la vida moderna” (López Morales, 1978: 76–77).

---

<sup>1</sup> Para un análisis de los hallazgos y reflexiones de diversos estudios de léxico disponible en español e inglés como lengua extranjera, consúltese en Jiménez Catalán (2014).

<sup>2</sup> Las dificultades de acceso a algunos artículos y léxicos publicados han ocasionado, muy a nuestro pesar, la ausencia de sus hallazgos y reflexiones en estas páginas.

Este trabajo ha pretendido analizar los fundamentos teóricos y metodológicos de los estudios de disponibilidad léxica, así como lo variado y fecundo de sus aportaciones dentro del campo de la ciencia lingüística. Como hemos visto, el Proyecto Panhispánico de Disponibilidad Léxica constituye un sólido marco de trabajo que garantiza unos criterios metodológicos comunes a todos los grupos de investigación que a él se acogen. La coherencia metodológica se consigue a través del uso de una misma prueba escrita asociativa para la recogida de los datos con los mismos dieciséis centros de interés clásicos, el mismo tiempo de respuesta para cada categoría temática –dos minutos–, los mismos criterios de edición de Samper Padilla (1998), las mismas variables sociolingüísticas, etc.

El orden de las respuestas en la hoja de la encuesta es un factor importantísimo en la investigación en disponibilidad léxica, ya que es un indicio decisivo sobre el lugar que ocupa la palabra en la organización mental. Para ello se ha desarrollado el índice de disponibilidad léxica, que combina la posición que cada palabra alcanza en las listas con su frecuencia de aparición. La fórmula de López Chávez y Strassburguer Frías (1991) es el algoritmo fiable que evalúa la probabilidad de que una palabra venga a la mente cuando una conversación trata sobre un tema específico. Los diferentes cálculos estadísticos, indispensables para un estudio de estas características, se han visto facilitados con el desarrollo de programas como *Lexidisp* (Moreno *et al.*, 1995), y más recientemente con las herramientas informáticas de la página Web *Dispolex.com*, creada por la Universidad de Salamanca (cf. Bartol Hernández y Hernández Muñoz, 2003). La léxico-estadística, en fin, ha sido una disciplina que ha servido al tratamiento de las listas de palabras, y que ha aportado oposiciones terminológicas tan importantes como palabras temáticas *vs.* palabras no temáticas, léxico frecuente *vs.* léxico disponible, palabra *vs.* vocablo, de las que aquí hemos tratado.

Entre los cálculos estadísticos también se cuenta el índice de cohesión, propuesto por Max Echeverría (1991), el cual hace referencia al grado de coincidencia entre las respuestas de los informantes, y que muestra cómo de abierto (difuso) o cerrado (compacto) es un centro de interés. Con independencia de la comunidad estudiada, la calidad de compacto o difuso de cada centro suele coincidir. Así, por ejemplo, las categorías ‘Partes del cuerpo humano’ y ‘La ropa’ suelen poseer un alto porcentaje de vocablos compartido por los informantes. En cambio, las categorías ‘Trabajos del campo y del jardín’ y ‘El campo’ presentan una gran diversidad en las respuestas de los informantes, por lo que se convierten en centros abiertos.

Evidencias como esta, además de la diferente productividad y composición léxicas de los centros de interés del Proyecto Panhispánico han llevado a Hernández Muñoz (2006: 134) a afirmar la disparidad semántica de los centros de interés clásicos a partir de su diferente ontogenia, organización interna y alcance. Entre los 16 centros de interés fijados por Gougenheim *et al.* (1964) se distinguen fácil-

mente subgrupos caracterizados por una diferente cantidad de vocabulario y por el grado de dispersión de sus componentes. Si los estudios de disponibilidad léxica, sobre todo los pertenecientes al citado macroestudio usan estos centros, es sobre todo por la conveniencia de realizar comparaciones entre las diferentes sintopías.

A raíz de estas diferencias entre categorías semánticas, Hernández Muñoz (2006: 150–156) propone un esbozo de una taxonomía de los centros de interés, que posteriormente Sánchez-Saus (2011: 196–209) completa y perfecciona, en función de la estructura interna y del nivel de inclusividad de los centros. Se descubre así la fuerte relación existente entre los conceptos de disponibilidad y prototipicidad. La teoría de los prototipos es válida para explicar las relaciones en los centros que responden a las características clásicas de una categoría semántica (límites más o menos definidos relacionados con características físicas o abstractas que nacen de la experiencia del mundo, identificación sencilla de los elementos que la constituyen, etc.). La mayoría de los centros de interés clásicos pertenecen a este tipo de categoría: ‘La ropa’, ‘Partes de la casa’, ‘Los muebles de la casa’, ‘Comidas y bebidas’, ‘La cocina y sus utensilios’, ‘Calefacción e iluminación’, ‘Medios de transporte’, ‘Trabajos del campo y del jardín’, ‘Animales’ y ‘Profesiones y oficios’. Sin embargo, la organización de las unidades léxicas es diferente en el resto de centros de interés, por lo que, para Hernández Muñoz (2006: 155) “la hipótesis de que los términos disponibles son aquellos más cercanos al prototipo (o a los prototipos) de una categoría semántica es cuestionable, ya que no todos los centros de interés poseen una estructura prototípica identificable”.

La tesis doctoral de Hernández Muñoz (2006) constituye un hito muy importante en los estudios de léxico disponible, al dar respuesta a la petición de López Morales (1999: 26), quien reclamaba no solo una teoría psicológica de la disponibilidad léxica, que conjugara factores como memoria semántica, familiaridad, etc., sino toda una teoría integrada de la disponibilidad. Después de Hernández Muñoz (2006), la psicolingüística no es solo un ámbito de aplicación, sino que pasa a convertirse en uno de los fundamentos de este tipo de estudios. Ello va a tener consecuencias, además, en una fase crucial de la investigación: el proceso de edición del corpus. Las normas de edición de Samper Padilla (1998) han pasado por el tamiz de los reenfoques habidos en estas investigaciones (pedagógico, descriptivo, psicológico). Sucesivamente, el grado de intervención y filtro del investigador se ha ido flexibilizando, hasta imponerse un necesario compromiso entre el objetivo de descripción de la norma léxica culta de una comunidad y el estudio de las implicaciones psicolingüísticas del léxico disponible. Ello implica que las tareas de estandarización y lematización no deben ir nunca en detrimento de la fidelidad al corpus, ni se puede descartar información léxica válida bajo una perspectiva psicolingüística, e incluso etnolingüística.

Abundando en los aspectos de procedimiento, la disponibilidad léxica no se ha mantenido ajena al debate metodológico. De hecho, ya Borrego Nieto (2004)

planteó una serie de cuestiones que movían a duda y controversia. Estas cuestiones afectaban a:

1) la influencia de la encuesta en el léxico producido debido a factores tales como las condiciones físicas de realización de la prueba, las pequeñas variaciones en el enunciado de los centros de interés, las pautas que el encuestador da a los informantes, el nivel de formalidad que implica realizar la encuesta en un centro académico, etc.;

2) la comparación cuantitativa de los datos entre diferentes repertorios léxicos, que se ve influida por el proceso de edición de los materiales, el diferente tamaño de las muestras, la compatibilidad entre sintopías; y a la comparación cualitativa, que depende del criterio que adoptemos a la hora de clasificar las palabras;

3) el tipo de informante –jóvenes preuniversitarios en el Proyecto Panhispánico–, y la posibilidad de ampliarlo a otro tipo de sujetos (según la edad, la extracción social, dialectal, etc.);

4) los centros de interés, en relación con su selección, cantidad, definición y el tipo de vocabulario que se obtiene o se persigue obtener;

5) el estatuto mental del léxico disponible, en cuanto a su relación con el lexicón mental y con otras realidades psicológicas o psicolingüísticas como el orden de aprendizaje, la prototipicidad, la familiaridad y el carácter abstracto o concreto de los términos.

Precisamente esas dudas han venido a ser resueltas dentro del Proyecto Panhispánico, por una parte, con unas pautas comunes, fruto de las reuniones de coordinación y de la red de investigación subsiguientes, que permiten “tanto las comparaciones directas entre los datos de distintas comunidades de habla como la obtención de unos resultados generales que facilitarán, primero, la elaboración del léxico disponible de España y, posteriormente, del léxico disponible panhispánico” (Samper Padilla y Samper Hernández, 2006: 84). De otra parte, se han explorado y, en su caso, afianzado nuevos enfoques metodológicos que enriquecen estas investigaciones con nuevas propuestas, como son las de tipo cognitivo que ya hemos comentado.

Probablemente la discusión de mayor interés es la relacionada con las características de las áreas temáticas elegidas. Como se ha dicho, las ventajas de manejar un conjunto común de áreas temáticas ha favorecido el desarrollo de esta línea de investigación en el ámbito hispano, gracias a las posibilidades de uso de los datos, las comparaciones y los análisis desde diversas perspectivas. Sin embargo, parece necesaria, de una parte, la incorporación de nuevos núcleos temáticos que permitan registrar, además de sustantivos de carácter concreto, otro tipo de vocabulario que incluya otras tipo de categorías verbales. En esta línea van las propuestas de nuevos centros de interés para recoger léxico abstracto<sup>3</sup> de

---

<sup>3</sup> También habían recabado vocabulario abstracto con diferentes centros de interés en México, Cañizal Arévalo (1987) y, en Chile, Vargas Sandoval (1991) y Valencia y Echeverría (1998 y 1999).

Hernández Muñoz (2004) con ‘La inteligencia’, Prado Aragonés *et al.* (2006) con ‘La salud’ y Paredes García *et al.* (en prensa) con ‘Sensaciones y sentimientos’, y las que recaban verbos con ‘Acciones y actividades habituales’, y adjetivos con ‘Aspectos físicos y de carácter’ de Sánchez-Saus (2011), etc. No obstante, como venimos diciendo, cualquier variación en la nómina de los centros de interés hace imposible la comparación interdialectal.

Por lo que respecta al tamaño y composición de las muestras, dado el carácter sociolingüístico del Proyecto Panhispánico, las investigaciones de disponibilidad léxica suelen incluir un capítulo en el que primero se delimita la zona de estudio, dando los datos pertinentes que ayuden a explicar la complejidad social de la región estudiada. De esta manera, las características geográficas, económicas y demográficas se entrelazan entre sí en la conformación de la red social de una sintopía. Estos tres aspectos definen así el universo estudiado y, en consecuencia, la composición de la muestra. Por este motivo, los estudiosos se esfuerzan en cumplir con unos requisitos estadísticos objetivos en la selección de la muestra del léxico disponible, siguiendo para ello unos criterios apropiados de exhaustividad y de representatividad proporcional coincidentes con los de otros trabajos similares. El tamaño recomendado de las muestras (400 informantes) cumple plenamente con el criterio de representatividad expresado por Labov (1972) de una relación de 0,025% en los estudios sociolingüísticos. No obstante, como hemos visto, el tamaño de la muestra ha variado de unos investigadores a otros y, si bien no afecta a la comparación de la productividad de palabras por informante, sí parece difícil establecer una comparación precisa en relación con el número de vocablos, ya que esta cifra se ve directamente afectada por el tamaño de las muestras.

Otra contribución novedosa tiene que ver con la elección de sujetos de otros grupos sociales, lo que puede hacer necesario la realización oral de la encuesta. Esto ya fue puesto en práctica por Sánchez Corrales y Murillo Rojas (1993) en su investigación con niños de nivel preescolar, y, recientemente lo ha hecho Borrego Nieto (2008) en su estudio con jubilados de edad avanzada y de bajo nivel educativo.

Junto al desarrollo del enfoque psicolingüístico, hay que destacar la atención que, desde el principio, el Proyecto Panhispánico ha dedicado a la incidencia de los factores sociales, con el establecimiento de cuatro variables sociales de obligado estudio en cada investigación: ‘sexo’, ‘tipo de centro escolar’, ‘zona geográfica’ y ‘nivel sociocultural’, en la línea de la sociolingüística variacionista. Como afirma Samper Padilla (2006: 117), “el análisis de los condicionantes sociales ha puesto de manifiesto que el factor que suele explicar la variación comunitaria de forma más acusada es el nivel sociocultural”. En este sentido, los análisis dejan ver regularmente una menor disponibilidad en los hablantes de los niveles socioculturales bajos, coincidiendo en cierto modo con algunas de las conclusiones de la teoría del déficit de Bernstein. Ligado al nivel sociocultural, también muestra un

papel muy relevante en algunas zonas el tipo de centro escolar (público o privado), con mejores porcentajes de disponibilidad en los alumnos de centros privados.

En cuanto al factor ‘sexo’, Samper Padilla (2006: 117) se sorprende de que aún hoy en día el resultado de este factor ofrezca “en muchas sintopías unas divergencias léxicas que reflejan la separación de actividades e intereses entre hombres y mujeres”, lo que, en sus palabras “puede interpretarse como la pervivencia residual de una diferencias tradicionales que van perdiéndose”. Lo cierto es que, como demuestra el Cuadro 5.1, en general, los informantes femeninos superan a los masculinos en la mayoría de los centros de interés –excepto en aquellos ámbitos culturalmente orientados al sexo masculino (‘Calefacción, iluminación y medios de airear un recinto’, ‘Medios de transporte’, ‘Trabajos del campo y del jardín’)–, lo cual parece indicar que la variable ‘sexo’ está también relacionada con la productividad léxica. Estas divergencias, cualitativamente, afectan también el tipo de léxico actualizado. Así, es sintomático que en Sevilla, Trigo Ibáñez (2011: 208) observe una tendencia a que las mujeres ofrezcan listados léxicos más estandarizados que los hombres, puesto que, aunque las informantes femeninas consignen palabras de uso coloquial o familiar, eufemismos y disfemismos, lo hacen con menor índice de disponibilidad léxica que estos.

Todavía dentro del enfoque sociolingüístico, son de singular importancia en el marco del estado español las investigaciones que estudian las comunidades bilingües del País Vasco, Navarra, Valencia, Galicia y Cataluña, además de las ciudades autónomas de Ceuta y Melilla, en las que se dan situaciones de diglosia con las lenguas árabe y bereber. Además del estudio obligatorio de la incidencia de los factores ‘lengua materna’, ‘lengua de uso familiar’ y ‘modelo de enseñanza (monolingüe/bilingüe)’, el análisis de las transferencias léxicas en las situaciones de contacto lingüístico aporta información sobre vitalidad de las mismas –más frecuentes en los hablantes bilingües–, así como de las áreas temáticas que facilitan el trasvase lingüístico. Por ejemplo, en Navarra, las áreas que facilitan el contacto lingüístico son, en su mayoría, las relacionadas con el ámbito doméstico y cotidiano: ‘La ropa’, ‘Partes de la casa’, ‘La comida y bebida’, ‘La escuela’, ‘El campo’, ‘Los trabajos de campo y jardín’ y ‘Los juegos y distracciones’.

Sin duda, como afirma López Morales (2008: 10), “el análisis cuantitativo [y cualitativo] de la disponibilidad de una comunidad dialectal dada ofrece ya una buena descripción de esta parcela de su norma léxica [...]. Sin embargo, donde estos estudios ofrecen su mejor y más valiosa contribución es en las comparaciones interdialectales”.

El análisis cuantitativo, apoyado en la léxico-estadística, permite establecer para cada comunidad listas ordenadas de las unidades léxicas según los distintos grados de disponibilidad. Las diferentes tablas estadísticas distribuyen jerárquicamente las áreas temáticas encuestadas, en un caso teniendo en cuenta el volumen de respuestas que se obtiene para cada una de ellas, en otro según la productividad

en palabras diferentes (vocablos), mostrando comportamientos divergentes en un caso y otro, como se aprecia en los Cuadros comparativos 3.13 y 3.14. Este diferente comportamiento en la producción de palabras por informante y palabras diferentes en el total de la muestra es el que cuantifica el índice de cohesión, que nos dice en qué medida convergen las respuestas de los informantes en los distintos centros de interés. Por fin, el análisis de todos estos índices en las variables que contemplan las investigaciones permite extraer conclusiones de orden sociolingüístico, que llevan al establecimiento de estratificaciones de niveles de habla a partir de los resultados de las listas de disponibilidad.

Con ser la estadística importante –e imprescindible– para obtener conclusiones significativas y fiables dentro de la investigación lingüística, hay que recordar con Guiraud (1959) que la estadística no es más que un procedimiento que nos acerca al conocimiento de nuestro objeto de estudio, la realidad lingüística léxica. El análisis cualitativo de la naturaleza y organización jerárquica del léxico actualizado por los distintos sociolectos dentro de cada variable social y del grado de compatibilidad y divergencia de las variables léxicas disponibles entre las distintas submuestras, dentro de un mismo factor social, permite delimitar las unidades léxicas que componen los recuentos y nos acerca al uso real que de esas unidades hace la comunidad hablante. El conjunto de asociaciones que despiertan los estímulos temáticos es muy variado, lo que lleva a que también sea muy variada la naturaleza del léxico actualizado. De ahí la importancia de mantener también en los listados las formas que aluden solo de manera indirecta a un centro de interés, ya que pueden ser utilizadas para estudios interdisciplinarios (psicolingüísticos, etnolingüísticos, etc.). Como afirma Mateo García (1998: 250), “organizarlas según criterios restrictivos habría supuesto desterrar los términos abstractos que, de este modo, si pueden incorporarse a los listados; términos que, además, en buena medida ofrecen valiosa información psicosocial de la que el elemento lingüístico es un vehículo muy directo”. Se refiere la investigadora a la aparición en el léxico disponible de Almería de términos como *grande* en el centro de interés ‘La ciudad’ o la inclusión de términos como *no atasco, sin contaminación* en el centro ‘El campo’.

En cualquier caso, estas unidades léxicas –u otras que presenten relaciones más alejadas con respecto a la etiqueta del estímulo temático y con respecto a los elementos prototípicos de una categoría dada– suelen presentar unos índices de disponibilidad léxica bajos, lo que hace que, en general, no se hallen dentro de la norma léxica de la comunidad. Por ello, su inclusión en los listados definitivos no entorpece el objetivo final de descripción de la norma léxica adulta de una comunidad dada. Como ya hemos reflejado, existe una cercanía entre los conceptos de disponibilidad y (proto)tipicidad, que puede venir dada por el valor del índice de disponibilidad. Pero además, como explica Bartol Hernández (2008), un valor concreto del índice de disponibilidad léxica (entre 0,1 y 0,01 en función del grado

de restricción que se quiera imponer) sirve como delimitación de las unidades que constituyen la norma léxica de una comunidad. Las palabras incluidas dentro de la norma léxica son las más representativas y con mayor frecuencia de uso dentro del campo de que se trate. Los vocablos que se hallan por debajo del punto de corte establecido y, por tanto, excluidas de esa norma léxica, suelen serlo por factores psicolingüísticos que afectan a pocos individuos o incluso a uno solo, como es el efecto de la memoria semántica, el *priming* o facilitación semántica, la familiaridad, etc., sociolingüísticos, como los derivados de los procesos de cambio lingüístico que conllevan la aparición o desaparición de determinadas palabras, y estilísticos, si se trata de términos muy marcados, bien por su grado de formalidad (tecnicismos) o de informalidad (eufemismos, vulgarismos). Al respecto, es de destacar la propuesta de análisis de Borrego Nieto (2009), que incorpora criterios estilísticos, ausentes en muchas de las investigaciones realizadas.

Por lo demás, un análisis lexicográfico que describa detalladamente las unidades de los diccionarios de disponibilidad parece necesario de cara a un uso posterior de esos repertorios léxicos, sea para tener un conocimiento exacto de la estructura y composición del léxico realmente utilizado, sea para un uso pedagógico de esos listados en la enseñanza de la lengua. La situación comunicativa de la encuesta –escrita– en un centro de enseñanza secundaria de un núcleo urbano, el grupo de individuos seleccionado para la prueba –estudiantes de nivel preuniversitario– la neutralización del factor ‘edad’ en el Proyecto Panhispánico, el mismo efecto uniformizador de los medios de comunicación y de la propia instrucción escolar determina, de manera clara, la pertenencia del léxico recogido a la variedad estándar de la comunidad de habla. Por eso, la mayoría del léxico disponible es léxico no marcado social ni estilísticamente, con predominio del léxico patrimonial, tanto entre las unidades simples como entre los grupos léxicos de carácter muy dispar. En este sentido, la descripción de su estructura y de los componentes que los caracterizan pone de manifiesto la manera en la que las palabras se asocian en el lexicón mental.

Entre el léxico disponible llama la atención el alto índice de disponibilidad de compuestos sintagmáticos, en su mayoría, a partir de un término nuclear evocado en primer lugar. Tales sintagmas constituyen, en muchos casos, auténticas colocaciones gramaticales y léxicas –combinaciones prefabricadas en la norma que presentan cierta fijación interna y que carecen de idiomatidad (cf. Castillo Carballo, 1998; Corpas Pastor, 1996)– tales como *montar a caballo*, *ver la televisión*, *estación de trenes*, *fiesta de disfraces*, etc. Como vimos con relación a ‘La ropa’ en § 6.3.2 y en otros centros en § 7.2, dichos compuestos son más frecuentes en su manifestación nominal, predominando el empleo de nexos para realizar las construcciones (*bota de agua*, *mesita de noche*, *arroz con leche*...). Las construcciones verbales se concentran principalmente en los centros de interés 13 ‘Trabajos del campo y del jardín’ –donde abundan las com-

plementaciones directas (*cortar el césped, recoger los frutos...*)– y 15 ‘Juegos y distracciones’ –con mayor presencia de construcciones con nexos (*montar en bicicleta, salir con amigos*).

También merece una atención especial la lexicalización de las marcas comerciales, muchas de ellas ya plenamente genéricas o en vías de cobrar ese carácter, debido a la influencia de los medios de comunicación en los informantes. A lo largo de § 7.2 damos cuenta de su importancia en los diferentes centros de interés. La presencia de marcas lexicalizadas se concentra, sobre todo, en 05 ‘Comidas y bebidas’ (*Coca Cola, Cola cao, Danone, Donut, Fanta...*), 08 ‘La escuela’ (*Dyn A4, Fixo, Rottring, Tippex...*) y 15 ‘Juegos y distracciones’ (*Barbie, Lego, Monopoly, PlayStation...*).

Dentro del análisis del vocabulario recopilado, destacan los estudios sobre los anglicismos y dialectalismos. En el campo del contacto de lenguas, y en un mundo globalizado como el actual, el índice de disponibilidad de los extranjerismos cuantifica el nivel de influencia del inglés o de otra lengua extranjera en nuestra lengua y cultura españolas. Este ha sido un campo de estudio que ha despertado un gran interés entre los investigadores de la disponibilidad léxica. Los centros de interés más proclives a la inclusión de extranjerismos son: 02 ‘La ropa’ (*body, frac, jeans, maillot, panty, t-shirt...*), 10 ‘La ciudad’ (*bar, boutique, parking, pub, sex shop...*) y 12 ‘Medios de transporte’ (*autostop, ferry, footing, jet, mountain bike...*).

Por lo que se refiere a los dialectalismos, con una presencia residual en la mayoría de las investigaciones, las investigaciones dejan constancia del nivel de pervivencia de este léxico dialectal. En uno y otro caso, como indican Samper Padilla *et al.* (2003: 119), en las investigaciones de disponibilidad léxica “al tratarse de un léxico aportado por hablantes muy jóvenes, puede inferirse directamente el futuro inmediato de tales voces en las distintas comunidades”. Precisamente, a veces, la falta de un léxico típicamente autóctono es síntoma de una situación de mortandad lingüística, como la que en la provincia de Cádiz constata González Martínez (2002: 99), ante la ausencia en las listas de palabras del dialecto gaditano como *barrilete o pandorga* ‘cometa’, *piriñaca* ‘ensalada de tomate, pimienta y cebolla’, *migote* ‘sopas de pan mojadas en el café, leche, o cacao’.

En este punto, al referirnos a la presencia de dialectalismos, y dado que anteriormente hicimos mención a las comparaciones interdialectales, queremos expresar nuestro acuerdo con Martínez Olmos (2007: 53–58) acerca de la ambigüedad terminológica que conlleva el uso de “dialecto” o “sintopía”, entre otros términos, en la denominación del marco geográfico y social sobre el que se plantea una investigación sobre disponibilidad léxica. En su argumentación, Martínez Olmos (2007: 56) aduce dos razones para rechazar el término “dialecto”: por un lado, la sociolingüística, uno de los pilares teórico-metodológicos de los estudios de léxico disponible, ya posee una unidad mucho más apropiada y precisa que el dialecto

en dialectología y, por otro, según Gimeno *et al.* (en prensa), es imposible usar el nombre “dialecto” para designar el marco social y espacial a través del que se extiende la variable léxica disponible del sociolecto preuniversitario. Como anticipamos anteriormente, la situación comunicativa de la encuesta en un centro de enseñanza secundaria con estudiantes preuniversitarios determina una referencia clara a la variedad estándar de la comunidad de habla. Dado el carácter de representatividad de la norma léxica adulta por esta franja de edad (véase López Morales, 1999: 28), “es evidente que el término «dialecto» (hoy diríamos vernáculo) no sería nunca el apropiado para la denominación de la variedad social implicada (el estándar) en el sociolecto de los estudiantes preuniversitarios” (Martínez Olmos, 2007: 58). ¡Nótese el contrasentido que supone afirmar que en el léxico disponible de un dialecto aparecen pocos dialectalismos de ese dialecto!

Retomando el tema de las comparaciones entre los léxicos disponibles de diferentes comunidades, proceder metodológicamente de manera análoga a otros estudios semejantes en el marco del Proyecto Panhispánico permite la plena comparabilidad de los léxicos disponibles resultantes, y así obtener las unidades léxicas que presenten altos índices de disponibilidad y las más frecuentes entre los hispanohablantes. A ello hemos dedicado el capítulo VII, que presenta un esbozo del léxico más disponible –o norma léxica, en palabras de Samper Padilla (1999: 554)– del español peninsular, a partir de las 50 primeras unidades de los léxicos disponibles de Asturias, Castilla-La Mancha, Sevilla, Valencia y Zamora. Se acompaña la delimitación del léxico altamente disponible, compartido por dichas comunidades de habla, con el análisis descriptivo de las unidades léxicas que componen los centros de interés. Este apartado pretende corresponderse, de hecho, con el análisis cualitativo de una investigación sobre léxico disponible. Si en § 6.3.2 se describen, de manera exhaustiva, lexicográficamente las unidades léxicas que componen el centro de interés 02 ‘La ropa’ de la provincia de Córdoba, de la investigación de Bellón Fernández (2011), en § 7.2 se aprovecha el cotejo de cinco léxicos disponibles peninsulares para dar cuenta de la variedad temática dentro de los centros de interés y de los fenómenos lingüísticos más destacados. Es esta, también, una ocasión que aprovechamos para comentar someramente las pautas particulares de edición que siguen los investigadores, dada su repercusión en la composición final de los listados.

Al final de este capítulo de conclusiones quisiéramos volver nuestra atención a las palabras mismas, el objeto de estudio de la disponibilidad léxica. En toda investigación de disponibilidad, la parte más importante está constituida por los listados de vocablos que conforman el diccionario del léxico disponible de una comunidad dada. Ejemplo de estos listados es el listado íntegro del léxico disponible de ‘La ropa’ de Córdoba (Bellón Fernández, 2011), mostrado en § 6.2, y los listados parciales con las 50 primeras unidades de los léxicos disponibles de Asturias (Carcedo González, 2001), Castilla La Mancha (Hernández Muñoz, 2006),

Sevilla (Trigo Ibáñez, 2011), Valencia (Gómez Devís, 2004) y Zamora (Galloso Camacho, 1998a), con los que trabajamos en § 7.2. En estas, las palabras, en tanto que designan la realidad, se reflejan de manera directa e inmediata, por un lado, los patrones de comportamiento, las peculiaridades culturales y los modos de vivir de una comunidad, y por otro, “los cambios socioeconómicos y culturales ocurridos en la comunidad hablante, así como las necesidades cambiantes de intercambio comunicativo de informaciones, conocimientos enciclopédicos, opiniones y criterios” (Mateo García, 1998: 251). Por eso, el léxico disponible es objeto de estudio de la etnolingüística –aún no lo suficientemente aprovechado que pudiera ser– y de la sociolingüística, con la que mantiene un maridaje desde los comienzos de su estudio en el mundo hispánico. Desde el punto de vista historicista, estos cambios socioculturales también podrán ser estudiados a partir de la disponibilidad léxica, pues el cambio social tiene su correlato en el cambio lingüístico y en la evolución del léxico. que varía, además, con el paso de los años, como ya ha puesto de manifiesto Alba (2013), al contrastar los datos de dos encuestas similares realizadas en 1990 y 2008 en la República Dominicana<sup>4</sup>. Los listados léxicos, en fin, están formados por ítems léxicos, cada uno con una intrincada red de relaciones formales y semánticas entre ese ítem y otras palabras y morfemas, que constituyen las relaciones significativas que estudiamos en § 4.2.3, y que conforman, en definitiva, la competencia léxico-semántica<sup>5</sup> de cada individuo. Gracias a esta competencia léxico-semántica de las palabras, en combinación con otros dominios de la competencia lingüística, podemos expresarnos y comprender en español.

---

<sup>4</sup> En este estudio, Alba (2013) constata tanto la incorporación y desaparición como la variación en el grado de disponibilidad de distintas palabras en el léxico de los estudiantes dominicanos.

<sup>5</sup> “Capacidad para relacionar formas con significados y utilizarlas adecuadamente, esto es, comprender una palabra es un proceso mental que consiste en conocer su significado, su estructura y saber usarla” (Lahuerta y Pujol, 1996: 121).



## BIBLIOGRAFÍA

- Ahumada Lara, Ignacio (2006): *El léxico disponible de los estudiantes preuniversitarios de la provincia de Jaén*. Jaén: Universidad.
- Aitchison, Jean (1987): *Words in the Mind: An Introduction to the Mental Lexicon*. Oxford: Basil Blackwell, 1994.
- Alameda, José R. y Fernando Cuetos (1995): *Diccionario de frecuencias de las unidades lingüísticas del castellano*, 2 vols. Oviedo: Universidad.
- Alba Ovalle, Orlando (1995a): *Léxico disponible de la República Dominicana*. Santiago de los Caballeros: Pontificia Universidad Católica Madre y Maestra.
- Alba Ovalle, Orlando (1995b): “Disponibilidad léxica en el español dominicano: análisis cuantitativo”, en L. Rodríguez-Fonseca e I. N. Vázquez (eds.), *Actas del III Seminario Internacional sobre ‘Aportes de la lingüística a la enseñanza de la lengua materna’*, números especiales de la *Revista de Adquisición de la Lengua Española*.
- Alba Ovalle, Orlando (1995c): “Anglicismos léxicos en el español dominicano: análisis cuantitativo”, en *El español dominicano dentro del contexto americano*. Santo Domingo: Librería La Trinitaria, 11–38.
- Alba Ovalle, Orlando (1996): “Disponibilidad léxica en el español dominicano: aspectos sociolingüísticos”, en M. Arjona Iglesias, J. López Chávez *et al.* (eds.), *Actas del X Congreso Internacional de la ALFAL*. México: UNAM, 742–749.
- Alba Ovalle, Orlando (1998): “Variable léxica y dialectología hispánica”, *La Torre, Revista de la Universidad de Puerto Rico*, 7/8: 299–316.
- Alba Ovalle, Orlando (1999): “Densidad de anglicismos en el léxico disponible de la República Dominicana”, en J. A. Samper Padilla *et al.* (eds.), *Actas del XI Congreso Internacional de ALFAL*. Las Palmas de Gran Canaria: Universidad/Librería Nogal, vol. 2, 853–865.
- Alba Ovalle, Orlando (2000a): “Densidad de anglicismos en el léxico disponible”, en O. Alba, *Nuevos aspectos del español en Santo Domingo*. Santo Domingo: Librería la Trinitaria y Brigham Young University, 73–98.
- Alba Ovalle, Orlando (2000b): “Variable léxica y comparación dialectal”, en O. Alba (ed.), *Nuevos aspectos del español en Santo Domingo*. Santo Domingo: Librería la Trinitaria y Brigham Young University, 99–132.
- Alba Ovalle, Orlando (2013): “Variación diacrónica del léxico disponible dominicano”, *Lingüística Española Actual*, 35/1: 127–148.
- Alcalá Venceslada, Antonio (1998) [1951]: *Vocabulario andaluz*. Jaén: Universidad de Jaén y Cajasur.
- Almela Pérez, Ramón, Pascual Cantos Gómez, Aquilino Sánchez, Ramón Sarmiento y Moisés Almela (2005): *Frecuencias del español. Diccionario y estudios léxicos y morfológicos*. Madrid: Universitas.

- Alvar Ezquerro, Manuel (2003): *La enseñanza del léxico y el uso del diccionario*, Cuadernos de Didáctica del español/LE. Madrid: Arco Libros.
- Alvar Ezquerro, Manuel (2005): “La frecuencia léxica y su utilidad en la enseñanza del español como lengua extranjera”, en M.<sup>a</sup> A. Castillo Carballo *et al.* (eds.), *Las gramáticas y los diccionarios en la enseñanza del español como segunda lengua: deseo y realidad. Actas del XV Congreso Internacional de ASELE. Sevilla 22–25 de septiembre de 2004*. Sevilla: Universidad, 19–39.
- Álvarez de la Granja, María (2011): “La presencia del gallego en el léxico disponible del español de Galicia. Análisis formal y funcional”, en B. López Meirama (ed.), *Estudios sobre disponibilidad léxica en el español de Galicia*. Santiago de Compostela: Universidad, 17–102.
- Álvarez de la Granja, María y Belén López Meirama (2013): “A presenza do galego no léxico dispoñible do español de Galicia. Análise distribucional”, en E. Gugenberger, H. Monteagudo y G. Rei-Doval (eds.), *Contacto de linguas, hibrididade, cambio: contextos, procesos e consecuencias*. Santiago de Compostela: Consello da Cultura Galega/Instituto da Lingua Galega, 49–96.
- Areta Lara, María (2009): *El léxico disponible de los estudiantes preuniversitarios navarros. Estudio sociolingüístico*. Tesis doctoral inédita. Pamplona: Universidad.
- Aristizabal, Enrique (1938): *Détermination expérimentale du vocabulaire écrit pour servir de base à l'enseignement de l'orthographe à l'école primaire*. París: H. Champion.
- Arnal Purroy, M.<sup>a</sup> Luisa (ed.) (2008a): *Estudios sobre disponibilidad léxica en los jóvenes aragoneses*. Zaragoza: Institución “Fernando el Católico”.
- Arnal Purroy, M.<sup>a</sup> Luisa (2008b): “Los dialectalismos en el léxico disponible de los jóvenes aragoneses”, en M.<sup>a</sup> L. Arnal Purroy (ed.), *Estudios sobre disponibilidad léxica en los jóvenes aragoneses*. Zaragoza: Institución “Fernando el Católico”, 17–49.
- Arnal Purroy, M.<sup>a</sup> Luisa (2008c): “Niveles socioculturales y léxico dialectal en el vocabulario disponible de Aragón”, en J. L. Blas Arroyo *et al.* (eds.), *Discurso y sociedad II. Nuevas contribuciones al estudio de la lengua en contexto social*. Castellón de la Plana: Universitat Jaume I, 457–469.
- Arnal Purroy, M.<sup>a</sup> Luisa (2009): “Áreas lingüísticas y dialectalismos en los jóvenes aragoneses (materiales de disponibilidad léxica de Aragón)”, en V. Lagüéns Gracia (ed.), *Baxar para subir. Colección de estudios en memoria de Tomás Buesa Oliver*. Zaragoza: Institución “Fernando el Católico”, 265–296.
- Arnal Purroy, M.<sup>a</sup> Luisa (coord.), Rosa Castañer, José María Enguita, Vicente Lagüéns y Ana Beatriz Moliné (2004): *Léxico disponible de Aragón*. Zaragoza: Libros Pórtico.
- Ayora Esteban, Carmen (2003): *Disponibilidad léxica en Ceuta: aspectos sociolingüísticos*. Tesis doctoral inédita. Sevilla: Universidad.
- Ayora Esteban, Carmen (2006): *Disponibilidad léxica en Ceuta: aspectos sociolingüísticos*. Cádiz: Universidad.
- Ávila, Raúl (1995): “Difusión Internacional del Español por Radio y Televisión (DIES-RTV): Procedimiento para recopilar el corpus por países”, México D.F.: El Colegio de México. <http://www.colmex.mx/academicos/cell/ravila/docs/10crpest.pdf> [19.05.2014].
- Ávila Muñoz, Antonio M. (2005): “Estudios léxico-estadísticos en Málaga: La disponibilidad léxica”, Ms. Málaga: Universidad.
- Ávila Muñoz, Antonio M. (2006): *Léxico disponible de estudiantes preuniversitarios de Málaga*. Málaga: Universidad.
- Ávila Muñoz, Antonio M. (2007): “Léxico disponible y ortografía. Condicionantes sociales y hábitos culturales de influencia”, en J. A. Moya Corral y M. Sosinski (eds.), *Las hablas andaluzas y la enseñanza de la Lengua. Actas de las XII Jornadas sobre la enseñanza de la lengua española*. Granada: Universidad, 25–46.
- Ávila Muñoz, Antonio M. y José M.<sup>a</sup> Sánchez Sáez (2011): “La posición de los vocablos en el cálculo del índice de disponibilidad léxica: Procesos de reentrada en las listas del léxico disponible de la ciudad de Málaga”, *Estudios de Lingüística*, 25: 45–74.

- Ávila Muñoz, Antonio M. y Juan A. Villena Ponsoda, Juan A. [eds.] (2010): *Variación social del léxico disponible en la ciudad de Málaga. Diccionario y análisis*. Málaga: Sarriá.
- Ayora Esteban, Carmen (2002): “La disponibilidad léxica en una situación de contacto de lenguas”, en M.<sup>a</sup> D. Muñoz Núñez *et al.* (eds.), *IV Congreso de Lingüística General. Cádiz, del 3 al 6 de abril de 2000*. Cádiz: Universidad de Cádiz/Universidad de Alcalá, II, 159–170.
- Ayora Esteban, Carmen (2003): *Disponibilidad léxica en Ceuta: aspectos sociolingüísticos*. Tesis doctoral inédita. Sevilla: Universidad.
- Ayora Esteban, Carmen (2004): “Deficiencias en la adquisición del léxico como problema específico de las lenguas de contacto”, en M. Villayandre Llamazares (ed.), *Actas del V Congreso de Lingüística General*, Tomo 1. Madrid: Arco Libros, 281–293.
- Ayora Esteban, Carmen (2006): *Disponibilidad léxica en Ceuta: aspectos sociolingüísticos*. Cádiz: Universidad.
- Ayora Esteban, Carmen (2007): “Los estudios de disponibilidad léxica: ámbitos de aplicación”, en P. Cano López (coord.), *Actas VI Congreso de Lingüística General*, vol. 2, Tomo 2: *Las lenguas y su estructura* (IIB). Santiago de Compostela: Arco Libros, 2367–2378.
- Azurmendi Ayerbe, M.<sup>a</sup> José (1983): *Elaboración de un modelo para la descripción sociolingüística del bilingüismo y su aplicación parcial en la comarca de San Sebastián*. San Sebastián: Caja de Ahorros Provincial de Guipúzcoa.
- Bailey Victory, John (1971): *A Study of Lexical Availability among Monolingual-Bilingual Speakers of Spanish and English*, Tesina inédita. Houston: Rice University.
- Bailly, Danielle (1984): *Éléments de Didactique des Langues*. París: Association de professeurs de langues vivantes.
- Barsalou, Lawrence W. (1983): “Ad hoc categories”, *Memory and Cognition*, 11/3: 211–227.
- Bartol Hernández, José A. (1998): “Anglicismos en el español de finales de siglo. Anglicismos en el léxico disponible de dos regiones hispanas”, en K. Kent y M. D de la Calle (eds.), *Visiones salmantinas*. Salamanca: Universidad de Salamanca y Ohio State University.
- Bartol Hernández, José A. (2001): “Reflexiones sobre la disponibilidad léxica”, en J. A. Bartol Hernández (coord.), *Nuevas aportaciones al estudio de la lengua española. Investigaciones filológicas*. Salamanca: Luso-Española Ediciones, 221–235.
- Bartol Hernández, José A. (2004a): *Léxico disponible de Soria. Estudio y diccionarios*. Burgos: Instituto Castellano Leonés de la Lengua.
- Bartol Hernández, José A. (2004b): “Léxico disponible de Soria: Primeros datos”, *Archivo de filología aragonesa*, vol. 59–60, 2, 2002–2004: 1075–1104.
- Bartol Hernández, José A. (2005): “Extranjerismos disponibles: Su integración y vitalidad en una comunidad de habla (Provincia de Soria, España)”, *Spanish in Context*, 2/2: 203–229.
- Bartol Hernández, José A. (2006): “La disponibilidad léxica”, *Revista Española de Lingüística*, 36: 379–384.
- Bartol Hernández, José A. (2008): “Variación léxica del español: los léxicos disponibles de Aragón y Soria”, en M.<sup>a</sup> L. Arnal Purroy (ed.), *Estudios sobre disponibilidad léxica en los jóvenes aragoneses*. Zaragoza: Institución “Fernando el Católico”, 207–226.
- Bartol Hernández, J. A. (2011): “Léxicos disponibles de zonas bilingües: interferencias sobre el español”, en B. López Meirama (ed.), *Estudios sobre disponibilidad léxica en el español de Galicia*. Santiago de Compostela: Universidad, 157–187.
- Bartol Hernández, José A. y Natividad Hernández Muñoz (2004): “Dispolex: base de datos de la disponibilidad léxica”. Panel de investigación presentado en el *VI Congreso de Lingüística General*. Santiago de Compostela, 3–7 de mayo de 2004.
- Bellón Fernández, Juan J. (2003): *Léxico disponible de la provincia de Córdoba*, Tesis doctoral inédita. Las Palmas de Gran Canaria: Universidad.
- Bellón Fernández, Juan J. (2004–2005): “Anglicismos en el centro de interés ‘La ropa’ en el léxico disponible de la provincia de Córdoba”, *Philologica canariensis*, 10–11: 15–38.

- Bellón Fernández, Juan J. (2011): *Léxico disponible de la provincia de Córdoba*. Las Palmas de Gran Canaria/Córdoba: Universidad de Las Palmas de Gran Canaria/Universidad de Córdoba.
- Benítez Pérez, Pedro (1991): “Reseña a Azurmendi (1983)”, *Revista Española de Lingüística* 21/2: 355–357.
- Benítez Pérez, Pedro (1992a): “Listas abiertas y listas cerradas de palabras en disponibilidad léxica”, en *Actas del VIII Congreso Nacional de la Asociación Española de Lingüística Aplicada*. Vigo: AESLA, 117–127.
- Benítez Pérez, Pedro (1992b): “Disponibilidad léxica en la zona metropolitana de Madrid”, *Boletín de la Academia Puertorriqueña de la Lengua Española*, 1/1: 71–102.
- Benítez Pérez, Pedro (1993): “Anglicismos en la disponibilidad léxica de Madrid”, Comunicación presentada en el X Congreso Internacional de ALFAL.
- Benítez Pérez, Pedro (1994a): “Léxico real/irreal en los manuales de español para extranjeros”, en S. Montesa Peydró y A. Garrido Moraga (eds.), *Español para extranjeros: didáctica e investigación*, *Actas del II Congreso Nacional de la ASELE*. Málaga: ASELE, 325–333.
- Benítez Pérez, Pedro (1994b): “¿Qué vocabulario hay que enseñar en las clases de español como lengua extranjera?”, *Didáctica del español como lengua extranjera*. Fundación Actilibre: 9–12.
- Benítez Pérez, Pedro (1994c): “Convergencias y divergencias en el léxico de alumnos de COU”, *Revista de Estudios de Adquisición de la Lengua Española (REALE)*, 2: 39–45.
- Benítez Pérez, Pedro (1995): “Disponibilidad léxica en Madrid: análisis cuantitativo”, en L. Rodríguez-Fonseca y M. Vázquez (eds.), *Actas del III Seminario Internacional sobre ‘Aportes de la lingüística a la enseñanza de la lengua materna’*, números especiales de la Revista de Adquisición de la Lengua Española.
- Benítez Pérez, Pedro, Clara E. Hernández Cabrera y José A. Samper Padilla. (1995): “Léxicos básicos de España (LEBAES) y de Canarias (LEBAICan). Proyecto de investigación”, en *REALE*, 3: 9–17.
- Bernstein, Basil (1971): *Clases, códigos y control. Estudios teóricos para una sociología del lenguaje*, vol. 1. Madrid: AKAL, 1989.
- Bernstein, Basil (1975) *Clases, códigos y control. Hacia una teoría de las transmisiones educativas*, vol. 2. Madrid: AKAL, 1989.
- Blanco, Marta (2011): “La ortografía en el léxico disponible del español de Galicia”, en B. López Meirama (ed.), *Estudios sobre disponibilidad léxica en el español de Galicia*. Santiago de Compostela: Universidad, 189–216.
- Blas Arroyo, José L. y Manuela Casanova Ávalos (2003a): “La influencia de la red educativa y del entorno sociocultural en la disponibilidad léxica. Estudio de las comunidades de habla castellonenses”, en F. Sánchez Miret (ed.), *Actas del XXIII Congreso Internacional de Lingüística y Filología Románica*. Tübinga: Niemeyer, vol. 5, 17–33.
- Blas Arroyo, José L. y Manuela Casanova Ávalos (2003b): “Interacción de factores sociales y de adscripción lingüística en el léxico disponible de una comunidad escolar española”, en *Lenguas modernas*, 28–29: 165–189.
- Blas Arroyo, José L. y Manuela Casanova Ávalos (2003c): “La influencia de la lengua materna en algunos valores de la disponibilidad léxica en el español de una comunidad bilingüe”, en I. Doval Reixa y M.ª R. Pérez Rodríguez (eds.), *Adquisición, enseñanza y contraste de lenguas, bilingüismo y traducción*. Vigo: Universidad, 49–62.
- Borrego Nieto, Julio (2001): “La marcación diafásica en los diccionarios”, en J. A. Bartol Hernández et al. (eds.), *Nuevas aportaciones al estudio de la lengua española. Investigaciones filológicas*. Salamanca: Universidad, 237–245.
- Borrego Nieto, Julio (2002a): “Niveles de lengua y diccionarios”, en J. L. Blas Arroyo et al. (eds.), *Estudios sobre lengua y sociedad*. Castellón de la Plana: Universitat Jaume I, 105–151.
- Borrego Nieto, Julio (2002b): “Estudiantes universitarios y jubilados, ¿dos culturas léxicas diferentes?”, Comunicación presentada en las *Terceras Jornadas de Reflexión Filológica*. Salamanca, 28 de febrero – 2 de marzo de 2002.

- Borrego Nieto, Julio (2004): "Algunas preguntas en relación con el concepto de léxico disponible", en J. Prado Aragonés y M.<sup>a</sup> V. Galloso Camacho (eds.), *Diccionario, léxico y cultura*. Huelva: Universidad, 59–69.
- Borrego Nieto, Julio (2008): "Edad y cultura léxicas", en M.<sup>a</sup> L. Arnal Purroy (ed.), *Estudios sobre disponibilidad léxica en los jóvenes aragoneses*. Zaragoza: Institución "Fernando el Católico", 227–244.
- Borrego Nieto, Julio (2009): "Niveles de lengua en el léxico disponible", en M.<sup>a</sup> V. Camacho Taboada et al. (eds.), *Estudios de lengua española. Descripción, variación y uso. Homenaje a Humberto López Morales*. Madrid–Frankfurt: Iberoamericana-Vervuert, 53–76.
- Borrego Nieto, Julio y Carmen Fernández Juncal (2002): "Léxico disponible: aplicaciones a los estudios dialectales", en M.<sup>a</sup> D. Muñoz et al. (eds.), *Actas del II Congreso de Lingüística General*. Cádiz: Universidad, vol. 2, 297–306.
- Borrego Nieto, Julio y Carmen Fernández Juncal (2003): "¿En qué cambia la universidad la disponibilidad léxica de los preuniversitarios?", en F. Moreno et al. (coords.), *Lengua, variación y contexto. Estudios dedicados a H. López Morales*. Madrid: Arco Libros, 167–178.
- Bortolini, Umberta, Carlo Tagliavini y Antonio Zampolli. (1972): *Lessico di Frequenza della Lingua Italiana Contemporanea*. Milán: Garzanti.
- Briz Gómez, Antonio (1985): *El léxico de la vid en la comarca de Requena-Utiel*. Valencia: Institució Alfons el Magnànim, IVEI.
- Buchanan, Milton A. (1927): *A Graded Spanish Word Book*. Toronto: Toronto University Press.
- Butrón, Gloria (1987): *El léxico disponible: índices de disponibilidad*, Tesis doctoral inédita. Río Piedras: Universidad de Puerto Rico.
- Butrón, Gloria (1989): "Aspectos sociolingüísticos de la disponibilidad léxica", *Asomante*, 1/2: 29–37.
- Butrón, Gloria (1991): "Nuevos índices de disponibilidad léxica" en H. López Morales (ed.), *La enseñanza del español como lengua materna*. Río Piedras: Universidad de Puerto Rico.
- Cañizal Arévalo, Alma V. (1987): *Disponibilidad léxica en escolares de primaria terminada. Análisis de seis centros de interés*, Memoria de licenciatura inédita. México: UNAM.
- Cañizal Arévalo, Alma V. (1991): "Redes semánticas y disponibilidad léxica en el español de escolares mexicanos", en C. Hernández et al. (eds.), *El español de América*. Valladolid: Junta de Castilla y León, 631–637.
- Carcedo González, Alberto (1998a): "Sobre las pruebas de disponibilidad léxica para estudiantes de español/LE", *Revista del Instituto de Lengua y Cultura Españolas (RILCE)*, 14.2 (número monográfico: "Español como lengua extranjera: investigación y docencia"): 205–224.
- Carcedo González, Alberto (1998b): "Tradicición y novedad en las aportaciones hispánicas a los estudios de disponibilidad léxica", en *Lingüística*, 10, ALFAL: 5–68.
- Carcedo González, Alberto (1999a): "Desarrollo de la competencia léxica en español LE: análisis de cuatro fases de disponibilidad", *Pragmalingüística* 5–6: 75–94.
- Carcedo González, Alberto (1999b): "Estudio comparativo del vocabulario español (LE) disponible de estudiantes finlandeses y el de la sintopía madrileña: propuestas didácticas", *Documentos de Español Actual*, vol. 1: 73–87.
- Carcedo González, Alberto (1999c): "Análisis de errores léxicos del español en la interlingua de los finlandeses", en T. Jiménez et al. (eds.), *Español como lengua extranjera: enfoque comunicativo y gramática*, *Actas del IX Congreso Internacional de ASELE*. Santiago de Compostela: Universidad, 465–472.
- Carcedo González, Alberto (2000a): *Disponibilidad léxica en español como lengua extranjera: el caso finlandés (estudio del nivel preuniversitario y cotejo con tres fases de adquisición)*, *Anales Universitatis Turkuensis, Humaniora*, Ser. B, Tom. 238. Turku: Universidad de Turku.
- Carcedo González, Alberto (2000b): "Índices léxico-estadísticos y graduación del vocabulario en la enseñanza de E/LE", M. Franco Figueroa et al. (eds.), *Nuevas perspectivas en la enseñanza*

- del español como lengua extranjera, Actas del X Congreso de ASELE*. Cádiz: Universidad, 175–183.
- Carcedo González, Alberto (2000c): “La lengua como manifestación de otredad cultural (o convergencia intercultural)”, en Aguirre Romero (ed.), *Espéculo*, monográfico *Cultura e Intercultura en la enseñanza del español como lengua extranjera*. Universidad Complutense de Madrid. <http://www.ucm.es/info/especulo/ele/carcedo.html> [20.05.2014].
- Carcedo González, Alberto (2001): *Léxico disponible de Asturias*. Turku: Universidad.
- Carcedo González, Alberto (2003): “Unidad y variedad diatópica de la disponibilidad léxica del español: comparación de los inventarios de Puerto Rico, Cádiz y Asturias”, en F. Moreno Fernández et al. (eds.), *Lengua, Variación y Contexto, Estudios dedicados a Humberto López Morales*. Madrid: Arco/Libros, 199–225.
- Carcedo González, Alberto (2004): “La variable léxica disponible en la comparación interdialectal: compatibilidad de la norma asturiana con otras sintopías hispanohablantes”, en V. Sánchez Corrales (ed.), *Actas del XIII Congreso Internacional de la Asociación de Lingüística y Filología de América Latina (ALFAL)*. San José de Costa Rica: Universidad de Costa Rica, 661–679.
- Carrera de la Red, M<sup>a</sup> Fátima y Winifred E. Bradley (2004): “Extranjerismos en el Léxico Disponible de Cantabria”, en M. Villayandre Llamazares (coord.), *Actas del V Congreso de Lingüística General: León 5–8 de marzo de 2002*. León: Universidad, vol. 1: 535–545.
- Casanova Ávalos, Manuela (2006): “La disponibilidad léxica en la Comunidad Valenciana”, en J. L. Blas Arroyo et al. *Discurso y sociedad. Contribuciones al estudio de la lengua en contexto social*. Castellón de la Plana: Universidad Jaume I, 737–751.
- Casanova Ávalos, Manuela (2008): “Anglicismos en el léxico disponible de Castellón”, en J. L. Blas Arroyo et al., *Discurso y sociedad II: nuevas contribuciones al estudio de la lengua en contexto social*. Castellón de la Plana: Universitat Jaume I, 471–499.
- Casas Gómez, Miguel (2005): “Relaciones «significativas», relaciones semánticas y relaciones léxicas”, *Lingüística Española Actual*, 27: 5–31.
- Casas Gómez, Miguel (2011): “Problemas y criterios lingüísticos subyacentes a una tipología de relaciones en semántica”, *Lorenzo Hervás*, 20 (Extraordinario): 63–108.
- Castañer Martín, Rosa (2008): “Los extranjerismos en el léxico disponible de los estudiantes aragoneses”, en M.<sup>a</sup> L. Arnal Purroy (ed.), *Estudios sobre disponibilidad léxica en los jóvenes aragoneses*. Zaragoza: Institución “Fernando el Católico”, 51–73.
- Castillo Carballo, M.<sup>a</sup> Auxiliadora (1998): “El término ‘colocación’ en la lingüística actual”, *Lingüística Española Actual*, 20/1: 41–54.
- Collins, Allan M. y Elisabeth F. Loftus (1975): “A Spreading-activation theory of semantic processing”, *Psychological Review*, 82/6: 407–428.
- Collins, Allan M. y M. Ross Quilliam (1969): “Retrieval time from semantic memory”, *Journal of Verbal Learning and Verbal Behaviour*, 8: 240–247.
- Corominas i Vigneaux, Joan (1973): *Breve diccionario etimológico de la lengua castellana*, 3<sup>a</sup> ed. Madrid: Gredos.
- Corpas Pastor, Gloria (1996): *Manual de fraseología española*. Madrid: Gredos.
- Coseriu, Eugen (1977): *Principios de semántica estructural*. Madrid: Gredos.
- Cruz Alonso, Raquel (2012): *El léxico disponible de Castilla y León*. Tesis doctoral inédita. Salamanca: Universidad.
- Cruzado Romero, Juan (2003): *El léxico disponible de alumnos de 2º, 4º de ESO y 2º de Bachillerato-LOGSE de Moguer (Huelva)*. Tesis doctoral inédita. Huelva: Universidad.
- Dalurzo, M.<sup>a</sup> Julia y Luis A. González (2010): “Disponibilidad Léxica en Córdoba, Argentina”, en *Salud y aprendizajes lingüísticos. Complejidades en la enseñanza de la lengua*. Córdoba, Argentina: Editorial Comunicarte, Tomo 2, 279–344.
- Dabène, Louise (1994): *Repères sociolinguistiques pour l’enseignement des langues*. Paris: Hachette.

- Dimitrijevic, Naum R. (1969): *Lexical Availability*. Heidelberg: Julius Gross Verlag.
- Echenique, M<sup>a</sup> Teresa (2000): “Algunas consideraciones sobre conexiones románicas variadas en la configuración del léxico básico”, en C. Saralegui y M. Casado (eds.), *Estudios en homenaje al profesor Fernando González Ollé*. Pamplona: Universidad de Navarra, 449–464.
- Echeverría, Max S. (1991): “Crecimiento de la disponibilidad léxica en estudiantes chilenos de nivel básico y medio”, en Humberto López Morales (ed.), *La enseñanza del español como lengua materna, Actas del II Seminario Internacional sobre ‘Aportes de la lingüística a la enseñanza del español como lengua materna’*. Río Piedras: Universidad de Puerto Rico, 61–78.
- Echeverría, Max S. (1993): “Estructura y perfil de la competencia léxica”, *Revista de Lingüística Teórica y Aplicada (RLA)*, 31: 55–66.
- Echeverría, Max S. et al. (1985): “Disponibilidad léxica en Educación Media”, *Revista de Lingüística Teórica y Aplicada*, 25: 55–102.
- Echeverría, Max S., M.<sup>a</sup> Olivia Herrera, Patricio Moreno y Francisco Pradenas (1987): “Disponibilidad léxica en Educación Media: resultados cuantitativos”, *Revista de Lingüística Teórica y Aplicada (RLA)*, 25: 55–115.
- Echeverría, Max S., Roberto Vargas, Paula Urzua, y Roberto Ferreira (2008): “DispoGrafo: una nueva herramienta computacional para el análisis de relaciones semánticas en el léxico disponible”, *Revista de Lingüística Teórica y Aplicada (RLA)*, 46/1: 81–91.
- Enguita Utrilla, José M.<sup>a</sup> (2008): “Norma urbana y norma rural en el léxico disponible de los jóvenes aragoneses”, en M.<sup>a</sup> L. Arnal Purroy (ed.), *Estudios sobre disponibilidad léxica en los jóvenes aragoneses*. Zaragoza: Institución “Fernando el Católico”, 75–101.
- Escoriza Morera, Luis (2003): “Lenguas en contacto y disponibilidad léxica: La situación lingüística e intercultural de Ceuta y Gibraltar. Proyecto de investigación”, en *Linred, Lingüística en la red*, 2, [http://www.linred.es/informacion\\_pdf/informacion4\\_19102004.pdf](http://www.linred.es/informacion_pdf/informacion4_19102004.pdf), [20.05.2014].
- Escoriza Morera, Luis (2007): “Disponibilidad léxica y multilingüismo. El contacto entre inglés y español en Gibraltar”, en P. Cano López et al. (eds.), *Actas del VI Congreso de Lingüística General, Santiago de Compostela, 3–7 de mayo de 2004*. vol. 2, tomo 2, Madrid: Arco/Libros, 2477–2484.
- Etxebarria Arostegui, Maitena (1996): “Disponibilidad léxica en escolares del País Vasco: variación sociolingüística y modelos de enseñanza bilingüe”, *Revista Española de Lingüística*, 26/2: 301–325.
- Etxebarria Arostegui, Maitena (1999): “Disponibilidad léxica y enseñanza de la lengua materna en el área metropolitana del Gran Bilbao”, en J. A. Samper Padilla et al. (eds.), *Actas del XI Congreso Internacional de ALFAL*. Las Palmas de Gran Canaria: Universidad de Las Palmas de Gran Canaria/Librería Nogal, vol. 2, 1479–1494.
- Fajardo Aguirre, Alejandro (1993): “Los vocabularios básicos”, en *Actas del Tercer Congreso Nacional de ASELE. El español como lengua extranjera: De la teoría al aula*. Málaga: ASELE.
- Fernández Juncal, Carmen (2008): *Léxico disponible de Burgos*. Burgos: Fundación Instituto Castellano y Leonés de la Lengua.
- Fernández Juncal, Carmen (2013): *Léxico disponible de Cantabria. Estudio sociolingüístico*. Salamanca: Universidad.
- Fernández Smith, Gérard (2008): “Disponibilidad léxica en Ceuta: aspectos sociolingüísticos”, *Lingüística en la Red*, 6: 1–6. [http://www.linred.es/resenas\\_pdf/resena31\\_16112008.pdf](http://www.linred.es/resenas_pdf/resena31_16112008.pdf), [20.05.2014].
- Fernández Smith, Gérard et al. (2006): “Proyecto sobre la disponibilidad léxica en alumnos preuniversitarios de Melilla”, en J. L. Blas Arroyo et al. (eds.), *Discurso y Sociedad. Contribuciones al estudio de la lengua en contexto social*. Castellón de la Plana: Universitat Jaume I, 767–777.
- Fernández Smith, Gérard et al. (2008): *Léxico disponible de Melilla: estudio sociolingüístico y repertorios léxicos*. Madrid: Arco Libros.
- Ferreira Campos, R. y Max S. Echeverría (2010): “Redes semánticas en el léxico disponible de inglés L1 e inglés LE”, *Onomázein*, 21/1: 133–153.

- Frago Gracia, Juan A. (1999): “Español culto y pervivencias dialectales en el Aragón dieciochesco”, en *Jornadas de Filología aragonesa en el L aniversario del AFA*. Zaragoza: IFC, 139–164.
- Fuentes González, Antonio D. (1994): “Índices de mortandad lingüística del léxico cañero-azucarero. Análisis sincrónico”, en A. Malpica (ed.), *1492: Lo dulce a la conquista de Europa*. Granada: Diputación Provincial de Granada, 236–245.
- Galisson, Robert (1991): *De la langue à la culture par les mots*. París: Clé International.
- Galoso Camacho, M.<sup>a</sup> Victoria (1997): “Tratamiento y perspectivas del léxico disponible: el caso de la provincia de Zamora”, *Revista de Estudios de Adquisición de la Lengua Española*, 8: 23–33.
- Galoso Camacho, M.<sup>a</sup> Victoria (1998a): *El Léxico disponible en el nivel preuniversitario (provincia de Zamora)*. Memoria de licenciatura inédita. Salamanca: Universidad.
- Galoso Camacho, M.<sup>a</sup> Victoria (1998b): “Hacia una definición de léxico disponible: posibles aplicaciones”, en J. Fernández González et al. (eds.): *Lingüística para el siglo XXI. Actas del III Congreso de Lingüística General*. Salamanca: Universidad, vol. 1, 711–718.
- Galoso Camacho, M.<sup>a</sup> Victoria (2002a): *El léxico disponible en el nivel preuniversitario. Provincia de Zamora*. Huelva: Universidad.
- Galoso Camacho, M.<sup>a</sup> Victoria (2002b): *El léxico de los estudiantes preuniversitarios en el distrito universitario de Salamanca*. Tesis doctoral inédita. Salamanca: Universidad.
- Galoso Camacho, M.<sup>a</sup> Victoria (2002c): “¿Para qué sirven los estudios de léxico disponible?”, en M.<sup>a</sup> D. Muñoz Núñez et al. (eds.), *Actas del IV Congreso de Lingüística General. Cádiz, del 3 al 6 de abril de 2000*. Cádiz: Universidad de Cádiz/Universidad de Alcalá, vol. 3, 1143–1151.
- Galoso Camacho, M.<sup>a</sup> Victoria (2003a): *El léxico disponible de Ávila, Salamanca y Zamora*. Burgos: Fundación Instituto Castellano y Leonés de la Lengua.
- Galoso Camacho, M.<sup>a</sup> Victoria (2003b): “Lematización en masculino y femenino de los nombres de profesiones y oficios”, *DEA (Documentos de Español Actual)*, 5: 33–46.
- Galoso Camacho, M.<sup>a</sup> Victoria (2003c): “El léxico disponible de los estudiantes preuniversitarios en Salamanca”, *Revista de Estudios*, 50: 201–224.
- Galoso Camacho, M.<sup>a</sup> Victoria (2005a): “Léxico y vocabulario en los alumnos de ESO: pautas para su desarrollo”, en J. Prado Aragonés y J. Heras Borrero (eds.), *Aspectos de Lengua y Literatura y su enseñanza en Andalucía*. Granada: Grupo Editorial Universitario, 117–131.
- Galoso Camacho, M.<sup>a</sup> Victoria (2005b): “El sexo y el lenguaje”, en M.<sup>a</sup> M. Gallego Durán y R. García Gutiérrez (coords.), *El legado plural de las mujeres*. Sevilla: Alfar, 117–132.
- Galoso Camacho, M.<sup>a</sup> Victoria (2008): “La ortografía en las encuestas onubenses de disponibilidad léxica”, en M.<sup>a</sup> V. Galoso Camacho (ed.), *Estudios en torno a la lengua y literatura en Andalucía*. Granada: Grupo Editorial Universitario, 25–34.
- Galoso Camacho, M.<sup>a</sup> Victoria y Josefina Prado Aragonés (2004): “La estructura estadística del léxico disponible de informantes nativos e informantes de ELE”, en M.<sup>a</sup> A. Castillo Carballo et al. (coord.), *Las gramáticas y los diccionarios en la enseñanza del español como segunda lengua: deseo y realidad, Actas del XV Congreso Internacional de ASELE*, Sevilla: Universidad, 370–375.
- Galoso Camacho, M.<sup>a</sup> Victoria y Josefina Prado Aragonés (2005): “Análisis cuantitativo del léxico disponible de los preuniversitarios de Huelva y contraste con el de los preuniversitarios de Salamanca”, en L. Santos Río et al. (eds.), *Palabras, norma, discurso. En memoria de Fernando Lázaro Carreter*. Salamanca: Universidad, 491–502.
- García Casero, M.<sup>a</sup> José (2013): *El léxico disponible en estudiantes de 4º de Enseñanza Secundaria Obligatoria en Santander*. Tesis doctoral inédita. Universidad de Cantabria.
- García de Hoz, Víctor (1952): *Vocabulario usual, común y fundamental*. Madrid, CSIC.
- García Domínguez, M.<sup>a</sup> Jesús, Vicente Marrero Pulido, José A. Pérez Martín y Gracia Piñero Piñero (1994): “Estudio de la disponibilidad léxica en Gran Canaria. La variable geográfica y el tipo de educación”, *Revista de Estudios de Adquisición de la Lengua Española (REALE)*, 2: 65–72.
- García Marcos, Francisco J. (1999): *Fundamentos críticos de sociolingüística*. Almería: Universidad.

- García Megía, Antonio (2004): *La disponibilidad Léxica en la Ciudad de Almería entre los nueve y los doce años*. Almería: Universidad.
- García Moutón, Pilar (1999): *Cómo hablan las mujeres*. Cuadernos de lengua española. Madrid: Arco Libros.
- Garret, Henry E. (1980): *Estadística en Psicología y Educación*. Buenos Aires: Paidós (1973, 1ª edición).
- Giacobbe, Jorge (1990): “Le recours à la langue première: Une approche cognitive”, *Le Français dans le Monde. Recherches et Applications*. Número spécial Février–Mars: 115–123.
- Gómez Capuz, Juan (2005): *La inmigración léxica*. Madrid: Arco/Libros.
- Gómez Devís, M.ª Begoña (2004): *La disponibilidad léxica de los estudiantes preuniversitarios valencianos: metodología, análisis sociolingüístico y aplicaciones*. Tesis doctoral inédita. Valencia: Universidad.
- Gómez Devís, M.ª Begoña (2005): “Aprovechamiento del léxico disponible”. *Lenguaje y textos*, 23: 69–74.
- Gómez Devís, M.ª Begoña (2010): “Aportes de la disponibilidad léxica a la investigación sobre fenómenos de contacto lingüístico”, en M. García (ed.), *Interacció Comunicativa i Ensenyament de Llengües. Quaderns de Filologia*. Valencia: Universidad, 233–240.
- Gómez Devís, M.ª Begoña y Francesc Llopis Rodrigo (2010): “El léxico disponible de Valencia: convergencia conceptual”, *Revista de Lingüística y Lenguas Aplicadas*: 53–76.
- Gómez Molina, José R. (2006): “La variable sexo en los estudios de disponibilidad léxica”, *Revista de Lingüística Teórica y Aplicada*. 44/1: 47–65.
- Gómez Molina, José R. y M.ª Begoña Gómez Devís (2004): *La disponibilidad léxica de los estudiantes preuniversitarios valencianos. Estudio de estratificación sociolingüística*, Valencia: Universidad.
- González Becker, Marina y Patricia Vargas Sandoval (1992): “La disponibilidad léxica como instrumento de planificación curricular: un diagnóstico al hablante medio en lengua materna en seis centros de interés (arte, derecho, economía, política, sentimientos y teología)”, en *I Congreso Internacional sobre la Enseñanza del Español*. Madrid, 27–31 de enero. Ms.
- González Martínez, Adolfo E. (1997): *Disponibilidad léxica de Cádiz*. Tesis doctoral inédita. Cádiz: Universidad.
- González Martínez, Adolfo E. (1999): “Andalucismos en el léxico disponible de Cádiz”, *TAVIRA*, 17: 181–193.
- González Martínez, Adolfo E. (2001): “Compatibilidad y discrepancia entre los léxicos disponibles de Puerto Rico y Cádiz”, *Pragmalingüística*, 8, 2000–2001: 179–196.
- González Martínez, Adolfo E. (2002): *La disponibilidad léxica de los alumnos preuniversitarios de la provincia de Cádiz*. Cádiz: Universidad.
- González Martínez, Adolfo y Josefa Orellana Ramírez. (2000): “Cotejo de los léxicos disponibles de Cádiz y Zamora”, en A. Carcedo González (ed.), *Documentos de Español Actual (DEA)*, 2: 139–160.
- González Martínez, Adolfo y Josefa Orellana Ramírez (2006): “Anglicismos en el léxico disponible de la provincia de Cádiz (España)”, *Boletín de lingüística*, 18/25: 3–21.
- Gougenheim, George (1967): “La statistique de vocabulaire et son application dans l’enseignement des langues”, *Revue de l’enseignement Supérieur*, 3: 137–144.
- Gougenheim, George, René Michéa, Paul Rivenc y Aurélien Sauvageot (1954): *Le français élémentaire*. Paris: Ministerio de Educación Nacional.
- Gougenheim, George, René Michéa, Paul Rivenc y Aurélien Sauvageot (1956): *L’élaboration du français élémentaire. Étude sur l’établissement d’un vocabulaire et d’une grammaire de bas*. Paris: Didier.
- Gougenheim, George, René Michéa, Paul Rivenc y Aurélien Sauvageot (1958): *Français Fondamental (1er degré)*. Paris: Didier.

- Gougenheim, George, René Michéa, Paul Rivenc y Aurélien Sauvageot (1964): *L'élaboration du Français fondamental (I degré). Étude sur l'établissement d'un vocabulaire et d'une grammaire de bas*. Paris: Didier.
- Gougenheim, George, René Michéa, Paul Rivenc y Aurélien Sauvageot (1967): *L'élaboration du Français fondamental (I degré). Étude sur l'établissement d'un vocabulaire et d'une grammaire de base (nouvelle édition refondue et augmentée)*. Paris: Didier.
- Guiraud, Pierre (1954): *Les caractères statistiques du vocabulaire*. Paris: P.U.F.
- Guiraud, Pierre (1959): *Problèmes et méthodes de la statistique linguistique*. Dordrecht, Reidel.
- Henmon, V. A. C. (1924). *A French Word Book Based on a Count of 400.000 Running Words*. Madison: Madison Bureau of Educational Research, University of Wisconsin.
- Herdan, Gustav (1956): *Language as Choice and Chance*. Groningen: P. Noordhoff N. V.
- Herdan, Gustav (1962): *The Calculus of Linguistic Observations*. La Haya, Mouton.
- Hernández Cabrera, Clara E. y Marta Samper Hernández (2002–2004): “Léxico disponible, norma culta y norma popular”, *Archivo de Filología Aragonesa*, 59–60: 1341–1358.
- Hernández Cabrera, Clara E. y Marta Samper Hernández (2005): “Léxico regional vs. léxico general en los materiales de disponibilidad de Gran Canaria”, Comunicación presentada al *XIV Congreso Internacional de la ALFAL* (Monterrey, México), Ms.
- Hernández Cabrera, Clara E. y José A. Samper Padilla (2003): “Los dialectalismos en el léxico disponible de Gran Canaria. Análisis de un centro de interés”, en F. Moreno Fernández *et al.* (coords.), *Lengua, variación y contexto. Estudios dedicados a Humberto López Morales*. Madrid: Arco Libros, 1, 339–352.
- Hernández Cabrera, Clara E. y José A. Samper Padilla (2006): “Densidad de dialectalismos y condicionantes sociales en el léxico disponible de Gran Canaria”, en M. Sedano *et al.* (eds.), *Haciendo Lingüística. Homenaje a Paola Bentivoglio*. Caracas: Universidad Central de Venezuela, 537–554.
- Hernández Cabrera, Clara E. y José A. Samper Padilla (2007): “Léxico regional en los materiales de disponibilidad léxica”, *Revista de Filología de la Universidad de La Laguna. Homenaje a Antonio Lorenzo*, 25: 287–297.
- Hernández Muñoz, Natividad (2002): *El léxico disponible de los estudiantes conquenses. Del análisis sociolingüístico a la reflexión metodológica*. Trabajo de grado inédito. Salamanca: Universidad.
- Hernández Muñoz, Natividad (2004): *El léxico disponible de los estudiantes conquenses*. Salamanca: Universidad.
- Hernández Muñoz, Natividad (2005): “La disponibilidad léxica: una herramienta fronteriza para el estudio del léxico en Lingüística y Psicología”, en M.<sup>a</sup> A. Mayor Cinca *et al.* (eds.), *Estudios sobre adquisición del lenguaje. IV Congreso Internacional sobre la adquisición de las lenguas del Estado*. Salamanca: Universidad, 942–954.
- Hernández Muñoz, Natividad (2006): *Hacia una teoría cognitiva integrada de la disponibilidad léxica: el léxico disponible de los estudiantes castellano-manchegos*. Salamanca: Universidad.
- Hernández Muñoz, Natividad (2009a): “Aspectos sociolectales del léxico dialectal”. *Spanish in Context*, 6/2: 224–248.
- Hernández Muñoz, Natividad (2009b): “Variación léxica y zonas dialectales de Castilla-La Mancha”, *Revista de Filología Española*, 89 (2): 279–300.
- Hernández Muñoz, Natividad y Cristina Izura (2010): “¿Perro o musaraña? La recuperación léxica en las categorías semánticas”, *Ciencia Cognitiva: Revista de Difusión Científica*, 4/1: 1–3.
- Izquierdo Gil, M.<sup>a</sup> del Carmen (2003): *La selección del léxico en la enseñanza del español como lengua extranjera*. Tesis doctoral inédita. Valencia: Universidad.
- Jiménez Catalán, Rosa M.<sup>a</sup> (ed.) (2014): *Lexical Availability in English and Spanish as a Second Language*. Dordrecht, Heidelberg, New York, London: Springer.
- Josselson, Harry H. (1953): *Russian Word Count*. Detroit: Wayne University.

- Juilland, Alphonse G. y Eugenio Chang-Rodríguez (1964): *Frequency Dictionary of Spanish Words*. The Hague: Mouton.
- Juilland, Alphonse G., Prior M. H. Edwards y Ileana Juilland (1965): *Frequency Dictionary of Rumanian Words*. La Haya: Mouton.
- Juilland, Alphonse G., Dorothy Brodin y Catherine Davidovitch (1971): *Frequency Dictionary of French Words*. La Haya: Mouton.
- Juilland, Alphonse G. y Vincenzo Traversa (1973): *Frequency Dictionary of Italian Words*. La Haya: Mouton.
- Justicia Justicia, Fernando (1995): *El desarrollo del vocabulario. Diccionario de frecuencias*. Granada: Universidad.
- Justo Hernández, Hortensia G. (1986). *Disponibilidad léxica en colores*. Memoria de licenciatura inédita. México: UNAM.
- Karelovic Burotto, Bruno F. (2010): *Estudio de la estructura de grafos de disponibilidad léxica*. Memoria de grado inédita. Universidad de Concepción, Chile.
- Kleiber, Georges (1995): *La semántica de los prototipos. Categoría y sentido léxico*. Madrid: Visor.
- Labov, William (1972): *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.
- Lagüéns Gracia, Vicente (2008): “La variable sexo en el léxico disponible de los jóvenes aragoneses”, en M.<sup>a</sup> L. Arnal Purroy (ed.), *Estudios sobre disponibilidad léxica en los jóvenes aragoneses*. Zaragoza: Institución “Fernando el Católico”, 103–162.
- Lahuerta Galán, Javier y Mercè Pujol Vila (1996): “El lexicón mental y la enseñanza del vocabulario”, en C. Segoviano (ed.), *La enseñanza del léxico español como segunda lengua: homenaje a Anton e Inge Bemmerlein*. Frankfurt–Madrid: Iberoamericana-Vervuert, 117–129.
- Lehmann, Fritz (1992): “Semantic networks”, en F. Lehmann (ed.), *Semantic Networks in Artificial Intelligence*. Oxford: Pergamon Press, 1–50.
- Llopis Rodrigo, Francesc (2005): “Transferències lèxiques en una comunitat de llengües en contacte”, *Lenguaje y textos*, 23: 59–68.
- Llopis Rodrigo, Francesc (2008): “Disponibilitat lèxica del valencià: classificació dels fenòmens de contacte lingüístic”, en J. L. Blas *et al.* (eds.), *Discurso y sociedad II. Nuevas contribuciones al estudio de la lengua en un contexto social*. Castellón de la Plana: Universitat Jaume I, 571–582.
- Llopis Rodrigo, Francesc y M.<sup>a</sup> Begoña Gómez Devís (2010): *El lèxic disponible de València*. Valencia: Denes.
- Llorente Pinto, M.<sup>a</sup> del Rosario (2005): “Léxico disponible y léxico dialectal en la provincia de Ávila”, en L. Santos Río *et al.* (eds.), *Palabras, norma, discurso. En memoria de Fernando Lázaro Carreter*. Salamanca: Universidad, 681–694.
- López Chávez, Juan (1988): “Por una programación en la enseñanza del vocabulario”, en *IV Simposio: La computación en la educación infantil y juvenil*. México D. F.: SOMECE.
- López Chávez, Juan (1991): “Préstamos, extranjerismos y anglicismos en el español de México. Valores lexicométricos (planteamientos previos)”, en C. Hernández Alonso *et al.* (eds.), *El español de América. Actas de III Congreso Internacional de «El Español de América»*. Valladolid: Junta de Castilla y León, vol. 2: 775–784.
- López Chávez, Juan (1992): “Alcances panhispánicos del léxico disponible”, *Lingüística*, 4: 26–124.
- López Chávez, Juan (1993): *El léxico disponible de escolares mexicanos*. México: Alhambra Mexicana.
- López Chávez, Juan (1994): “Comportamiento sintáctico de algunos verbos ordenados según su grado de disponibilidad léxica”, *Revista de Estudios de Adquisición de la Lengua Española (REALE)*, 1: 67–84.
- López Chávez, Juan (1995a): “El léxico en estudiantes mexicanos de primaria”, en Rodríguez-Fonseca y Vázquez (eds.), *Actas del III Seminario Internacional sobre “Aportes de la lingüística a la enseñanza de la lengua materna”*, números especiales de la *Revista de Adquisición de la Lengua Española*.

- López Chávez, Juan (1995b): “Léxico fundamental panhispánico: realidad o utopía”, en *Actas del III Congreso Internacional sobre el español de América*. Santiago de Chile: Universidad Católica de Chile, 1006–1114.
- López Chávez, Juan y Carlos Strassburguer Frías (1987): “Otro cálculo del índice de disponibilidad léxica”, en *Actas del IV Simposio de la Asociación Mexicana de Lingüística Aplicada, Presente y perspectiva de la lingüística computacional en México*. México: UNAM, 1006–1014.
- López Chávez, Juan y Carlos Strassburguer Frías (1991): “Un modelo más para el cálculo de disponibilidad léxica individual”, en H. López Morales (ed.), *La enseñanza del español como lengua materna*. Río Piedras: Universidad de Puerto Rico, 91–112.
- López González, Antonio M.<sup>a</sup> (2014): “La estructura interna del léxico disponible en español como lengua extranjera (ELE) de los preuniversitarios polacos”, *Studia Romanica Posnaniensis*, 41/1, *Variedad e interdisciplinariedad en los estudios de lengua y literatura en español*: 45–61.
- López Meirama, Belén (2008): *Léxico disponible en el español de Galicia*. Santiago de Compostela: Universidad.
- López Meirama, Belén (ed.) (2011): *Estudios sobre disponibilidad léxica en el español de Galicia*. Santiago de Compostela: Universidad.
- López Morales, Humberto (1973): *Disponibilidad léxica en escolares de San Juan*, Ms.
- López Morales, Humberto (1978): “Frecuencia léxica, disponibilidad y programación curricular”, en H. López Morales (ed.), *Aportes de la Lingüística a la Enseñanza del Español como Lengua Materna*, edición especial del *Boletín de la Academia Puertorriqueña de la Lengua Española (BAPLE)*, 6: 73–86.
- López Morales, Humberto (1979): “Disponibilidad léxica y estratificación socioeconómica”, en H. López Morales (ed.), *Dialectología y Sociolingüística. Temas puertorriqueños*. Madrid: Hispanova de Ediciones, 173–181.
- López Morales, Humberto (1983): “Lingüística Estadística”, en H. López Morales (coord.), *Introducción a la lingüística actual*. Madrid: Playor, 209–225.
- López Morales, Humberto (1986): *La enseñanza de la lengua materna, Lingüística para maestros de español*. Madrid: Playor, 5<sup>a</sup> ed., 1990.
- López Morales, Humberto (1989): *Sociolingüística*. Madrid: Gredos, 2<sup>a</sup> ed., 1993.
- López Morales, Humberto (1991a): “Muestra de léxico pan-antillano: el cuerpo humano”, en *Investigaciones léxicas sobre el léxico antillano*. Santiago de los Caballeros: Pontificia Universidad Católica y Maestra, 45–80.
- López Morales, Humberto (1991b): “Anglicismos léxicos en el habla culta de San Juan de Puerto Rico”, en *Investigaciones léxicas sobre el léxico antillano*. Santiago de los Caballeros: Pontificia Universidad Católica Madre y Maestra, 125–144.
- López Morales, Humberto (1993): “En torno al aprendizaje del léxico. Bases psicolingüísticas de la planificación curricular”, en S. Montesa y A. Garrido (eds.), *El español como lengua extranjera: de la teoría al aula, Actas del III Congreso Nacional de ASELE*. Málaga: ASELE, 9–22.
- López Morales, Humberto (1995): “Disponibilidad léxica en Puerto Rico. Primera aproximación”, en L. Rodríguez-Fonseca y M. Vázquez (eds.), *Actas del III Seminario Internacional sobre ‘Aportes de la lingüística a la enseñanza de la lengua materna’*. Carolina: Universidad de Puerto Rico.
- López Morales, Humberto (1996): “Los estudios de disponibilidad: pasado y presente”, *Boletín de Filología de la Universidad de Chile*. 35: 245–259.
- López Morales, Humberto (1999a): *Léxico disponible de Puerto Rico*. Madrid: Arco Libros.
- López Morales, Humberto (1999b): “Anglicismos en el léxico disponible de Puerto Rico”, en L. A. Ortiz López (ed.), *El caribe hispánico: perspectivas lingüísticas actuales. Homenaje a Manuel Álvarez Nazario*. Frankfurt–Madrid: Vervuert-Iberoamericana: 147–170.
- López Morales, Humberto (2001): “Tendencias actuales del léxico hispanoamericano”, *Revista de Occidente*, 240: 5–24.

- López Morales, Humberto (2005): “Un nuevo corpus para el estudio del español: la disponibilidad léxica”, *Oralia*, 8: 141–159.
- López Morales, Humberto (2008): “Un alto en el camino: el léxico disponible de Aragón”, en M.<sup>a</sup> L. Arnal Purroy (ed.), *Estudios sobre disponibilidad léxica en los jóvenes aragoneses*. Zaragoza: Institución “Fernando el Católico”, 7–15.
- López Morales, Humberto (2009): “Los estudios de disponibilidad léxica”, en D. Corbella y J. Dorta (eds.), *La investigación dialectológica en la actualidad*. Santa Cruz de Tenerife: Agencia Canaria de Investigación, Innovación y Sociedad de la Información, 23–52.
- Lorán, R. y H. López Morales. (1983): *Nouveau calcul de l'indice de disponibilité*. Ms.
- Lorenzo Ramos, Antonio, Marcial Morera y Gonzalo Ortega (1994): *Diccionario de canarismos*. La Laguna: Francisco Lemus.
- Mackey, William C. (1965): “Bilingual Interference. Its Analysis and Measurement”, *Journal of Communication*, 15/4: 239–249.
- Mackey, William C. (1971a): *Le vocabulaire disponible du français*. Paris, Didier, 2 vols.
- Mackey, William C. (1971b): *Le sondage dans les enquêtes de disponibilité*. Québec: CIRB.
- Manjón-Cabeza Cruz, Antonio (2008a): “Determinación de redes semánticas naturales por medio de encuestas de disponibilidad léxica”, en A. Moreno Sandoval (coord.), *El valor de la diversidad (meta)lingüística: Actas del VIII congreso de Lingüística General*. Edición en CD. Madrid: Universidad Autónoma de Madrid.
- Manjón-Cabeza Cruz, Antonio (2008b): “Redes semánticas naturales en escolares de 5 a 16 años: los colores”, *Docencia e investigación, Revista de la Escuela Universitaria de Magisterio de Toledo*, 18: 1–12.
- Martínez Olmos, Esther (2007): *Disponibilidad léxica en las comunidades de habla alicantinas*. Tesis doctoral inédita. Alicante: Universidad.
- Mateo Garda, M.<sup>a</sup> Victoria (1995): “Disponibilidad léxica: posibles aplicaciones”, en J. D. Luque Durán y A. Pamies Bertrán (eds.), *Primeras Jornadas sobre el estudio y la enseñanza de] léxico*. Granada: Método Ediciones, 141–148.
- Mateo García, M.<sup>a</sup> Victoria (1996): *Estratificación social del léxico de la alimentación en Motril (Granada)*. Almería: GRUSTA.
- Mateo García, M.<sup>a</sup> Victoria (1998): *Disponibilidad léxica en el COU almeriense. Estudio de estratificación social*. Almería: Universidad.
- Mena Osorio, Mónica (1986): *Disponibilidad léxica infantil en tres niveles de enseñanza básica*. Tesina inédita. Concepción: Universidad de Concepción.
- Mesa Canales, Rosa M.<sup>a</sup> (1989): *Disponibilidad léxica en preescolares*. Tesina inédita. México: UNAM.
- Michéa, René (1950): “Vocabulaire et culture”, *Les langues modernes*, 44: 188–189.
- Michéa, René (1953): “Mots fréquents et mots disponibles. Un aspect nouveau de la statistique du langage”, *Les langues modernes*, 47: 338–344.
- Minsky, Marvin (1975): “A framework for representing knowledge”, P. H. Winston (ed.), *The Psychology of Computer Vision*. New York: McGraw Hill, 211–278.
- Moliné Juste, Ana B. (2008): “Creación léxica en el vocabulario disponible de los jóvenes aragoneses: aspectos morfológicos”, en M.<sup>a</sup> L. Arnal Purroy (ed.), *Estudios sobre disponibilidad léxica en los jóvenes aragoneses*. Zaragoza: Institución “Fernando el Católico”, 163–194.
- Moliner, María (1998): *Diccionario de uso del español*. 2 vols. Madrid: Gredos.
- Morales, Amparo (1986): *Léxico básico del español de Puerto Rico*. San Juan de Puerto Rico: Academia Puertorriqueña de la Lengua Española.
- Moreno Fernández, Francisco (2012): “Disponibilidad léxica: cuestiones metodológicas. A propósito de disponibilidad léxica de los estudiantes hispanos de Redwood City, CA”. *Revista Nebrija de Lingüística Aplicada*. 11/6: 51–68.
- Moreno Fernández, Francisco *et al.* (1995): “Cálculo de disponibilidad léxica. El programa Lexi-disp”, *Lingüística*. 7: 243–249.

- Moreno Fernández, Francisco (2004): “Los estudios dialectales sobre el español de España (1979–2004)”, *Lingüística Española Actual*, 26/2: 65–100.
- Müller, Charles (1964): *Essai de statistique lexicale*. París: Klincksieck.
- Müller, Charles (1973): *Estadística lingüística*. (Antonio Quilis, trad.). Madrid: Gredos. (Obra original: 1968. *Initiation a la statistique linguistique*. París: Librairie Larousse).
- Murillo Rojas, Marielos (1993): “Disponibilidad léxica en los preescolares: estudio de cinco campos semánticos”, en *Káñina, Revista de Artes y Letras*, Universidad de Costa Rica, 17/2: 117–127.
- Murillo Rojas, Marielos (1994): “Comidas y bebidas: estudio de disponibilidad léxica en preescolares”, *Káñina, Revista de Letras y Artes*, Universidad de Costa Rica, 18/2: 117–133.
- Murillo Rojas, Marielos (1998): “Crecimiento de la disponibilidad léxica: niños de preescolar y primer ciclo de la educación básica costarricense”, *Revista de Filología y Lingüística de la Universidad de Costa Rica*, 25/2: 187–203.
- Najt Rajtlemsky, Myriam y M.<sup>a</sup> Victoria Reizábal Rodríguez (1989): “La adquisición del vocabulario”, en J. García Padrino y A. Medina (dirs.), *Didáctica de la lengua y la literatura*. Madrid: Anaya, 212–232.
- Nattinger, James R. y Jeanette S. De Carrico (1992): *Lexical Phrases and Language Teaching*. Oxford: Oxford University Press.
- Nelson, Douglas L., David J. Bennett y Todd W. Leibert (1997): “One step is not enough: making better use of association norms to predict cued recall”, *Memory and Cognition*, 25: 785–706.
- Ogden, Charles K. (1944): *Basic English. A general Introduction with Rules and Grammar*. Londres: Routledge and Kegan Paul.
- Ortolano Ríos, Bárbara (2005): “Disponibilidad léxica de los estudiantes Ayamonte”, *Interlingüística*, 16/2: 847–857.
- Paredes García, Florentino (1999): “La ortografía en las encuestas de disponibilidad léxica”. *Revista de Estudios de Adquisición de la Lengua Española (REALE)*, 11: 75–97.
- Paredes García, Florentino (2001a): “Disponibilidad léxica en alumnos de Alcalá y su comarca: resultados generales”, en I. de la Cruz *et al.* (eds.), *La lingüística aplicada a finales del siglo XX. Ensayos y propuestas. Actas del IV Congreso de AESLA*. Alcalá de Henares: Universidad, 2, 721–728.
- Paredes García, Florentino (2001b): “Disponibilidad de los extranjerismos en estudiantes de Educación Secundaria”, en M.<sup>a</sup> A. Zorraquino y C. Díez Pelegrín (eds), *¿Qué español enseñar? Norma y variación lingüísticas en la enseñanza del español a extranjeros. Actas del XI Congreso Internacional de ASELE*. Zaragoza: ASELE/Universidad de Zaragoza, 567–576.
- Paredes García, Florentino (2005): “El campo léxico de los colores: convergencias y divergencias en grupos sociales de Madrid”, en L. Rodríguez Alfano (ed.), *Memorias del XIV Congreso Internacional ALFAL*. CD-ROM. Monterrey: Universidad Autónoma de Nuevo León, vol. 1, 131–141.
- Paredes García, Florentino (2012): “Desarrollos teóricos y metodológicos recientes de los estudios de disponibilidad léxica”, *Revista Nebrija de Lingüística Aplicada*. 11/6: 78–100.
- Paredes García, Florentino, Luis Guerra, y Elena Gómez (en prensa): *Léxico disponible de los estudiantes preuniversitarios madrileños*. Madrid: Universidad de Alcalá – Universidad Europea de Madrid.
- Pastor Milán, M.<sup>a</sup> Ángeles y Francisco J. Sánchez García (2008): *El léxico disponible de Granada y su provincia*. Granada: Universidad.
- Postigo, Rosa M.<sup>a</sup> (1998). “Modelos de educación lingüística en contextos de lenguas en contacto”, en A. Mendoza Fillola (coord.), *Conceptos clave en Didáctica de la Lengua y la Literatura*. Barcelona: Universitat de Barcelona, Institut de Ciències de l'Educació, ICE, 265–274.
- Prado Aragonés, Josefina (2008a): “La disponibilidad léxica en las zonas limítrofes de Andalucía y Extremadura (España) y Algarve y Alentejo (Portugal): aspectos metodológicos y primeros resultados”, en D. Azorín Fernández (dir.), *El diccionario como puente entre las lenguas y cul-*

- turas del mundo. Actas del II Congreso Internacional de Lexicografía Hispánica*. Alicante: Biblioteca Virtual Miguel de Cervantes, 568–574.
- Prado Aragonés, Josefina (2008b): “Incidencia de la variable social ‘Tipo de centro’ en la disponibilidad léxica de Huelva: Aspectos cuantitativos”, en M.<sup>a</sup> V. Galloso Camacho (ed.), *Estudios en torno a la lengua y la literatura en Andalucía*. Granada: Grupo Editorial Universitario, 147–164.
- Prado Aragonés, Josefina (2009): “Dialectalismos en el léxico disponible de Huelva en relación con el centro de interés *Alimentos y bebidas*”, en M.<sup>a</sup> V. Camacho Taboada *et al.* (eds.), *Estudios de lengua española. Descripción, variación y uso. Homenaje a Humberto López Morales*. Madrid–Frankfurt: Iberoamericana–Vervuert, 567–594.
- Prado Aragonés, Josefina y M.<sup>a</sup> Victoria Galloso Camacho (2005): *Léxico disponible de Huelva. Nivel preuniversitario*. Huelva: Universidad.
- Prado Aragonés, Josefina y M.<sup>a</sup> Victoria Galloso Camacho (2006a): *Diccionario del léxico disponible de alumnos de 6º curso de Educación Primaria de Huelva*. Huelva: Universidad. Grupo de Investigación “Lingüística andaluza”.
- Prado Aragonés, Josefina y M.<sup>a</sup> Victoria Galloso Camacho (2006b): *Diccionario del léxico disponible de alumnos de 2º curso de Bachillerato de Huelva*. Huelva: Universidad. Grupo de Investigación “Lingüística andaluza”.
- Prado Aragonés, Josefina, M.<sup>a</sup> Victoria Galloso Camacho y Manuel Célio Conceição (2009): *La disponibilidad léxica en situación de contacto de lenguas en las zonas limítrofes de Andalucía y Extremadura (España) y Algarve y Alentejo (Portugal)*. Huelva: Universidad.
- Real Academia Española (2001): *Diccionario de la lengua española*, 22.<sup>a</sup> ed. Madrid: Espasa Calpe.
- Richards, I. A. (1968): *Basic English. International Second Language. A rev. and expanded Version of Ogden: The System of Basic English*. Nueva York: Harcourt, Brace & World.
- Ríos González, Gabriela (2007): “Diferencias léxicas entre el hombre y la mujer en tres centros de interés: Saludos, Temas de conversación y Despedidas”, *Filología y Lingüística*, 33/1: 151–166.
- Ríos González, Gabriela (2010): *Características del lenguaje de los jóvenes costarricenses desde la disponibilidad léxica*. Tesis doctoral inédita. Salamanca: Universidad.
- Ríos González, Gabriela (2010): “Los saludos de los jóvenes costarricenses”, *Káñina, Revista de Artes y Letras de la Universidad de Costa Rica*. 34/2: 151–180.
- Ríos González, Gabriela (2011): “El tabú y el disfemismo en jóvenes colegiales costarricenses”, *Káñina, Revista de Artes y Letras de la Universidad de Costa Rica*. 35/1: 153–162.
- Rodríguez Bou, Ismael (1952): *Recuento de vocabulario español*. Río Piedras: Universidad de Puerto Rico.
- Romaine, Suzanne (1996): *El lenguaje en la sociedad*. (J. Borrego Nieto, trad.). Barcelona: Ariel. (Original: 1994. *Language in Society*. Oxford: Oxford University Press.)
- Román-Morales, Belén (1985): *Disponibilidad léxica en escolares de Dorado*. Tesina inédita. Río Piedras: Universidad de Puerto Rico.
- Romero Rubilar, Carlos (2000): “Variabilidad y prototipicidad en el vocabulario disponible”, *Revista de Lingüística Teórica y Aplicada (RLA)*, 38: 139–148.
- Rosch, Eleanor (1978): “Principles of categorization”, en E. Rosch y B. Lloyd (eds.), *Cognition and Categorization*. Hillsdale: Lawrence Erlbaum, 27–48.
- Ruiz Basto, Araceli (1987): *Disponibilidad léxica de los alumnos de primer ingreso en el Colegio de Ciencias y Humanidades, Plantel de Naucalpan*. Memoria de Licenciatura inédita. México: UNAM.
- Samper Hernández, Marta (2002): *Disponibilidad léxica en alumnos de español como lengua extranjera*, Colección Monografías, 4. Málaga: ASELE.
- Samper Hernández, Marta (2003a): *Evolución de la disponibilidad léxica en escolares grancanarios de enseñanza primaria y secundaria*. Tesis doctoral inédita. Salamanca: Universidad.
- Samper Hernández, Marta (2003b): “Datos comparativos entre el léxico disponible de estudiantes de español como lengua extranjera y el de la variedad grancanaria”, en C. Díaz Alayón *et al.*

- (coords.), *Estudios sobre el español de Canarias*, Islas Canarias: Academia Canaria de la lengua, 961–987.
- Samper Hernández, Marta (2005a): “La disponibilidad léxica de escolares grancanarios: resultados generales de la investigación”, en G. Santana Henríquez *et al.* (eds.), *Estudios en Homenaje a María del Prado Escobar Bonilla*. Las Palmas de Gran Canaria: Universidad, 577–587.
- Samper Hernández, Marta (2005b): “Dialectalismos en el léxico disponible de escolares grancanarios”, en L. Santos Río *et al.* (eds.), *Palabras, norma, discurso. En memoria de Fernando Lázaro Carreter*. Salamanca: Universidad, 1065–1077.
- Samper Hernández, Marta (2005c): “Incidencia del factor ‘nivel sociocultural’ en la disponibilidad léxica de escolares grancanarios”, *Interlingüística*, 15: 1243–1251.
- Samper Hernández, Marta (2008): “Datos comparativos entre dos léxicos disponibles: Aragón y Gran Canaria”, en M.<sup>a</sup> L. Arnal Purroy (ed.), *Estudios sobre disponibilidad léxica en los jóvenes aragoneses*. Zaragoza: Institución “Fernando el Católico”, 245–282.
- Samper Hernández, Marta (2009): *Evolución de la disponibilidad léxica en estudiantes grancanarios de enseñanza primaria y secundaria*. Las Palmas de Gran Canaria: Cabildo de Gran Canaria.
- Samper Hernández, Marta (2012): “Anglicismos en el léxico disponible de escolares grancanarios de enseñanza primaria y secundaria”, en J. A. Bartol Hernández y J. F. García Santos (eds.), *Estudios de filología española*. Salamanca: Luso-Española de Ediciones y Universidad de Salamanca.
- Samper Padilla, José A. (1998a): “Criterios de edición del léxico disponible”, *Lingüística*, 10: 311–333.
- Samper Padilla, José A. (dir.) (1998b): *Léxico del habla culta de Las Palmas de Gran Canaria*. Las Palmas de Gran Canaria: Cabildo Insular de Gran Canaria y Universidad de Las Palmas de Gran Canaria.
- Samper Padilla, José A. (1999): “Léxico disponible y variación dialectal: datos de Puerto Rico y Gran Canaria”, en A. Morales *et al.* (eds.), *Estudios de lingüística hispánica (Homenaje a María Vaquero)*. San Juan de Puerto Rico: Universidad de Puerto Rico, 550–573.
- Samper Padilla, José A. (2000): “El léxico de la norma popular de Las Palmas de Gran Canaria. Comparación con el de la norma culta”, en C. Corrales Zumbado y D. Corbella Díaz (coords.), *Estudios de dialectología dedicados a Manuel Alvar con motivo del XL aniversario de la publicación de El español hablado en Tenerife*. La Laguna: Instituto de Estudios Canarios, 39–74.
- Samper Padilla, José A. (2006): “Disponibilidad léxica y sociolingüística”, en J. L. Blas Arroyo *et al.* (eds.), *Discurso y Sociedad. Contribuciones al estudio de la lengua en contexto social*. Castellón de la Plana: Universitat Jaume I, 99–120.
- Samper Padilla, José A. y Clara E. Hernández Cabrera (1995a): “Disponibilidad léxica en Gran Canaria. Primeras estimaciones”, en *Jornadas Internacionales de Lingüística Aplicada*, Ms. Granada: Universidad.
- Samper Padilla, José A. y Clara E. Hernández Cabrera (1995b): “Disponibilidad léxica en Las Palmas: análisis cuantitativo”, en L. Rodríguez-Fonseca e I. N. Vázquez (eds.), *Actas del III Seminario Internacional sobre ‘Aportes de la lingüística a la enseñanza de la lengua materna’*, números especiales de la *Revista de Adquisición de la Lengua Española*. Carolina: Universidad de Puerto Rico.
- Samper Padilla, José A. y Clara E. Hernández Cabrera (1997): “Estudio de la disponibilidad léxica en Gran Canaria. Datos iniciales y variación sociolingüística”, en M. Almeida y J. Dorta (eds.), *Contribuciones al estudio de la lingüística hispánica (Homenaje a Ramón Trujillo Carreño)*. La Laguna: Universidad, vol. 2. 229–239.
- Samper Padilla, José A. y Marta Samper Hernández (2006): “Aportaciones recientes de los estudios de disponibilidad léxica”, *LynX: Panorámica de Estudios Lingüísticos*, 5: 5–95.

- Samper Padilla, José A. y Marta Samper Hernández (2009): “Estudios sobre anglicismos en el Proyecto Panhispánico de léxico disponible”, en M. Veyrat Rigat y E. Serra Alegre (eds.), *La Lingüística como reto epistemológico y como acción social. Estudios dedicados al profesor Ángel López García con ocasión de su sexagésimo aniversario*, vol. 2. Madrid: Arco Libros, 869–880.
- Samper Padilla, José A., Juan J. Bellón Fernández, y Marta Samper Hernández (2003): “El proyecto de estudio de la disponibilidad léxica en español”, en Gerd Wotjak (coord.), *Pautas y pistas en el estudio del léxico hispano (americano)*, Frankfurt–Madrid: Vervuert-Iberoamericana, 27–140.
- Samper Padilla, José A., Clara E. Hernández Cabrera y Juan J. Bellón Fernández (2003): “Comparación de los léxicos disponibles de Gran Canaria y Córdoba”, en C. Díaz Alayón *et al.* (eds.), *Estudios sobre el español de Canarias. Actas del I Congreso Internacional sobre el español de Canarias*. Islas Canarias: Academia Canaria de la Lengua, vol. 2, 1077–1099.
- Sánchez Corrales, Víctor M. y Marielos Murillo Rojas (1993): “Campos semánticos y disponibilidad léxica en preescolares”, *Revista de Educación* 17/2: 15–25.
- Sánchez Corrales, Víctor M. y Marielos Murillo Rojas (1997): “Léxico básico de los niños preescolares de la sub-región educativa de San José, Costa Rica”, *Kañina, Revista de Artes y Letras de la Universidad de Costa Rica*, 22/2: 77–124.
- Sánchez Corrales, Víctor M. y Marielos Murillo Rojas (1999): “Inventario de léxico disponible de los preescolares costarricenses”, *Kañina, Revista de Artes y Letras de la Universidad de Costa Rica*, 24/1: 103–126.
- Sánchez Corrales, Víctor y Marielos Murillo Rojas (2002): *Léxico básico de los niños preescolares costarricenses*. Costa Rica: Universidad de Costa Rica.
- Sánchez Corrales, Víctor M. y Marielos Murillo Rojas (2004): “Disponibilidad léxica y educación lingüística en niños preescolares costarricenses”, en J. Prado Aragonés y M.<sup>a</sup> V. Galloso Camacho (eds.), *Diccionario, Léxico y Cultura*. Huelva: Universidad, 215–232.
- Sánchez Corrales, Víctor M. y Marielos Murillo Rojas (2006): *Disponibilidad léxica de los niños preescolares costarricenses*. Costa Rica: Universidad de Costa Rica.
- Sánchez Gómez, Coronada (2005): “Naturaleza gramatical del léxico disponible en informantes de español como lengua extranjera”, *Interlingüística*, 16/2: 977–986.
- Sánchez-Saus Laserna, Marta (2011): *Bases semánticas para el estudio de los centros de interés del léxico disponible. Disponibilidad léxica de informantes extranjeros en las universidades andaluzas*. Tesis doctoral inédita. Cádiz: Universidad.
- Santiago Guervós, Javier de (2008): *El léxico disponible de Segovia: estudio y diccionarios*. Burgos: Fundación Instituto Castellano y Leonés de la Lengua.
- Santos Maldonado, M.<sup>a</sup> Jesús (2002): *El error en las producciones escritas de francés lengua extranjera: análisis de interferencias léxicas y propuestas para su tratamiento didáctico*. Tesis doctoral inédita. Escuela Universitaria de Educación de Palencia, Universidad de Valladolid.
- Saralegui, Carmen y Cristina Taberner (2002): *Navarrismos en el Diccionario de la Real Academia Española*. Pamplona: Institución Príncipe de Viana, 2<sup>a</sup> ed. aumentada.
- Saralegui, Carmen y Cristina Taberner (2008): “Aportación al Proyecto Panhispánico de Léxico Disponible: Navarra”, en I. Olza Moreno *et al.* (eds.), *Actas del XXXVII Simposio Internacional de la Sociedad Española de Lingüística (SEL)*. Pamplona: Universidad de Navarra, 745–761.
- Saura Ramí, José A. (2008): “La ortografía en las encuestas aragonesas de disponibilidad léxica”, en M.<sup>a</sup> L. Arnal Purroy (ed.), *Estudios sobre disponibilidad léxica en los jóvenes aragoneses*. Zaragoza: Institución “Fernando el Católico”, 195–205.
- Seco, Manuel, Olimpia Andrés y Gabino Ramos (1999): *Diccionario del español actual*. Madrid: Aguilar.
- Serrano Zapata, Maribel (2002): *Análisis sociolingüístico del léxico castellano disponible en la ciudad de Lleida*. Lleida: Universitat de Lleida.

- Serrano Zapata, Maribel (2003): "Disponibilidad léxica en la provincia de Lleida: aspectos metodológicos", *Interlingüística*, 14: 919–927.
- Serrano Zapata, Maribel (2004): "Aspectos sociolingüísticos del léxico disponible castellano de los preuniversitarios leridanos", *Pragmalingüística*, 12: 147–165.
- Serrano Zapata, Maribel (2006): "Consecuencias del contacto de lenguas en Lérida: interferencias detectadas en las encuestas de disponibilidad léxica", en J. L. Blas Arroyo *et al.* (eds.), *Discurso y Sociedad. Contribuciones al estudio de la lengua en contexto social*. Castellón de la Plana: Universitat Jaume I, 811–829.
- Siguán, Miguel y William F. Mackey (1986): *Educación y bilingüismo*. Madrid: Santillana.
- Taberero Sala, Cristina (2008a): "Disponibilidad léxica y contacto de lenguas", *Oihenart. Cuadernos de Lengua y Literatura*, 23: 545–565.
- Taberero Sala, Cristina (2008b): "El léxico disponible como fuente de aproximación al estudio de los regionalismos", en I. Olza Moreno *et al.* (eds.), *Actas del XXXVII Simposio Internacional de la Sociedad Española de Lingüística*. Pamplona: Universidad de Navarra, 811–824.
- Thorndike, Edward L. (1921): *Teacher's Word Book*. Nueva York: MacMillan.
- Tomé Cornejo, Carmela (2011): "Reflexiones en torno a los centros de interés", en M. V. Galloso Camacho (coord.), *La investigación de la Lengua y la Literatura en la Onubense*. Huelva: Universidad, 119–159.
- Torres González, Antonia N. y Ana B. García Falero (2009): "Planificación y primeros resultados de la disponibilidad léxica en Tenerife", en D. Corbella y J. Dorta (eds.), *La investigación dialectológica en la actualidad*. Santa Cruz de Tenerife: Agencia Canaria de Investigación, Innovación y Sociedad de la Información, 75–101.
- Trigo Ibáñez, Ester (2007): *El léxico disponible de la provincia de Sevilla: variación versus déficit*. Tesis doctoral inédita. Cádiz: Universidad.
- Trigo Ibáñez, Ester (2011): *Dialectología y cultura. El léxico disponible de los universitarios sevillanos*. Valencia: Aduana Vieja.
- Trigo Ibáñez, Ester y Adolfo E. González Martínez (2011): "Estudio del comportamiento de la variable *sexo* en el léxico disponible de los preuniversitarios sevillanos". *Diálogo de la Lengua*, 3: 28–41.
- Urrutia Martínez, Mabel A. (2001): *Edad y acceso léxico*. Tesis de magister. Universidad de la Concepción.
- Valencia Espinoza, Alba (1993): "El vocabulario infantil y la medición de las habilidades intelectuales", en *Actas del X Congreso de la ALFAL*. Veracruz: México.
- Valencia Espinoza, Alba (1994a): *El léxico de los estudiantes de 4º año de Educación Media. Centro de interés Procesos mentales*. Santiago de Chile, Serie documentos de estudio 26, CPEIP.
- Valencia Espinoza, Alba (1994b): "Disponibilidad léxica en Educación Media: III y IV", en *X Encuentro en torno a la Admisión Universitaria*. Santiago de Chile: Talleres gráficos DIVEST.
- Valencia Espinoza, Alba (1997): "Disponibilidad léxica. Muestreo y estadísticos", *Onomazein*, 2: 197–226.
- Valencia Espinoza, Alba (1998a): "Aspectos sociolingüísticos del léxico estudiantil de la Región Metropolitana", en M. Bernales y C. Contreras (eds.), *Por los Caminos del Lenguaje*. Temuco: Universidad de La Frontera, 39–45.
- Valencia Espinoza, Alba (1998b): "Léxico estudiantil. Comentario sobre un repertorio", *Boletín de Filología, Homenaje a Ambrosio Rabanales*, 37: 1211–1221.
- Valencia Espinoza, Alba (2000): "Ciencia y tecnología. Análisis del léxico estudiantil", *Nueva Revista del Pacífico*, 45: 89–98.
- Valencia Espinoza, Alba (2005): "Dialectalismos en el léxico disponible chileno", Comunicación presentada al *XIV Congreso Internacional de la ALFAL* (Monterrey, México), Ms.
- Valencia Espinoza, Alba y Max S. Echeverría (1998): *Disponibilidad léxica en estudiantes de cuarto año de Educación Media*. Santiago de Chile: Universidad de Chile y Universidad de La Concepción.

- Valencia Espinoza, Alba y Max. S. Echeverría (1999a): *Disponibilidad léxica en estudiantes chilenos*. Santiago de Chile: Universidad de Chile y Universidad de la Concepción.
- Valencia Espinoza, Alba y Max S. Echeverría (1999b): “El factor geográfico en la disponibilidad léxica chilena”, en J. A. Samper Padilla *et al.* (eds.), *Actas del XI Congreso Internacional de la Asociación de Lingüística y Filología de América Latina*. Las Palmas de Gran Canaria: Universidad de Las Palmas de Gran Canaria y Librería Nogal, Tomo 2: 1705–1711.
- Vander Beke, George E. (1935). *French Word Book*. Nueva York: MacMillan.
- Vargas Sandoval, Patricia (1991): “La disponibilidad léxica: un diagnóstico al hablante medio de la V Región en seis centros de interés”, *Nueva Revista del Pacífico*. Valparaiso: Universidad de Playa Ancha, 33–36: 115–123.
- Vega, Manuel de (1984): *Introducción a la Psicología Cognitiva*. Madrid: Alianza Psicología.
- Verlée, Léon (1963): *Basis Woordenboek voor de Franse Taal*. Amberes: De Sikkel.
- Wingeyer, Hugo R. (2007): *Léxico disponible de Resistencia*. Tesis doctoral inédita. Alcalá de Henares: Universidad.



## DOSTĘPNOŚĆ LEKSYKALNA. TEORIA, METODA I ANALIZA. STRESZCZENIE

Niniejsza monografia przedstawia krytyczną analizę podstaw teoretycznych i metodologicznych dla badań dostępnej leksyki w języku hiszpańskim. Na podstawie szerokiego przeglądu bibliograficznego dokonano analizy najbardziej znaczących danych w ramach Projektu Panhispanistycznego, przedstawiając różne obszary lingwistyczne, w ramach których są prowadzone niniejsze badania. W tym kontekście niniejsza monografia ma na celu uzupełnienie i uaktualnienie poprzednich prac takich autorów, jak Samper Padilla *et al.* (2003), czy Samper Padilla i Samper Hernández (2006), które przedstawiały szeroką panoramę osiągnięć w dziedzinie dostępności leksykalnej w języku hiszpańskim.

Jak wiadomo, dostępność leksykalna narodziła się w związku z celami pedagogicznymi we Francji (Gougenheim *et al.*, 1956), a konkretniej w celu zebrania powszechnie stosowanego słownictwa tematycznego, które należało włączyć do podstawowego słownictwa w ramach nauczania. Te same cele zainspirowały Humberta Lópeza Moralesa (1973) do przeprowadzenia pierwszego zbadania dostępności leksykalnej w obszarze hiszpańskojęzycznym, w ramach którego analizuje on akwizycję słownictwa i różnych zmiennych występujących w jego zakresie, na podstawie próbki języka ośrodków edukacyjnych z obszaru miejskiego w San Juan w Portoryko.

Perspektywa dydaktyczna ma nadrzędne znaczenie dla opracowania tego rodzaju badań, jednak w niniejszej pracy tymi rodzajami zastosowań zajmujemy się jedynie drugorzędnie, jako że celem niniejszej monografii jest analiza podstaw teoretycznych i metodologicznych tych prac. Mamy tu więc do czynienia z badaniem o charakterze epistemologicznym, ponieważ wnosi ono wgląd w podstawy, strukturę, metody i samą zasadność badania w zakresie dostępnej leksyki. W tym celu niniejsza praca zgłębiła wstępne badania w ramach prac wymienionych w niniejszej monografii, porównując podstawy teoretyczne dla tychże prac, metodologię zastosowaną w każdej z nich oraz wyniki uzyskane w wyniku jej zastosowania.

Strukturalnie niniejsza monografia składa się z siedmiu rozdziałów, *Wstępu*, zawierającego zakres badania i jego podstawowe tezy, *Wniosków końcowych* (konkluzji) oraz *Bibliografii*. Rozdział I (*Badania nad dostępnością leksykalną*), z jednej strony, poświęcony jest historycznej ewolucji w zakresie głównych procedur pomiaru istniejącego słownictwa w społeczności języka mówionego. Wśród tych procedur dostępność leksykalna dotyczy słownictwa o bardzo konkretnej zawartości semantycznej, które zawiera się w mentalnym leksykonie, ale ma charakter potencjalny. Z drugiej strony, rozdział I pokazuje, w jaki sposób język hiszpański najlepiej wykorzystał badania ilościowe leksyki w oparciu o dostępność leksykalną. Po rozpoczęciu wstępnych prac w 1973 r. Humberto López Morales nadal badał różne aspekty deskryptywne i metodologiczne dostępności leksykalnej, przy czym jego badania osiągnęły fazę kulminacyjną w momencie publikacji pracy zatytułowanej *Dostępna leksyka Portoryko* (López Morales, 1999a). W okresie pomiędzy dwiema datami wspomnianymi powyżej, badania dostępnej leksyki zaczęły być prowadzone w różnych kra-

jach Ameryki Łacińskiej (w Portoryko, w Republice Dominikany, w Kostaryce, w Meksyku, w Chile), a w Hiszpanii po raz pierwszy przeprowadzili je Benítez Pérez (1992) w Madrycie oraz Samper Padilla i Hernández Cabrera (1992) w Las Palmas de Gran Canaria. Od tej pory badania dostępnej leksyki rozpowszechniły się w całym obszarze hiszpańskojęzycznym, przede wszystkim dzięki ambitnemu „Projektowi Badań nad Dostępnością Leksykalną w Języku Hiszpańskim”, prowadzonemu przez profesora Lópeza Moralesa.

Rozdziały II (*Podstawy teoretyczne i metodologie dostępności leksykalnej*) i III (*Ośrodki zainteresowań*) poruszają kwestię odnowy epistemologicznej i metodologicznej, której doświadczyły badania w zakresie dostępnej leksyki w ramach Projektu Panhispanistycznego, poświęcone zbadaniu i zdefiniowaniu normy w zakresie dorosłego społeczeństwa w odniesieniu do różnych hiszpańskojęzycznych społeczności. Tym niemniej, dyskusja metodologiczna nie była wolna od kontrowersji, np. kontrowersji w zakresie opracowania wzoru matematycznego, stosowanego w badaniach statystycznych dostępności leksykalnej (Lórán i López Morales, 1983; López Chávez i Strassburger, 1987), lub kontrowersji w postaci różnych propozycji na temat ilości i charakteru pól leksykalnych, które powinny zostać zbadane. Nawet same kryteria redakcyjne w ramach Projektu Panhispanistycznego (Samper Padilla, 1998a: 311–313), nadające temu projektowi jednolity charakter, były postrzegane jako podlegające różnym zmianom i adaptacjom.

Nieodłączną i konstytutywną częścią badań nad dostępną leksyką jest zastosowanie narzędzi i procedur analizy statystycznej, które obejmują statystykę opisową oraz zastosowanie specjalistycznych wzorów (współczynnika dostępności, współczynnika spójności, analizy wieloczynnikowej itp.). Jak twierdzi Mateo García (1998: 14), „w pewne zdziwienie wprawia analiza, zawierająca w ramach liczb, współczynników i wielkości socjolingwistycznych tak znaną i zwyczajną rzeczywistość, jak ta w postaci dostępności leksykalnej”. W każdym razie ten rodzaj naukowej obróbki danych służył temu, aby wykryć te cechy statystyczne, które są mniej więcej stałe na przestrzeni badań wykonanych w obrębie różnych społeczności, mających wpływ na produkcję leksykalną w ośrodkach zainteresowań, jak również na zachowanie różnych badanych zmiennych socjolingwistycznych. Te aspekty są analizowane w rozdziale III wspomnianym powyżej i w rozdziale V (*Pozalingwistyczne uwarunkowania dostępności leksykalnej*), na podstawie porównania wyników badań hiszpańskich w ramach Projektu Panhispanistycznego.

Rozdział IV (*Powiązania między słowami w ramach dostępnej leksyki*) jest zainspirowany badaniami Hernándezza Muñoza (2006) oraz Ávili Muñoza i Villena Ponsody (2010), a zgłębia kierunki badań zainicjowane przez tych badaczy, zmierzające w kierunku wykrycia struktury leksykonu mentalnego na podstawie danych pochodzących z badań dostępności, oraz z wykorzystaniem modeli psycholingwistycznych mających największe zastosowanie do produkcji leksyki. Rozdział ten pokazuje, że analiza powiązań pomiędzy słowami, np. wewnątrz listy słownictwa, stosowanego przez pojedynczą badaną osobę, lub wewnątrz zbioru leksykalnego danej grupy, jest niezwykle użytecznym narzędziem, pozwalającym na wykrycie rzeczywistości psychologicznej dostępnej leksyki. Badania, takie jak prowadzone przez Hernándezza Muñoza (2006), które postrzegają dostępną leksykę, jako złożone zjawisko poznawcze wnoszą najbardziej znaczący wkład w ramach osiągnięć zintegrowanej teorii dostępności leksykalnej, promowanej przez Lópeza Moralesa (1999a: 26), w obrębie której dochodzą do głosu czynniki, takie jak: pamięć semantyczna, znajomość itp.

Rozdział V (*Uwarunkowania pozalingwistyczne dostępności leksykalnej*) analizuje występowanie czynników społecznych w produkcji leksykalnej. Jedną z głównych cech Projektu Panhispanistycznego dotyczącego Dostępności Leksykalnej jest przyjęcie perspektywy socjolingwistycznej o zróżnicowanym nacechowaniu, która próbuje ukazać społeczną stratyfikację leksykalną w odniesieniu do dostępności leksykalnej badanej społeczności. Zgodnie z ogólnymi zasadami Projektu Panhispanistycznego, obowiązkowe zmienne społeczne w ramach każdego badania to: ‘płeć’, (mężczyzna/kobieta), ‘ośrodek edukacyjny’ (publiczny/prywatny), ‘strefa geograficzna’ (miejska/wiejska) i ‘poziom społeczno-kulturowy’ (wysoki/średni/niski). Rozdział ten zajmuje się również analizą ilościową tych zmiennych w ramach różnych badań. W obecnym opracowaniu, opierając

się na danych pochodzących ze wspomnianych badań, wskazano tendencje, którymi cechują się różne badania w odniesieniu do każdej zmiennej, a wśród których można zaobserwować wyraźny wpływ poziomu społeczno-kulturowego na zróżnicowanie leksykalne w ramach danej społeczności. Oprócz tego analizie poddano inne zmienne, takie jak „wiek”, wraz z uwarunkowaniami pozalingwistycznymi, które muszą być uwzględnione w sytuacjach, w których dochodzi do kontaktów językowych.

Rozdział VI (*Analiza jakościowa dostępnej leksyki*) bada szereg czynników lingwistycznych, które są kluczowe dla struktury wykazów dostępnego słownictwa, zawartych w różnych słownikach, będących owocem zróżnicowanych badań. Zgodnie z tym, co twierdzi García Marcos (1997: 25), oprócz czynników społecznych należy również uwzględnić w ramach badań dostępnego słownictwa „ściśle lingwistyczne fakty, które nie tylko są – lub mogą być – odpowiedzialne za konfigurację, osiągającą określony poziom dostępnego słownictwa, ale za te, które leżą u podstawy dynamiki lingwistycznej, a co za tym idzie, również socjolingwistycznej określonego społeczeństwa posługującego się językiem”. Charakter analizowanych faktów lingwistycznych pozwala na zastosowanie różnych klasyfikacji słów wymienionych w ramach wykazów, odzwierciedlających zróżnicowanie (diatopiczne, diastatyczne, diafazywne), zapożyczenia leksykalne, morfologię i słowotwórstwo itp.

Na koniec, przed podaniem wniosków końcowych rozdział VII (*Porównanie dostępnych leksykonów*) analizuje metodologię stosowaną w przypadku porównań międzydialektycznych. Rozdział ten przedstawia również porównanie pięciu rodzajów leksyki dostępnych w Hiszpanii – co jest ostatecznym celem Projektu Panhispanistycznego Dostępności Leksykalnej; innymi słowy jest to próba ustanowienia normy leksykalnej dla języka hiszpańskiego w Hiszpanii. Zadanie to może również posłużyć nam do ukazania wspólnych cech i różnic pomiędzy dostępnymi rodzajami słownictwa, oraz do zilustrowania prawdopodobnego końcowego etapu wspomnianego makroprojektu: opracowania słowników zawierających dostępną leksykę, uwzględniających wyniki wcześniejszych badań.



## LEXICAL AVAILABILITY: THEORY, METHOD AND ANALYSIS. SUMMARY

This monograph presents a critical analysis of the theoretical and methodological bases of the studies of available lexicon in Spanish. It analyses the most significant contributions of the Pan-Hispanic Project on the basis of a broad bibliographical revision, and it presents different linguistic areas in which these studies have developed. In this respect, it attempts to complete and update the previous works by Samper Padilla *et al.* (2003), and by Samper Padilla and Samper Hernández (2006), which presented a broad panorama of the achievements of lexical availability in the Spanish language.

As it is already known, lexical availability originated in relation to pedagogical objectives in France (Gougenheim *et al.*, 1956), with the view to collecting thematic vocabulary of common use which was to be included in the fundamental lexis to be taught. These very objectives inspired Humberto López Morales (1973) to conduct the first investigation of lexical availability in the Spanish-speaking world, in which he analysed the acquisition of vocabulary and the variables present within it on the basis of a sample from scholarly centres in the metropolitan zone of San Juan in Puerto Rico.

Didactic perspective is of maximum importance for the development of these studies, nevertheless this time we have only briefly mentioned the applications of this type, since the objective of this monograph is the analysis of theoretical and methodological foundations of these works. Therefore we have to do with the investigation of epistemological nature, which means that this study results in the contribution informing about the origin, the structure, the methods and the validity of the study of available lexicon itself. To this end, this work has delved into the introductory studies related to the works mentioned herein, and it has compared theoretical assumptions from which they originate, the applied methodology that derives from them, and the results obtained thanks to their application.

Structurally, this monograph consists of seven chapters, Introduction, containing a range of research and its basic thesis, Final Conclusions and Bibliography. Chapter I (*The study of lexical availability*) is dedicated, on the one hand, to the historical evolution of the main procedures which serve for the measurement of the existing vocabulary in a society using a given language. Among these procedures, lexical availability refers to vocabulary with a very specific semantic content that exists in the mental lexicon, but which has a potential character. On the other hand, it offers a testimony on how Spanish has benefited most from quantitative investigation of lexicon using lexical availability. After the initial work in 1973, Humberto López Morales continued to investigate different descriptive and methodological aspects of lexical availability, and his investigations reached a peak point upon the publication of *The available lexicon of Puerto Rico* (López Morales, 1999a). Between these two dates, the studies on available lexicon spread across different countries of Latin America (Puerto Rico, the Dominican Republic, Costa Rica, Mexico, Chile), and they reached Spain

in 1992 with the first studies conducted by Benítez Pérez (1992) in Madrid, and by Samper Padilla and Hernández Cabrera (1992) in Las Palmas de Gran Canaria. Since then, the studies of available lexicon have proliferated in the whole geographical region of Spanish-speaking countries, which is mainly due to the existence of an ambitious pan-Hispanic project, i.e. the Project of Study on Lexical Availability in Spanish, co-ordinated by professor López Morales.

Chapters II (*Theoretical and methodological foundations for lexical availability*) and III (*Centres of interest*) treat about epistemological and methodological renovation that the studies of available lexicon, which were dedicated to the investigation and determination of the communitarian adult standard within different Spanish-speaking communities, underwent under the framework of the Pan-Hispanic Project. Nevertheless, the discussion has not been free of controversies, such as the one that affected the design of the mathematical formula used in statistical studies of lexical availability (Lorán and López Morales, 1983; López Chávez and Strassburger, 1987), or the one that manifested itself through multiple proposals concerning the quantity and the nature of lexical fields to be investigated. Even the edition criteria of the Pan-Hispanic Project (Samper Padilla, 1998a: 311–313), which add homogeneity to the project, have been perceived as subject to different adjustments and adaptations.

An inherent and constitutive part of the studies of available lexicon is the use of tools and procedures of statistical analysis, which include both the descriptive statistics and the use of specialist formulas (the index of availability, the index of cohesion, the multi-variant analysis, etc.). As Mateo García stated (1998: 14), “it seems quite surprising that the analysis includes the reality in the form of lexical availability which is so familiar and commonplace, together with the numbers, indexes and sociolinguistic dimensions”. In any case, this scientific treatment helps to discover statistical characteristics, which are more or less constant within the conducted studies concerning different communities, and which affect both the lexical production of the centres of interest and the behaviour of sociolinguistic variables taken into consideration. These aspects are analysed in the above-mentioned chapter III and in chapter V (*Extra-linguistic determinants of lexical availability*) on the basis of the comparison of the results of Spanish investigation under the Pan-Hispanic Project.

Chapter IV (*Relations between the words and available lexicon*) has been inspired by the studies by Hernández Muñoz (2006), as well as by Ávila Muñoz and Villena Ponsoda (2010); it looks into the directions of investigation initiated by these scholars, which aim at discovering the structure of the mental lexicon on the basis of the data offered by the studies on availability, and through the application of the most relevant psycholinguistic models in the area of lexical production. The chapter shows that the analysis of associations produced between the words, either within a lexical list produced by an informer, or within a lexical group attributable to a collective, is a powerful tool which can be used for the deduction of a psychological reality of available lexicon. The studies, such as the ones by Hernández Muñoz (2006), which view the available lexicon as a complex cognitive phenomenon, are the most serious contribution to the integrated theory of lexical availability, promoted by López Morales (1999a: 26), in which factors such as semantic memory, familiarity etc. play an important part.

Chapter V (*Extra-linguistic determinants of lexical availability*) analyse the impact of social factors on lexical production. One of the main characteristics of the Pan-Hispanic Project of Lexical Availability is the adoption of a sociolinguistic perspective of variationist nature, which tries to depict sociolectal stratification of the studied community. According to the common standards of the Pan-Hispanic Project, the obligatory social variables within each investigation are ‘sex’ (man/woman), ‘scholarly centre’ (public/private one), ‘geographical zone’ (urban/rural one), and ‘socio-cultural level’ (high/medium/ low one). This chapter also treats about quantitative analysis of these variables under different investigations. This study indicates the trends on the basis of the comparison of data coming from these investigations, which are present in different studies in relation to each variable, within which a strong impact of sociocultural level on the communitarian lexical

variation is quite visible. Apart from this variable, other variables are also analysed, such as 'the age' and the extra-linguistic determinants which have to be taken into account in a situation of contact between the languages.

In chapter VI (*Qualitative analysis of available lexis*) we are studying a series of linguistic factors that determine the composition of lists of available lexicon, included in the dictionaries which are the result of different investigations. As García Marcos said (1997: 25), apart from social factors, it is necessary to take into account the studies of available lexicon, "strictly linguistic facts which are not only (or which can be) responsible for the configuration assumed by a specific level of available lexicon, but also those that are at the very basis of linguistic and sociolinguistic dynamics, which of course corresponds to a specific community using a given language". The nature of linguistic facts that we analyse allows for different classifications of words included in the lists, which reflect the variation (diatopic, diastratic or diaphasic one), lexical borrowing, the formation and creation of words, etc.

Finally, before the presentation of final conclusions, chapter VII (*The comparison of available lexicons*) analyses the methodology used for inter-dialectal comparisons. This chapter thus presents the comparison of the five lexicons available in Spain, which has been the ultimate objective of the Pan-Hispanic Project, and which corresponds to an attempt of establishing a lexical norm of Spanish in Spain. Thanks to this activity we can show the common and differing aspects of these available lexicons and to illustrate the final phase of the aforementioned macro project: the development of dictionaries of available lexicons that will include the results of preceding investigations.



## OD REDAKCJI

Antonio María López González, urodzony w 1969 r. w Almerii (Hiszpania), jest absolwentem filologii hiszpańskiej Uniwersytetu w Almerii, gdzie w roku 2001 uzyskał tytuł doktora filologii hiszpańskiej. Jego praca doktorska dotyczyła analizy socjolingwistycznej języka radiowego w rodzinnym mieście – Almeria oraz standardu języka hiszpańskiego w mediach. Dzięki swojej dysertacji dołączył do projektu DIES-M, *Difusión Internacional del Español por los Medios* [„Międzynarodowe medialne rozpowszechnienie języka hiszpańskiego”], koordynowanego przez prof. Raúlą Ávilę z Colegio de Mexico.

W Polsce działalność akademicką rozpoczął w 2001 r. na Uniwersytecie im. A. Mickiewicza oraz w Wyższej Szkole Języków Obcych w Poznaniu. Od 2007 r. związany jest z Uniwersytetem Łódzkim.

Dorobek naukowy dr. Antonio M.<sup>a</sup> López’a González’a to ponad 40 prac naukowych, z których znaczna większość została opublikowana jako artykuły lub rozdziały w czasopismach i monografiach w Hiszpanii, Polsce i Niemczech. Oprócz tego dr López González zaprezentował ponad 20 referatów na konferencjach naukowych w kraju i za granicą.

Jego główne obszary zainteresowań naukowych to zmienność stylistyczna i socjolingwistyczna w języku hiszpańskim, zastosowanie języka hiszpańskiego w środkach masowego przekazu, akwizycja języka (dyspozycyjność leksykalna w języku hiszpańskim jako języku obcym) oraz dydaktyka języka hiszpańskiego jako języka obcego.

W latach 2005–2007 był stypendystą Ministerstwa Spraw Zagranicznych Hiszpanii, na Uniwersytecie im. Adama Mickiewicza w Poznaniu. Tam też realizował projekt badawczy pod tytułem „Aspekty dydaktyki leksykalnej w szkołach dwujęzycznych z języka hiszpańskiego: zakres zdolności użycia przyswojonego słownictwa hiszpańskiego”, który pozwolił mu na włączenie się do prestiżowego międzynarodowego projektu badawczego „Dyspozycyjność leksykalna w świecie hiszpańskojęzycznym”, koordynowanego przez prof. Humberto López’a Morales’a z Real Academia Española w Madrycie. Od 2007 r., pracując już na Uniwersytecie Łódzkim, dr López González zajmuje się perspektywą językoznawczą, jak i dydaktyczną w badaniach nad kompetencją i dyspozycyjnością leksykalną młodzieży uczącej się języka hiszpańskiego w gimnazjach i liceach w Polsce.

Od lat dr López González regularnie prowadzi warsztaty metodyczne dla nauczycieli języka hiszpańskiego. Współpracuje również z Biurem Radcy ds. Edukacji przy Ambasadzie Hiszpanii w Warszawie w kontekście zadań sekcji dwujęzycznych.

