

DISPONIBILIDAD LÉXICA EN LA PROVINCIA DE LLEIDA: ASPECTOS METODOLÓGICOS¹

MARIBEL SERRANO ZAPATA
UNIVERSITAT DE LLEIDA

1. Introducción

No es nada nuevo en lingüística emprender el análisis del léxico que emplea una comunidad de habla determinada. Con anterioridad, muchos investigadores desde enfoques muy distintos han pretendido estudiar el conjunto de palabras que integran una lengua. La investigación que aquí detallamos nace también del deseo de conocer cuál es el léxico real que usa una cierta comunidad lingüística.

Como es sabido, suele haber una notable distancia entre las numerosísimas palabras que quedan recogidas en los diccionarios generales y las pocas que utiliza un hablante «corriente» en sus interacciones cotidianas². Los estudios de léxico-estadística han tratado, mediante diferentes métodos, de establecer índices estadísticos para las unidades léxicas de una lengua.

Los primeros intentos se hicieron mediante *listas de frecuencia* que recogían los vocablos ordenados según rangos, es decir, en función de las veces que aparecían en las muestras estudiadas. Más tarde, y con la finalidad de intentar paliar los problemas que presentaba la aplicación única del criterio de frecuencia, surgieron los *léxicos básicos*, que ponderaban la frecuencia con la dispersión; de este modo, se obtenían las palabras más usuales de una comunidad, caracterizadas por un alto grado de estabilidad que les permitía aparecer frecuentemente en todo tipo de discursos (H. LÓPEZ MORALES 1986: 53-66, M^a V. MATEO 1998: 27-31).

Por otro lado, en los años 50, un grupo de investigadores franceses que intentaban, con una finalidad esencialmente pedagógica, establecer cuál era el léxico "fundamental" de la lengua francesa³ se percataron de que muchos de los términos más habituales y cotidianos no aparecían en las listas de frecuencia que se utilizaban. Muy pronto se descubrió que existía una parcela del léxico, las palabras "temáticas", que sólo se actualizaban si el tema del discurso así lo permitía. Estas voces, principalmente sustantivos de contenido semántico específico, eran poco frecuentes e inestables a pesar de ser las más conocidas y utilizadas por los hablantes. Así nació el concepto de *léxico disponible*, o, lo que es lo mismo, el caudal léxico utilizable en una situación comunicativa dada.

Para determinar el grado de disponibilidad de cada palabra, los mismos investigadores franceses llevaron a cabo una investigación empírica. Bajo unas mismas condiciones de reacción, propusieron a sus informantes una serie de estímulos asociativos denominados "centros de interés"⁴. Las listas de palabras que los informantes escribían en torno a los citados centros, constituían las palabras más disponibles, es decir, aquéllas que sin ser particularmente frecuentes acudían a su mente con mayor rapidez cuando tenían la necesidad de utilizarlas.

En realidad, los léxicos básicos y los disponibles son dos procedimientos complementarios de recabar el léxico de una comunidad de habla. Mientras que los léxicos básicos ofrecen una

¹ Con el apoyo del Departamento de Universidades, Investigación y Sociedad de la Información de la Generalitat de Cataluña.

² Parece ser que mientras un hablante "corriente" o "común" no emplea más de 2000 palabras, un hablante culto es capaz de manejar 4000 ó 5000 vocablos (M^a V. MATEO 1998: 36-37, A. CARCEDO 2000: 11).

³ Nos referimos a G. GOUGENHEIM *et alii* 1964.

⁴ En realidad, tanto la distinción entre palabras "temáticas" y "atemáticas" como la definición de disponibilidad léxica y la metodología ideada para su cuantificación, son obra de René Michéa, uno de los autores de *L'élaboration du Français Fondamental* (G. GOUGENHEIM *et alii* 1964: 144-156).

proporción real del uso de las diferentes clases de palabras, la disponibilidad léxica señala el vocabulario concreto que el hablante podría utilizar en un contexto comunicativo dado. De la suma de ambos repertorios se conformará el *léxico fundamental* de una lengua (H. LÓPEZ MORALES 1986: 63).

2. Presupuestos teóricos

Los estudios de disponibilidad léxica, por tanto, se enmarcan dentro de la léxico-estadística o lexicometría. Sin embargo, los resultados que ofrecen estos estudios se han aplicado, con mayor o menor éxito, a numerosas disciplinas lingüísticas, como la psicolingüística, la dialectología, la etnolingüística, la sociolingüística o la lingüística aplicada. Concretamente, desde el enfoque sociolingüístico se han realizado las principales investigaciones españolas de disponibilidad léxica⁵. La inmensa mayoría de estas investigaciones ha demostrado, con anterioridad, que el hecho de someter la disponibilidad léxica a análisis de covariación con distintos factores sociales permite establecer diferentes estratificaciones en las comunidades de habla, al mismo tiempo que caracterizar léxicamente los sociolectos que las integran (H. LÓPEZ MORALES 1999: 23-25).

De entre todas las variables sociales que se manejan, sin lugar a dudas, es la del nivel sociocultural o socioeconómico la que arroja los resultados más contundentes. En este sentido, los estudios sobre disponibilidad léxica parecen coincidir con la teoría del déficit de Bernstein (Cf. H. LÓPEZ MORALES 1979). Otros estudios han demostrado también la incidencia de otros factores como el sexo, el tipo de educación o la ubicación geográfica en el análisis cuantitativo y cualitativo de los resultados⁶.

Además, otra disciplina que se ha beneficiado enormemente de los estudios de léxico disponible es la lingüística aplicada. Las investigaciones sobre disponibilidad léxica son una herramienta imprescindible a la hora de planificar la enseñanza de esa lengua, ya sea como lengua materna o como lengua extranjera.

3. Delimitación del objeto de estudio

La investigación que aquí presentamos tiene como objetivo principal conocer el léxico disponible en lengua castellana y en lengua catalana en diferentes campos léxicos de una muestra representativa de escolares de la provincia de Lleida. Al mismo tiempo, pretendemos establecer una estratificación social a partir de los datos que nos aporte el análisis conjunto de los campos léxicos y de las distintas variables sociales. Partimos de la hipótesis de que las variables sociales que hemos elegido covariarán en mayor o menor medida con los resultados de los listados y que, tras realizar el análisis cuantitativo y cualitativo, podremos llegar a conclusiones significativas en cuanto a la distribución social del léxico. También pretendemos observar las características del léxico disponible de los escolares en lengua castellana teniendo en cuenta que nuestro estudio se realiza en una zona bilingüe, y comparar estos resultados con los que obtengamos del catalán. Nuestra propia experiencia como miembros de la comunidad de habla que estudiamos nos hace predecir que observaremos en los resultados finales una influencia importante de la lengua catalana y, por lo tanto, una notable presencia de convergencias e interferencias lingüísticas. Con el fin de analizar con precisión cuál es el influjo en la lengua castellana del catalán incorporamos como variables sociales tanto la lengua materna como la lengua habitual del informante.

4. Pasos metodológicos

⁵ De hecho, ya los pioneros franceses, desde una perspectiva muy avanzada para su época, dieron cuenta de la importancia que algunos factores sociales como el sexo o el medio geográfico cobraban al analizar las palabras obtenidas en las listas de disponibilidad (véase G. GOUGENHEIM *et alii* 1964: 163-182).

⁶ Véase un resumen de la influencia que han tenido las distintas variables en los estudios anteriores de léxico disponible en M^a V. MATEO 1998: 47-54.

En cuanto a la metodología que utilizamos, podemos afirmar que, *grosso modo*, hemos seguido las pautas de las investigaciones españolas en curso o ya finalizadas⁷. El motivo de la uniformidad metodológica es la existencia de un proyecto panhispánico, coordinado e impulsado por Humberto López Morales, que tiene como finalidad establecer el léxico disponible general del español. Todas las investigaciones que se suman a este proyecto comparten el mismo modo de hacer para posibilitar la comparación interdialectal de los resultados y su utilización en una planificación adecuada de la enseñanza del léxico español, ya sea como lengua materna o como lengua extranjera⁸. Nuestro trabajo constituye una novedad en la disciplina en el sentido de que es el primero que se realiza en Cataluña.

Siguiendo las pautas del proyecto, nuestro universo ha quedado integrado por la totalidad de alumnos de 2º de bachillerato de la provincia de Lleida⁹. Aunque se han realizado estudios de disponibilidad léxica con informantes de edades muy diversas, las investigaciones hispánicas que se llevan a cabo desde los años 90 prefirieron trabajar con esta franja poblacional. La elección de esta edad responde al deseo de querer estudiar el vocabulario de hablantes ya formados léxicamente pero que aún no han adquirido los tecnicismos propios de los estudios universitarios (P. BENÍTEZ 1992: 73, H. LÓPEZ MORALES 1999: 28).

La muestra constituye alrededor del 13,85% del universo estudiado, ya que la media de los alumnos matriculados durante los cursos académicos 2001/2002 y 2002/2003 en este nivel educativo ascendía a 1834. Han cumplimentado la encuesta exactamente 253 alumnos en castellano (13,79% del universo) y 255 en catalán (13,9% del universo). En realidad se trata de dos muestras, la del castellano y la del catalán, en las que la mayor parte de los informantes coinciden y que representan un mismo universo.

Debido a que la media de edad de los encuestados era de 17 años, hemos eliminado de la muestra los resultados de aquellos individuos que llevaban menos de 10 años viviendo en la provincia de Lleida. Con esta medida nos hemos asegurado de que nuestros informantes han pasado la mayor parte de su vida en el lugar de la provincia donde los hemos encontrado. También de este modo y, siguiendo fielmente nuestra intención de estudiar las características del léxico español en un territorio bilingüe, tenemos la certeza de que sólo hemos recabado los datos de informantes bilingües o que han estado en contacto con el catalán durante la mayor parte de su vida¹⁰.

Hemos realizado la encuesta durante el año 2002 en un total de doce centros educativos; seis de la ciudad de Lleida y seis de diferentes puntos de la provincia. En principio, la elección de los centros urbanos no fue aleatoria sino que la realizamos tras solicitar en una carta ciertos datos que nos interesaban, como el número total de alumnos del nivel preuniversitario y su lugar de nacimiento. Seleccionamos tres centros públicos de la ciudad (*I.E.S. Ronda*, *I.E.S. Joan Oró* e *I.E.S. Escola del Treball*) y tres privados (*Col·legi Lestonnac*, *Col·legi Episcopal* y *Col·legi Maristes Montserrat*) distribuidos estratégicamente en la ciudad. Con la elección de estos centros hemos procurado que se encontraran representados los distintos barrios de Lleida.

Por otro lado, los centros rurales se escogieron principalmente en función de su ubicación territorial para que pudieran ser representativos de la provincia. Para realizar tal elección tuvimos en cuenta las diversas fronteras territoriales y lingüísticas de Lleida, ya que ésta limita al oeste con

⁷ Véase una panorámica general de estas investigaciones en la página web: www3.usal.es/dispollex

⁸ Cabría preguntarse, en este punto, si la comparación interdialectal justifica por sí sola la inmovilidad de la metodología. Si bien es cierto que la uniformidad metodológica permite el cotejo de los distintos trabajos, también es verdad que cualquier metodología puede necesitar variar conforme se desarrolla la disciplina en la que nace.

⁹ Este nivel educativo corresponde al antiguo C.O.U., ya totalmente extinguido en nuestra provincia.

¹⁰ Además de este modo reflejamos fielmente la propia realidad de la provincia de Lleida donde la población bilingüe es mayoritaria y en la que los habitantes no nacidos en Cataluña constituyen sólo el 16,9% de la población total.

el castellano y al este con el catalán oriental. Predomina en nuestra provincia la variedad dialectal del catalán denominada *nordoccidental*. A pesar de que ésta es la variedad con mayor extensión, es posible encontrar otras subvariedades dentro de la provincia como el *pallarés* o el *ribagorzano*. Mención aparte merece el caso del aranés, variedad de la lengua occitana propia del Valle de Arán¹¹. Teniendo en cuenta estas circunstancias los centros finalmente elegidos son los situados en las siguientes poblaciones: Borjas Blancas (*I.E.S. Josep Vallverdú*), Tárrega (*I.E.S. Alfons Costafreda*), Solsona (*I.E.S. Francesc Ribalta*), Seo de Urgel (*I.E.S. Joan Budrieu*), Viella (*I.E.S. d'Aran*) y Almacellas (*I.E.S. Canigó*). Las tres primeras poblaciones se corresponden con las capitales de comarca que más se aproximan al límite geográfico este de la provincia y limitan respectivamente con las provincias de Tarragona y Barcelona. Nos interesaban puesto que la primera limitaba lingüísticamente con la variedad del catalán central hablada en Tarragona (que algunos han denominado *tarragoní*) y las otras dos poblaciones eran los puntos más poblados que delimitan lingüística y geográficamente con la provincia de Barcelona y el catalán central.

Por otro lado, Viella y Seo de Urgel son los núcleos más importantes situados en el noroeste y noreste respectivamente de la provincia. En el Valle de Arán son oficiales tres lenguas: el aranés, el catalán y el castellano, aunque la lengua propia de Viella es el aranés; eso explica que, lógicamente, ésta apareciera como lengua materna y habitual en algunos de los informantes encuestados en ese centro escolar. La última población escogida, Almacellas, supone el límite oeste provincial y limita con la franja lingüística extendida verticalmente entre las comunidades autónomas de Cataluña y Aragón donde se extiende una variedad lingüística de transición que actúa de puente entre el catalán y el aragonés. Esperamos que la elección de estas poblaciones refleje las diversas influencias lingüísticas que pueden observarse en nuestra provincia.



¹¹ En realidad, el aranés es un dialecto del occitano, al igual que el catalán

En cuanto a las variables sociológicas que hemos utilizado para la estratificación de la muestra, éstas coinciden, en general, con las utilizadas por el resto de investigadores en otros lugares del mundo hispánico. De este modo hemos tenido en cuenta el sexo del informante, el tipo de enseñanza, la ubicación del centro y el nivel sociocultural de la familia. Como ya se ha dicho, por realizarse nuestro trabajo en un territorio bilingüe, incluimos dos variables más, la lengua materna y la lengua de uso habitual del informante, que no se incluyen, como es lógico, en los territorios hispánicos monolingües.

La variable *sexo* no resulta problemática. Por el contrario, la variable *tipo de enseñanza* (pública/privada) sólo ha podido ser considerada en el caso de los informantes de la ciudad de Lleida. El motivo de la desestimación de los informantes del resto de la provincia es que en la mayoría de los centros rurales escogidos no existía ningún centro educativo privado que impartiera el nivel educativo que estudiábamos. Con el fin de no entorpecer el análisis de los datos y de que la variable *tipo de enseñanza* resultara realmente significativa, desechamos el estudio de esta variable en los informantes procedentes de centros rurales.

Por otro lado, en un primer momento, pretendíamos reflejar en esta variable el tipo de enseñanza que el alumno había recibido durante sus años de escolaridad. Por esta razón, incluimos en la encuesta una pregunta para determinar si el alumno había realizado la mayoría de sus estudios en el centro en el que lo encontramos. Muy pronto nos percatamos de que las características del actual sistema educativo en el que el bachillerato constituye un ciclo educativo de dos años no nos permitían establecer esta concreción: mientras que la mayoría de los alumnos de los centros educativos privados sí habían realizado allí la educación secundaria y, por tanto, llevaban seis años como mínimo de escolaridad privada, en los centros públicos no ocurría lo mismo y contábamos con muchos informantes que habían realizado la secundaria en un centro privado¹². Esta circunstancia nos obligó a tomar como determinación asegurar, por lo menos, la realización del bachillerato de nuestros informantes en el centro educativo en el que lo encontramos. Por todas estas razones, no estamos seguros de que esta variable muestre finalmente una covariación significativa en relación con otros estudios que preceden al nuestro.

En cuanto a la variable *ubicación del centro educativo*, los informantes se distribuyen entre centros urbanos y centros rurales. Al contrario de lo que ocurre en otras provincias, Lleida no contiene, exceptuando la capital, ningún núcleo urbano superior a los 20.000 habitantes. Por lo tanto, hemos considerado como centros urbanos los seis escogidos en la ciudad de Lleida y como rurales, el resto de centros de la provincia. Por consiguiente, hemos eliminado aquellos informantes que, a pesar de asistir a un centro educativo de la ciudad de Lleida, no residían en la misma ciudad sino en poblaciones cercanas. Consideramos que los centros urbanos deben representar a la colectividad urbana y que la inclusión de estos informantes podría distorsionar el peso de esta variable.

Para determinar el *nivel sociocultural* de la familia del alumno hemos tenido en cuenta la profesión y los estudios del padre y de la madre¹³, y el nivel cultural de la familia. Este último punto supone una novedad metodológica ya que los estudios de disponibilidad léxica llevados a cabo con anterioridad sólo habían considerado los dos primeros parámetros¹⁴. Nos parecía que, por lo menos en nuestra comunidad de habla, no siempre el nivel cultural de una familia tiene un paralelismo exacto con el grado de instrucción de los padres ni, mucho menos, con el estatus social de las profesiones que tienen¹⁵.

¹² El volumen de estos informantes nos impedía eliminarlos.

¹³ Para la clasificación de los niveles de estudios y de las distintas profesiones seguimos el índice que detalla J. A. SAMPER 1997: 230 nota 6 y 1990: 40-41, especialmente el artículo que cita en la nota 10.

¹⁴ Otros estudios, realizados principalmente en Hispanoamérica, tienen en cuenta además los ingresos económicos de la familia (vg. H. LÓPEZ MORALES 1999: 30).

¹⁵ M^a V. MATEO (1998: 77) también expresa sus reservas a esta forma concreta de establecer el factor sociocultural de un informante y constata, siguiendo a García Marcos, que otros autores, en otro tipo de

Con el fin de establecer el nivel cultural de la familia incorporamos en la encuesta una serie de preguntas que versaban sobre distintos aspectos: la cantidad de libros que leía el informante o que se compraban en su casa, la pertenencia del alumno a alguna asociación de carácter cultural o la frecuencia con la que asistían sus padres a distintos acontecimientos culturales como espectáculos teatrales, de danza o conciertos musicales¹⁶. Más tarde, consideramos los tres factores estudiados (las profesiones y los estudios de los progenitores y el nivel cultural de la familia) mediante una sumatoria no ponderada y establecimos un *continuum* que dividimos en tres grupos correspondientes al nivel sociocultural alto, medio y bajo.

Las otras dos variables restantes se refieren a la adscripción lingüística del informante. Respecto a la *lengua materna*, preguntamos directamente cuál era. En cambio, para establecer la *lengua de uso habitual*, creímos más oportuno detallar los principales ámbitos en los que se mueve un estudiante de 2º de bachillerato y, por tanto, preguntamos qué lengua utilizaba en su casa, con sus amigos y en la escuela. Una vez obtenidas las respuestas observamos cuál era la lengua habitualmente empleada por el alumnado. En principio estas dos variables iban a tener sólo tres variantes, catalán, castellano o ambas. Sin embargo, al pasar las encuestas en el Valle de Arán tuvimos que añadir una variante más para aquellos informantes cuya lengua materna o habitual era el aranés. No obstante, debido al poco volumen de estos informantes no creemos que esta variante pueda covariar de manera significativa con los datos.

Otra variable que hubiéramos podido tener en cuenta es la de *modelo de enseñanza o lengua utilizada en la enseñanza*¹⁷ 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 de nuestra comunidad lo impedía, puesto que no existe ningún centro educativo en nuestra provincia que no use el catalán como lengua vehicular de la enseñanza¹⁸. No obstante, hemos recopilado de cada centro información sobre el número de horas semanales que recibe el alumnado en catalán y en castellano para considerarlo en nuestro estudio.

Finalmente, la muestra en castellano queda distribuida sociológicamente de la siguiente forma:

- a) Sexo: varón, 109; mujer, 144.
- b) Tipo de centro: público, 56; privado, 72.
- c) Ubicación del centro: urbano, 128; rural, 125.
- d) Nivel sociocultural: bajo, 96; medio, 120; alto, 37 ;
- e) Lengua materna: castellano, 66; catalán; 172; ambas (castellano/catalán), 11; aranés, 4.
- f) Lengua de uso habitual: castellano, 56; catalán, 179; ambas(castellano/catalán), 16; aranés, 2.

La muestra para el catalán se reparte del modo que sigue:

- a) Sexo: varón, 104; mujer 151.
- b) Tipo de centro: público, 65; privado, 74.
- c) Ubicación del centro: urbano, 139; rural, 116.
- d) Nivel sociocultural: bajo, 101; medio, 113; alto, 35.
- e) Lengua materna: castellano, 69; catalán; 168; ambas (castellano/catalán), 12; aranés, 5.
- f) Lengua de uso habitual: castellano, 57; catalán, 174; ambas (castellano/catalán), 20; aranés, 2.

El mecanismo de obtención de los datos ha sido el siguiente. Como ya dijimos, enviamos una carta a todos los centros que impartían 2º de bachillerato en la capital y a los situados en las

estudios, distinguen con mayor nitidez entre clase social y cultura. En la ciudad de Lleida ya se ha realizado una tesis doctoral en la que se plantea la influencia del estatus sociocultural de la familia en las construcciones sintácticas del lenguaje infantil (véase D. MAYORAL 1994).

¹⁶ Estas preguntas están inspiradas en la noción de "capital cultural" de D. MAYORAL (1994: 109-110 y 122-125).

¹⁷ Sobre la incidencia de esta variable en otro territorio bilingüe, Bilbao, véase M. ETXEBERRÍA 1996.

¹⁸ La *Llei de Normalització Lingüística* (Ley 7/1983 de 18 de abril, D.O.G.C. de 22.04.83) establece que el catalán es la lengua propia de la enseñanza en todos los niveles educativos.

poblaciones que nos interesaban del resto de la provincia. En esta carta explicábamos las directrices básicas de la investigación y solicitábamos ciertos datos estadísticos sobre el alumnado que nos resultaban de interés. Después, nos pusimos en contacto con los centros seleccionados y el director de cada uno de ellos o el tutor del curso nos asignó al azar un grupo del centro y estableció el día y la hora en que podíamos pasar las encuestas. La hora lectiva destinada a tal fin acostumbro a ser la de tutoría o la de lengua castellana o catalana. Una vez en el aula, se entregó a cada alumno un grupo de fotocopias grapadas y se explicó brevemente la finalidad del estudio y el mecanismo de funcionamiento de la encuesta. En las dos primeras páginas del cuadernillo se encontraban las preguntas encaminadas a determinar las variables sociológicas de cada alumno y, en la tercera y siguientes, se distribuían los distintos centros de interés, en torno a los cuales los alumnos tenían que producir las listas de unidades léxicas; los que hemos utilizado en nuestra investigación, han sido:

- | | |
|--|--|
| 01. El cuerpo humano | <i>01. El cos humà</i> |
| 02. Ropa y complementos | <i>02. Roba i complementes</i> |
| 03. Partes de la casa (sin los muebles) | <i>03. Parts de la casa (sense els mobles)</i> |
| 04. Muebles de la casa | <i>04. Mobles de la casa</i> |
| 05. Alimentos y bebidas | <i>05. Aliments i begudes</i> |
| 06. Objetos colocados en la mesa para la comida | <i>06. Objectes col·locats a la taula per menjar</i> |
| 07. La cocina y sus utensilios | <i>07. La cuina i els seus utensilis</i> |
| 08. La escuela (muebles y materiales) | <i>08. L'escola (mobles i materials)</i> |
| 09. Iluminación, calefacción y medios para airear un recinto | <i>09. Il·luminació, calefacció i mitjans per airejar un recinte</i> |
| 10. La ciudad | <i>10. La ciutat</i> |
| 11. El campo | <i>11. El camp</i> |
| 12. Los medios de transporte | <i>12. Mitjans de transport</i> |
| 13. Trabajos del campo y del jardín | <i>13. Treballs del camp i del jardí</i> |
| 14. Animales | <i>14. Animals</i> |
| 15. Juegos y distracciones | <i>15. Jocs i distraccions</i> |
| 16. Profesiones y oficios | <i>16. Professions i oficis</i> |
| 17. Colores | <i>17. Colors</i> |

Como se verá, estos centros son, salvando algunas pequeñas diferencias, los mismos que se utilizaron en el estudio pionero de G. GOUGENHEIM *et alii* 1964. Aunque algunos investigadores han utilizado otros, la mayor parte de los estudios españoles ha optado por mantener los dieciséis centros clásicos en aras de la uniformidad metodológica a la que antes aludíamos. Las únicas modificaciones que hemos realizado han sido la inclusión —al igual que han hecho otras investigaciones— del centro “colores”¹⁹ así como la ampliación del segundo centro de interés a “ropa y complementos”²⁰. Siguiendo una vez más las pautas del proyecto panhispánico, se le concedió al alumno un tiempo límite de dos minutos para escribir todas las palabras que asociara al centro en cuestión en las líneas numeradas debajo de cada uno de ellos, pudiendo sobrepasarlas si el tiempo se lo permitía. La duración total de la encuesta osciló entre 45 y 50 minutos durante los cuales los alumnos contestaron en silencio a cada uno de los centros de interés. No todas las

¹⁹ H. LÓPEZ MORALES y F. GARCÍA MARCOS (1995: 70) justifican la inclusión de este centro con las siguientes palabras: «“los colores” nos introducen en un aspecto básico de la verbalización de la percepción del entorno, de constante referencia desde los trabajos de los antropólogos lingüistas para el examen de las relaciones lengua/pensamiento/realidad».

²⁰ A pesar de que el centro clásico era simplemente “ropa” (*vêtement*), otros investigadores ya han constatado con anterioridad que, en realidad, los alumnos asocian a este campo tanto prendas de vestir como accesorios y complementos (véase A. CARCEDO 2000: 76, J. A. SAMPER 1997: 318).

encuestas nos fueron de utilidad, algunas quedaron automáticamente eliminadas por la evidente falta de seriedad que demostraban²¹.

Este mecanismo se repitió al cabo de dos semanas de forma idéntica con el cuestionario sobre disponibilidad léxica catalana. El hecho de dejar dos semanas de tiempo entre una prueba y otra respondía a nuestro deseo de que el alumno se viera obligado a realizar de nuevo el mecanismo asociativo sin traducir las palabras que, de otro modo, pudiera haber recordado de la primera prueba hecha en castellano, toda vez que los grupos de 2º de bachillerato a los que se pasó la encuesta en catalán fueron los mismos que los que habían respondido a la encuesta en castellano. Sin embargo, se observan diferencias entre una muestra y otra que se explican porque aproximadamente 30 alumnos de los diferentes centros educativos no se hallaron en el aula, por causas indeterminadas, uno de los dos días en que pasamos la encuesta.

5. Conclusión

En este momento estamos estudiando los criterios de edición que adoptaremos al vaciar las encuestas. El próximo paso metodológico será el tratamiento estadístico de los datos mediante el programa informático *Lexidisp*, empleado habitualmente en los estudios sobre disponibilidad léxica, que utiliza para el cálculo de la disponibilidad una fórmula matemática que tiene en cuenta no sólo la frecuencia de las unidades léxicas sino también la posición que ocupan en las listas. Esperamos que un futuro próximo podamos contar ya con algunos resultados que nos ayuden a conocer la disponibilidad léxica de la lengua castellana y catalana en territorio bilingüe.

Referencias bibliográficas

- ORLANDO, ALBA, *Léxico disponible de la República Dominicana*, Santiago de los Caballeros, Pontificia Universidad Católica Madre y Maestra, 1995.
- BENÍTEZ PÉREZ, PEDRO, "Disponibilidad léxica en la zona metropolitana de Madrid", *Boletín de la Academia Puertorriqueña de la Lengua Española*, 1, 1, 1992, 71-102.
- CARCEDO GONZÁLEZ, ALBERTO, *Disponibilidad léxica en español como lengua extranjera: el caso finlandés (estudio del nivel preuniversitario y cotejo con tres fases de adquisición)*, Turku, Annales Universitatis Turkuensis, 2000.
- ECHVERRÍA, MAX S., *et alii*, "Disponibilidad léxica en Educación Media", *Revista de Lingüística Teórica y Aplicada*, 25, 1987, 55-115.
- ETXEBARRÍA, MAITENA, "Disponibilidad léxica en escolares del País Vasco. Variación sociolingüística y modelos de enseñanza bilingüe", *Revista Española de Lingüística*, 26, 2, 1996, 301-325.
- GALLOSO CAMACHO, M^a VICTORIA, "Tratamiento y perspectivas del "léxico disponible": el caso de la provincia de Zamora", *REALE*, 8, 1997, 23-33.
- GONZÁLEZ MARTÍNEZ, ADOLFO, *La disponibilidad léxica de los alumnos preuniversitarios de la provincia de Cádiz*, Universidad de Cádiz, 2002.
- GOUGHENHEIM GEORGES, RENE MICHEA, PAUL RIVENC y AURELIEN SAUVEGOT, *L'élaboration du Français Fondamental (1er degré). Étude sur l'établissement d'un vocabulaire et d'une grammaire de base*, Paris, Didier, 1964.
- LÓPEZ CHÁVEZ, JUAN, "Alcances panhispánicos del léxico disponible", *Lingüística*, 4, 1992, 26-124.
- LÓPEZ MORALES, HUMBERTO, "Disponibilidad léxica y estratificación socioeconómica", en *Dialectología y Sociolingüística. Temas puertorriqueños*, Madrid, Hispanova de Ediciones, 1979, 173-181.
- LÓPEZ MORALES, HUMBERTO, *La enseñanza de la lengua materna. Lingüística para maestros de español*, Madrid, Playor, 1986.
- LÓPEZ MORALES, HUMBERTO, (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, 1991.

²¹ Curiosamente todas las encuestas que fueron eliminadas por estas causas fueron de informantes varones.

- LÓPEZ MORALES, HUMBERTO y FRANCISCO GARCÍA MARCOS, "Disponibilidad léxica en Andalucía. Proyecto de investigación", *REALE*, 3, 1995, 65-76.
- LÓPEZ MORALES, HUMBERTO, *Léxico disponible de Puerto Rico*, Madrid, Arco/Libros, 1999.
- MACKAY, WILLIAM C., *Le vocabulaire disponible du français*, Paris-Bruxelles-Montréal, Didier, 1971.
- MAYORAL DOLORS, *Lenguaje y estatus sociocultural: un estudio empírico sobre el lenguaje infantil en los barrios de Lleida* (tesis doctoral inédita), Universitat de Lleida, 1994.
- MATEO, M^a VICTORIA, *Estratificación social del C.O.U. almeriense*, Universidad de Almería, Grusta, 1998.
- SAMPER PADILLA, JOSÉ ANTONIO, *Estudio sociolingüístico del español de Las Palmas de Gran Canaria*, Las Palmas de Gran Canaria, La Caja de Canaria, 1990.
- SAMPER, JOSÉ ANTONIO y CLARA EUGENIA HERNÁNDEZ, "Estudio de la disponibilidad léxica en Gran Canaria. Datos iniciales y variación sociolingüística", en *Contribuciones al estudio de la lingüística hispánica. Homenaje al profesor Ramón Trujillo*, Universidad de la Laguna, 1997, vol. II, 229-239.
- SAMPER PADILLA, JOSÉ ANTONIO, "Criterios de edición del léxico disponible: sugerencias", *Lingüística*, 10, 1997, 311-333.