

Diccionarios de léxico disponible en español en comparación con los diccionarios de lengua extranjera (inglés, portugués) de alumnos universitarios hispanohablantes¹

Marisela Colín Rodea
María Noemí Alfaro Mejía
Eréndira Dolores Camarena Ortíz
María de los Ángeles Cervantes Amador
CENTRO DE ENSEÑANZA DE LENGUAS EXTRANJERAS, UNAM
colinr@hotmail.com

Resumen

En este artículo analizamos los resultados del contraste realizado entre los diccionarios de léxico disponible en español y los diccionarios de lengua extranjera (inglés, francés, portugués y ruso) de alumnos universitarios hispanohablantes. En los estudios de disponibilidad léxica se denomina diccionario al conjunto léxico en forma de listas organizadas por centro de interés, elaboradas por un hablante ante la presencia de un estímulo. Michéa (1953: 342) considera como léxico disponible las palabras que vienen a la mente de inmediato ligadas a una situación dada y que la caracterizan. Una palabra disponible es una palabra que, sin ser particularmente frecuente, está, sin embargo, siempre lista para ser usada y surge en el instante en que se le necesite. Las recopilaciones del léxico disponible fueron obtenidas de estudiantes mexicanos universitarios que aprenden una lengua extranjera. Los datos fueron investigados en la literatura de la lingüística aplicada. Los diccionarios resultantes fueron introducidos en Ledical, una base de datos diseñada para comparar y analizar las relaciones entre las recopilaciones en español y la otra lengua.² En el diccionario de léxico disponible del español como lengua materna se realizó un análisis semántico de cada uno de los centros de interés contemplados y, posteriormente, este diccionario sirvió como base para hacer comparaciones con los otros diccionarios de lengua extranjera (inglés, francés, portugués y ruso).

PALABRA CLAVE: léxico disponible

¹ Para mayor información sobre los antecedentes de este proyecto, véase Anexo 1.

² Más información sobre Ledical en Anexo 2.

Introducción

El léxico ha sido tradicionalmente el plano lingüístico que se ha puesto más en relación con factores culturales. Principalmente debido a su valor denominativo: las palabras nombran el conjunto de valores, creencias, objetos, actividades y personas que configuran una cultura; nos ofrecen datos sobre la diversidad sociocultural en una misma comunidad de habla, señalan Helena Calsamiglia y Amparo Tusón (2002: 60).

Siguiendo la idea de Lara (2006: 100) sobre el término cultura, discutimos en el análisis de esta investigación cómo el léxico disponible arroja datos acerca de la lengua histórica, en lo que se refiere a la acumulación y selección de experiencias lingüísticas que una comunidad valora y selecciona como pertinente y con sentido para los miembros de esa comunidad. Asimismo, analizamos cómo la competencia léxica de un hablante depende de varios factores, entre ellos, de la memoria. Un hablante puede disponer del léxico a partir del tipo de estímulo que le exige la comunicación cotidiana. Estamos frente al léxico disponible. Por ejemplo, en el centro de interés de la encuesta: *animales*, los alumnos mexicanos reportan léxico de animales domésticos; cuando los alumnos son de provincia, el léxico reportado se asocia al subcampo: animales de granja; en el caso de los alumnos que estudian portugués se reportan animales de la selva: *tatú, anta, piranha*; los alumnos de chino, animales del zodiaco chino; o los de ruso, léxico del paisaje ruso.

El estudio del léxico y la cultura se inscribe también en la antropología lingüística. Esta disciplina se podría definir en general como el estudio del lenguaje y la cultura, y son vitales sus contribuciones para comprender la naturaleza de la lengua como parte del habla, como instrumento social y como práctica cultural (Duranti, 2000). Por otro lado, el léxico disponible son las palabras que los hablantes tienen en el lexicón y cuyo uso está condicionado por el tema concreto de la comunicación; se diferencia del léxico básico en que éste lo componen las palabras más frecuentes de una lengua independientemente del tema tratado (Taberner, 2008).

Disponibilidad léxica

El estudio del léxico disponible de una lengua materna se inicia a partir de 1950 por la escuela francesa. Marta Samper (2001) refiere estos inicios de la manera siguiente:

Los primeros estudios se realizan en Francia en los años 50 cuando los autores de lo que sería el *Français Fondamental* (Gougenheim, Michéa), se dan cuenta de que la ponderación de la frecuencia con que aparecen las palabras en los textos no es suficiente para conocer la nómina léxica usual de una comunidad lingüística. En las investigaciones que sólo consideraban este factor, no aparecían palabras muy conocidas, que eran normalmente usadas por los hablantes cuando se trataba un tema en concreto, sobre todo aquellas pertenecientes a la categoría de los sustantivos. Es por ello por lo que se empieza a trabajar con unos estímulos, los llamados *centros de interés* y se combina el análisis de la frecuencia de aparición de las unidades con la posición en las que éstas aparecen (Samper, 2001: 277).

Señala esta autora que así fue como surgió el estudio del léxico disponible, que no es otra cosa sino las palabras que vienen a la mente del hablante *cuando la conversación gira alrededor de un tema*. De esta manera, informa la autora, en los estudios de léxico disponible se considera que la suma de este léxico –obtenido por frecuencia y dispersión– más el léxico básico permite obtener el *léxico fundamental* de una lengua.

Lara (2006: 148) nos recuerda que en los primeros estudios de vocabulario fundamental se encuentran los sistemas taquigráficos, que necesitaban saber qué palabras eran imprescindibles en una lengua; y agrega que también podemos considerar como antecedentes las listas de ideogramas chinos y japoneses que las tradiciones pedagógicas determinaron como el mínimo necesario para que un alumno de la escuela elemental pudiera aprender a leer y a escribir el chino. El vocabulario fundamental es aquel que permite hablar una lengua en condiciones mínimas de comunicación. El vocabulario activo de una persona, señala Lara (2006: 148), supera las 2000 palabras y puede llegar a 8000 o 9000 palabras, dependiendo de la competencia del hablante. Existe un vocabulario aprendido que no se usa de manera activa, a éste se le denomina vocabulario pasivo. Y existe un léxico que puede pasar de un vocabulario pasivo a uno activo, mostrando la disponibilidad de este léxico en la memoria. Se trata de vocablos que pueden ser usados en cualquier acción verbal, bajo el estímulo de sus necesidades de comunicación (Lara, 2006: 173). Los factores que pueden influir este fenómeno de la memoria son muy diversos: el sexo, la edad, la educación, el conocimiento de la tradición verbal, la especialidad de cada quien y, quizá, incluso los flujos represivos del inconsciente.

Un dato relevante que nos ofrece este autor (Lara, 2012) para el caso del español de México y para la noción de vocabulario fundamental es el referente al *vocabulario fundamental de tradición culta*, correspondiente a 1451 vocablos, y especifica que 1131 corresponden al número de vocablos común a la tradición culta y a la tradición popular. Estos datos son el resultado de la experiencia cuantitativa para el tratamiento de léxico seguido en el *Corpus del español mexicano contemporáneo* (CEMC). Lara señala varios aspectos importantes sobre el *vocabulario fundamental*:

Una enseñanza notable de nuestro estudio cuantitativo, que se hizo patente al reflexionar acerca de los vocablos cuya frecuencia fue menor de diez y mayor o igual a ocho en el CEMC, es que existe un núcleo léxico de nuestro español cercano a los 15 000 vocablos, que es el que constituye el “español mexicano nacional”, es decir, el que se utiliza en todo el país, sin distinción de regiones y de correlaciones exclusivas con grupos sociales determinados. Más allá de esa cifra, la aleatoriedad constitutiva del *Corpus* se ha vuelto evidente. Se puede explicar con una imagen biológica o astronómica: notamos un núcleo léxico como el de un sol o como el de una célula, en torno al cual el resto del vocabulario de frecuencia más baja aparece como un deshilachado de su orilla como una periferia caótica. A partir de esa “membrana”, encontramos muchos nahuatlismos específicos, muchas voces propias de diversas ciencias y técnicas, muchos nombres de seres de la naturaleza, que no forman campos léxicos completos y bien estructurados, sino que apuntan a un horizonte cambiante, que depende completamente de las características y de los textos que forman este CEMC (Lara, 2012).³

Posteriormente, los estudios de disponibilidad léxica se han desarrollado con mayor profundidad a partir de 1970, gracias al impulso de Humberto López Morales, quien fundó el Proyecto Panhispánico. En México, el proyecto de Juan López Chávez en la Universidad de Zacatecas se reporta en el libro *¿Qué te viene a la memoria? La disponibilidad léxica: teoría, métodos y aplicaciones*.⁴ Al realizar una comparación con los resultados del grupo francés, se observa que el léxico obtenido es el mismo y mantiene el mismo orden; esto quiere decir que los hablantes de cada grupo realizan el mismo tipo de asociaciones. Un trabajo relacionado con esta investigación es la tesis de doctorado en lingüística de Ricardo Arriaga (2010),

³ Véase “Introducción”, <<http://dem.colmex.mx/moduls/Default.aspx?id=8>> (consulta: 8 de noviembre de 2013).

⁴ Este libro fue publicado por la Universidad Autónoma de Zacatecas en 2004.

intitulada “Involución de la complejidad lingüística. Análisis de la complejidad léxico-semántica de estudiantes de primaria y secundaria”. Además, López Chávez y Strassburguer (2000), desarrollaron el índice de disponibilidad léxica, el cual engloba todos los factores y una medida de competencia léxica de carácter individual.

Respecto al estudio del léxico disponible de lenguas extranjeras, algunas de las investigaciones más importantes han sido realizadas por Alberto Carcedo González en la Universidad de Turku (Finlandia) y por Marta Samper Hernández en la Universidad de Salamanca (España).

Almacenamiento de las encuestas y de los diccionarios en una base de datos. Esto permite la recuperación de esa información y facilita la ampliación del corpus. Esta línea fue desarrollada a lo largo de la investigación doctoral y del proyecto posdoctoral de Helen Frey Pereyra (2008: 697), en su trabajo de investigación doctoral *El léxico disponible en los escritos de alumnos de español como lengua extranjera. Estudio comparativo de dos corpora*.⁵ La investigación se orienta al análisis del léxico disponible en los corpus de español, uno del léxico disponible y el otro de los textos escritos, ambos obtenidos de los mismos estudiantes de español como lengua extranjera, que fueron investigados en una comparación única en la literatura de lingüística aplicada. Los diccionarios resultantes se introdujeron en una base de datos diseñada para comparar y analizar las relaciones entre los corpus. En esta investigación, la autora muestra que el estudio del léxico disponible proporciona una metodología científica para recoger y analizar palabras mediante procedimientos controlados, y permite obtener, a través de una serie de datos matemáticos como el índice de disponibilidad, las características del léxico de una comunidad lingüística determinada.

Utilidad y funciones del léxico disponible

Nos hemos referido al léxico disponible como a un tipo de vocabulario que todo hablante de una lengua conoce, pero que no usa más que en circunstancias con-

⁵ En el área de lenguas extranjeras, J. A. van Ek (1975), autor de *Systems development in adult language learning. The threshold level in a european unit/credit system for modern language learning by adults*, da inicio a proyectos en diversas lenguas, coordinados por J. L. M. Trim, director del Consejo de Europa en idiomas modernos desde 1975 a 1997, que especificaron el nivel de umbral inicial para el inglés, junto con la especificación desarrollada para el francés (Niveau Seuil), generando los modelos básicos para otros idiomas.

textuales muy precisas (Frey, 2008: 13). En el apartado anterior mencionamos los alcances mostrados por diversas investigaciones para el estudio del léxico de una comunidad específica. Queremos referirnos ahora al papel didáctico que representan los estudios de disponibilidad léxica.

Para sólo listar algunas de estas funciones, veamos las siguientes:

- Recolección del léxico disponible de una población para la elaboración de un diagnóstico de conocimiento.
- Recolección del léxico disponible para la medición de la adquisición de léxico en contextos de enseñanza específicos.
- Análisis de errores en las encuestas de léxico disponible.
- Uso de las encuestas de léxico disponible como herramienta de enseñanza en el salón de clase.
- Elaboración de diccionarios de léxico disponible.
- Análisis de aspectos culturales en los diccionarios de léxico disponible.
- Comparación de los diccionarios de léxico disponible de lenguas extranjeras con diccionarios de nativos.
- Propuestas didácticas para la enseñanza de léxico basadas en investigaciones de léxico disponible.
- El léxico disponible como herramienta para los estudios de lexicografía y traducción.

Los diccionarios de léxico disponible

La encuesta es el método para recoger el léxico que está disponible en una comunidad. Se constituyen así listas de palabras organizadas por centro de interés. Son estos datos los que constituyen los diccionarios individuales y de grupo. Como ya mencionamos, la aparición de este léxico es la respuesta a un estímulo en la comunicación. La disponibilidad léxica se obtiene a partir de los centros de interés de la encuesta. El informante tiene dos minutos para llenar un centro de interés. En nuestro caso, aplicamos una encuesta en español y otra en la lengua que el alumno estudia. De esta manera, pudimos recuperar el léxico; esto es, logramos activar su presencia en la memoria del hablante.

Desde una perspectiva semántica, los centros de interés son núcleos semánticos y las listas de palabras se ordenan de acuerdo a su mayor o menor presencia de

estas palabras en la memoria; intervienen diferentes factores, como la actividad del hablante o las experiencias más recientes que haya tenido.

El tipo de relaciones que se observan entre las palabras de las listas son asociaciones. Lara (2006: 183) dice que estas palabras recogidas por asociación libre nos dan un *campo asociativo*. El campo asociativo está formado por a) vocablos relacionados por el tema o los objetos de la experiencia a los que refieren (incluso términos de una especialidad, conocidos por los informantes); b) vocablos con la misma etimología; vocablos agrupados por su categoría gramatical; vocablos organizados por sus categorías morfológicas y fonológicas, y además c) vocablos de diferentes dialectos o variedades regionales o sociales de la misma lengua. Por ejemplo, si a un hispanohablante le proponemos la palabra *alimentos* como parte inicial de la tarea de asociación libre, obtenemos una constelación de vocablos como la siguiente: chocolate, pan, miel, nueces, plátano, de un informante que seguramente tenía como referencia el desayuno de ese día.

Cuando hemos obtenido una constelación, es decir significados alrededor de un núcleo o centro de interés, podemos organizarla siguiendo diferentes criterios, tales como: tipos de alimentos, procedencia de los alimentos, frutas, verduras, lácteos, postres, condimentos, ingredientes. Así, una vez organizados, entonces se puede dar paso al análisis.

Samper (2001: 279) nos permite observar la manera en que nos acercamos a las asociaciones realizadas por un hablante: a) el campo léxico en el que aparece la palabra: *cabeza, ojo, brazo, mano y nariz*; b) referencias a una categoría semántica en las respuestas; c) información sobre el hablante; d) la localización de una palabra que se encuentra entre otras que hacen ciertas referencias; e) la existencia de otras referencias del mismo tipo.

Lara (2006) señala que el análisis semántico del léxico es un método interpretativo cuya meta es garantizar que los resultados sean *intersubjetivamente verificables* y den lugar a una *verosimilitud ampliamente aceptable*. Entonces, y dado que se trata de un fenómeno de la mente humana y su análisis plantea un problema de objetividad, debido a que “los signos pueden usarse de maneras ilimitadamente diferentes, ampliando o reduciendo su significado o modificándolo metafóricamente, los significados obtenidos varían de acuerdo al contexto, a los conocimientos y sensibilidad del lingüista que realiza el análisis” (Lara, 2006: 185).

La metodología

El proyecto al que pertenece este estudio es Corpus Multilingüe de Léxico Disponible en Español, Inglés, Francés, Portugués y Ruso del Centro de Enseñanza de Lenguas Extranjeras (CELE) de la Universidad Nacional Autónoma de México (UNAM). Como mencionamos, en él se analiza el léxico que viene rápida y fácilmente a la memoria del estudiante de lenguas extranjeras cuando se trata un tema específico, de ahí la conveniencia de basar su método de estudio en pruebas asociativas al modo de encuestas por tiempo (un minuto y medio) y por 16 centros de interés organizados en columnas, que en este estudio fueron: partes del cuerpo, ropa, alimentos, escuela, ciudad, campo, transporte, animales, distracciones, profesiones, colores, comunicaciones, adjetivos, sensaciones, vocabulario especializado y verbos (Frey, 2008). Estas encuestas se realizaron en una misma sesión en inglés y español en el año de 2011 con dos grupos de inglés de nivel intermedio del CELE; en total fueron 15 estudiantes cuyas listas serán incluidas en la base de datos del proyecto Ledical,⁶ que eventualmente integrará el diccionario de léxico disponible en lenguas extranjeras del CELE.

A su vez, estas encuestas estuvieron precedidas de otra sociolingüística en la que se contemplaron algunas variables también establecidas por el proyecto conjunto como: nombre, sexo, carrera, semestre, grupo, profesor y años de estudio del inglés. Así sabemos que todos los informantes son de la misma universidad, de distintas carreras; nueve hombres y seis mujeres en un rango de edades entre 22 y 26 años. Esta parte también fue de gran utilidad a la hora de analizar las encuestas, pues así logramos saber si determinadas palabras están presentes en un determinado grupo de personas.

Una vez recopilado el corpus de datos para su análisis, se clasificaron los vocablos en campos semánticos, lo cual nos permitió reconocer algunas de las asociaciones mentales, principales y secundarias, que se ponen en marcha en la respuesta de cada centro. Analizamos las listas escritas producidas por los alumnos en cuanto a número de palabras, categoría gramatical, campo y subcampo semántico. Aquí se hizo presente también la memoria y las condiciones de asociación verbal que están estrechamente relacionadas con el contexto y, por tanto, con la cultura. Así, se hicieron evidentes algunas cuestiones relacionadas con la edad,

⁶ Véase <<http://www.cele.unam.mx/ledical>>.

género, carrera a la que pertenecen, el entorno al cual tienen acceso, los dialectismos locales, extranjerismos, uso de jerga, interlengua, invenciones e influencias de una lengua sobre otra.

Diccionarios de alumnos universitarios hispanohablantes español vs. diccionarios de (inglés, francés, portugués y ruso)

Contraste en las encuestas de léxico disponible de español/inglés

El estudio de los múltiples aspectos socioculturales que aparecen en el léxico disponible de los alumnos pertenece al ámbito de la antropología lingüística y merece analizarse tanto por sus contenidos, como por las perspectivas de futuro que se vislumbran en nuevos hallazgos sobre cómo los alumnos aprenden una lengua y cómo construyen el conocimiento de las otras culturas a las que pertenece la lengua extranjera, en este caso el inglés.

En nuestro estudio español-inglés, el léxico disponible en lengua materna se organiza sobre la base de categorías y subcategorías semánticas evidentemente en mayor número, con más grado de precisión y especificidad que en inglés. Mientras que el léxico en lenguas extranjeras aparece muy relacionado con aspectos visuales que se presentan en los libros de texto con los que se ha realizado el aprendizaje. Por otro lado, hay ciertos centros donde aparecen diferencias de género visibles como es el de alimentos y distracciones. Mientras que otros, como comunicaciones y verbos, aparecen muy semejantes en las dos lenguas y en todos los alumnos con independencia de su género. Se observaron asociaciones de palabras de uso muy arraigado tanto en español como en inglés, por ejemplo, en el número y orden en que se mencionan los colores. Asimismo, se observó en ambas lenguas que en el centro de interés de los verbos aparecen en primer lugar los de acción y, en segundo, los de pensamiento. Hay otros centros donde prevalece algún elemento, como en el de animales, donde predominan en ambas lenguas los mamíferos. Y también existen muchos términos cuya justificación es sencilla, son los relacionados con el ámbito doméstico, escolar y personal.

Este estudio –aún no concluido– y los resultados presentados aquí son preliminares. Sin embargo, hoy los investigadores parecen estar de acuerdo con la idea de que el léxico disponible es un complejo conjunto de redes mentales y culturales que abarca una gran cantidad de palabras a las cuales se accede de manera muy

rápida y certera en el momento en que se requiere. Evidentemente conocer su naturaleza es de gran importancia para la enseñanza, pues cada vez que un alumno evoca una palabra necesariamente debe recorrer un camino que lo lleva a esta estructura mental para activar, seleccionar y, finalmente, producir la palabra.

Pudimos observar que el léxico se organiza de diferentes maneras según la lengua que se esté utilizando. Como vemos a continuación en la lengua materna se advierte una influencia del contexto determinante:

- Centro de interés *escuela*: maestros, alumnos, sillas, escritorios, pluma, goma, lápiz, sacapuntas, corrector. En inglés: *desk, eraser, blackboard, piece of chalk, pencils, pens, note-books, TV*.
- Centro de interés *campo*: animales, sol, trabajo, granjeros, vacas, burros, borregos, perros, gallina, establo, árboles, camión, sombrero, pala, pico, semillas, trigo. En inglés: *animals, plants, mountains, lakes, rivers, people, farm, machines, mushroom, trees*.
- Centro de interés *profesiones y oficios*: albañil, cartero, arquitecto, científico, ingeniero, abogado, doctor, profesor, vendedor, chofer, taxista, director, futbolista. En inglés: *football player, architect, scientist, lawyer, doctor, teacher, postman, researcher*.

En el caso de la lengua extranjera se evidencia el contexto escolar, y sabemos también que es necesario el reciclaje de ciertos elementos lingüísticos a partir de algunos errores que se repiten constantemente, aunque sabemos que estos errores pueden también ser lapsus por falta de atención, o bien derivados de la falta de conocimiento de la lengua. Asimismo, debemos considerar la dimensión personal y la competencia individual del uso de la lengua y desarrollar estrategias orientadas a la observación, reflexión y crítica. Finalmente, esperamos al término del estudio tener más encuestas incluidas con sus respectivas estadísticas.

Contraste en las encuestas de léxico disponible de español/portugués

En el análisis de las encuestas aplicadas a los aprendientes de portugués como lengua extranjera y cuya lengua materna es el español –estudiantes del CELE-UNAM– hemos observado que en la adquisición del léxico de portugués se presentan dos aspectos importantes que debemos considerar en el proceso de enseñanza-aprendizaje.

Por un lado, al ser el portugués una lengua próxima al español, un problema que los aprendientes y los docentes enfrentan es el *interlenguaje*. Veamos algunos ejemplos:

ESPAÑOL	PORTUGUÉS	PRODUCCIÓN DEL APRENDIENTE
Rodilla	Joelho	Rodela, goelho
Amarillo	Amarelo	Amarelho
Pizarrón	Quadro	Pizarrão
Abogado	Advogado	Advocado
Hoja	Folha	Folia

Por otro lado, la falta de una sistematización continua y eficiente del léxico no permite que los aprendientes, en su mayoría, logren contar con un verdadero ‘léxico disponible’ que puedan usar libremente en situaciones comunicativas específicas en las que ellos necesiten transmitir información sobre un tema determinado.

1. Español: Campo, aire, maíz, frijol, espacio, bosque, árbol, piñón, ganado, vaca, cerdo, rancho finca, caballo, siembra, sembrador, campesino, segar, lluvia, sequía, río, riachuelo, roca, vegetación, abierto, (24).
2. Portugués: árvores, tigres, elefantes, cão, galo, galinha, estornio, espaço, deserto, nuvens, chuva, março, sertão, povoado, povo, pó, vacas, boi, carneiro, cabra, chimarrão, (21)

En el caso de nuestros aprendientes universitarios, frecuentemente hemos notado que cuando ellos están cursando un nivel intermedio-avanzado de su aprendizaje, ya no hacen uso del léxico que aprendieron en los cursos previos. Un ejemplo de esto es que les resulta muy difícil llevar a cabo tareas como “describirse o describir otra persona tanto físicamente como psicológicamente” o “hablar de su rutina”.

Comentarios finales

Consideramos necesario: a) elaborar materiales didácticos diversos que aborden el léxico de múltiples centros de interés y de las diferentes áreas de estudio de los estudiantes universitarios; b) promover el uso de estrategias para la adquisición del léxico, tales como la utilización del contexto, la revisión de aspectos morfoló-

gicos, la función gramatical de la palabra en la frase y la construcción de campos semánticos; c) desarrollar actividades que integren el léxico aprendido a lo largo del curso de portugués. De esa manera, al concluir el curso, el aprendiente efectivamente podrá contar con el léxico disponible que le permita tener una alta competencia comunicativa en diferentes situaciones y en diferentes contextos sociales.

Cabe mencionar que los resultados obtenidos a lo largo del periodo que hemos trabajado en el proyecto Estudios del léxico disponible en el CELE, en las áreas de francés, inglés, portugués y ruso, los resultados han sido muy satisfactorios; sin embargo, debemos ampliar la aplicación de encuestas, continuar el análisis de los datos y elaborar diccionarios que apoyen a los aprendientes en la adquisición del léxico en las diferentes lenguas que se imparten en el Centro.

Referencias

- ARRIAGA CAMPOS, R. (2010). *Involución de la complejidad lingüística. Análisis de la complejidad lexico-semántica de estudiantes de primaria y secundaria*. Tesis de doctorado en lingüística. México: Facultad de Filosofía y Letras, UNAM.
- CALSAMIGLIA H. & A. TUSÓN (2002). *Las cosas del decir*. Barcelona: Ariel.
- CARCEDO, A. (2000). Índices léxico-estadísticos y graduación del vocabulario en la enseñanza de E/LE (aspectos culturales). En M. Franco Figueroa et al. (eds.), *Nuevas perspectivas en la enseñanza del español como lengua extranjera*. Cádiz: Universidad, pp. 175-183.
- DURANTI, A. (2000). *Antropología lingüística*. Madrid: Cambridge University Press.
- FREY PEREYRA, L. H. (2008). *El léxico disponible en los textos de alumnos de ELE. Estudio comparativo de dos corpora*. Tesis doctoral inédita. España: Universidad de Alcalá.
- FREY PEREYRA, L. H. (2009). *Corpus multilingüe de léxico disponible de los estudiantes del CELE. Proyecto posdoctoral*, Posgrado en Lingüística, UNAM/Conacyt.
- HERNÁNDEZ MUÑOZ, N. & J. BORREGO NIETO (2004). Cuestiones metodológicas sobre los estudios de disponibilidad léxica. En M. Villayandre (ed.), *Actas del V Congreso de Lingüística General*. Madrid: Arco/Libros.
- LARA, L. F. (2006). *Teoría del diccionario monolingüe*. México: El Colegio de México.
- _____ (2012). *Diccionario del español de México (DEM)* [en línea]. México: El Colegio de México. <<http://dem.colmex.mx>>. [Consulta: 10 noviembre de 2012.]

- LÓPEZ CHÁVEZ, J. (1992). Alcances panhispánicos del léxico disponible. *Lingüística*, 4, pp. 26-124.
- _____ & C. STRASSBURGER, (2000). El diseño de una fórmula matemática para obtener un índice de disponibilidad léxica confiable. En J. A. Bartol, S. Crespo, C. Fernández Juncal, C. Pensado y E. Prieto. *Dificultades de los estudios de disponibilidad léxica en ELE: los criterios de selección*, Anuario de Letras, pp. 38, 227-251.
- MICHÉA, R. (1953). Mots fréquents et mots disponibles. Un aspect nouveau de la statistique du langage. *Les Langues Modernes*, 47: 338-344.
- SAMPER, M. (2001). Dificultades de los estudios de disponibilidad léxica en ELE: los criterios de selección. Nuevas aportaciones al estudio de la lengua española. En *Investigaciones filológicas (277-286)*. Salamanca: Luso.
- SAMPER PADILLA, J. A. (1998). Criterios de edición del léxico disponible. *Lingüística*, 10, pp. 311-333.
- TABERNEO, C. (2008). El léxico disponible como fuente de aproximación al estudio de regionalismos. Actas del XXXVII Simposio Internacional de la Sociedad Española de Lingüística (SEL). Olza, I. (ed.). Moreno, Casado y González Ruiz. Departamento de Lingüística Hispánica y Lenguas Modernas. Pamplona: Publicaciones de la Universidad de Navarra.
- VAN EK, J. A. (1975). *The Threshold Level in a European Unit/Credit System for Modern Language Learning by Adults*. Estrasburgo: Publicaciones del Consejo de Europa.

Anexo 1

Como parte de los antecedentes de este trabajo se encuentra el proyecto de posdoctorado Corpus Multilingüe de Léxico Disponible de los Alumnos del CELE de la UNAM, a cargo de la Dra. Ma. Luisa Helen Frey Pereyra (+), con la asesoría de la Dra. Marisela Colín Rodea. Este proyecto –que contó con una beca del Conacyt (2009)– dio origen tanto al proyecto Estudios de Léxico Disponible, CBHC091127, cuyo responsable fue la Dra. Colín Rodea, como a varios proyectos adjuntos de profesores del CELE: María Noemí Alfaro Mejía, Barbara Lou Byer Clark, Eréndira Dolores Camarena Ortiz, María de los Ángeles Cervantes Amador, Leonardo Herrera González, Jorge Tapia, Patricia Cánovas Corrales. Además, durante el primer año se contó con la asesoría del Dr. Santiago Cortés Hernández y del Ing. James Mallory Frey, este último encargado del diseño de la base de datos del proyecto (Ledical). A fin de llevar a cabo esta propuesta, se trabajó dentro del Seminario de Estudios de Léxico Disponible. Cabe añadir que el proyecto cobijó la realización de una tesis de maestría de lingüística aplicada, elaborada por la alumna Anahí Ramírez Alfaro.

Anexo 2

Ledical es un programa, actualmente en desarrollo, elaborado por James Mallory Frey. Aplicando la fórmula de López Chávez y Strassburguer (2000), calcula la disponibilidad léxica y, en caso de que varios vocablos cuenten con la misma disponibilidad, los ordena alfabéticamente. En la actualidad este programa se está ampliando para que realice también el cálculo de índices de cohesión y pueda ofrecer la comparación de datos de acuerdo con distintas variables. *Ledical* tiene la particularidad de no pertenecer a ningún proyecto de investigación, está en línea de forma gratuita y es de código libre, además de que funciona con cualquier sistema operativo compatible con Java, tales como Mac OS X, Linux, Solaris o Windows.