

DISPONIBILIDAD LÉXICA DE 52 ESTUDIANTES TAIWANESES UNIVERSITARIOS DE ELE A1

BIODATA

Miguel Rubio Lastra, miguelrubio@thu.edu.tw, profesor y coordinador del programa de Estudios hispánicos, Departamento de Literatura y lenguas extranjeras, Tunghai University. Sus investigaciones se enmarcan dentro de la lingüística aplicada a la enseñanza y aprendizaje ELE, centrándose principalmente en el análisis conversacional y la enseñanza de la competencia oral. Desde 2017, ha dirigido su labor investigadora hacia la disponibilidad léxica de los alumnos de ELE. Además, es editor en Ediciones Catay, así como en las revistas *Interplay* y *Encuentros en Catay*.

RESUMEN

A lo largo del proceso de aprendizaje de una lengua, el alumno va completando diferentes etapas que pueden ser definidas por la riqueza, el control y el alcance del vocabulario. La relevancia de estos parámetros en la determinación del nivel competencial del estudiante es el principal motivo de este trabajo, cuyo objetivo general es analizar con detalle el vocabulario que conoce un grupo de 52 alumnos taiwaneses con un nivel A1 de español. Para ello, hemos aplicado un test de disponibilidad léxica conforme a la metodología practicada por los estudios que surgieron en Francia dentro del marco del *Centre d'étude du français élémentaire* a mediados del siglo XX.

El análisis de los resultados del test servirá como punto de partida para enfrentar futuras investigaciones que aborden la suficiencia o insuficiencia del léxico de los alumnos, así como la eficacia o ineficacia de los materiales de aprendizaje e, incluso, de la metodología de enseñanza seguida en el aula.

PALABRAS CLAVE: léxico, disponibilidad léxica, alumnos taiwaneses, nivel A1.

LEXICAL AVAILABILITY OF 52 TAIWANESE STUDENTS WITH ELE A1

Throughout the process of learning a language, the student will complete different stages which can be characterized by the richness, control and scope of the vocabulary. The relevance of these parameters in the determination of the competence level of the student is the main reason for this work, whose main objective is to analyze in detail the lexicon of a group of 52 Taiwanese students with an A1 level of Spanish. We have applied a lexical availability test according to the methodology that was practiced by the studies that arose in France within the framework of the *Centre d'étude du français élémentaire* in the middle of the 20th century.

Analysis of the results will serve as a starting point to face future research that might address the adequacy or inadequacy of the lexicon of the students, as well as the effectiveness or ineffectiveness of the learning materials and, even, of the teaching methodology followed in the classroom.

KEY WORDS: lexicon, lexical availability, Taiwanese students, ELE A1 level of language proficiency.

Resulta lógico que todo profesor de español como lengua extranjera (ELE) se preocupe de que sus alumnos alcancen un nivel de comunicación suficiente para ir completando las diferentes etapas de aprendizaje de la lengua. En nuestra experiencia como profesores de alumnos taiwaneses durante más de diez años, hemos venido observando que, independientemente del nivel lingüístico que estos alcancen, los propios alumnos se quejan de una "falta de vocabulario" que les impide expresar lo que realmente quieren decir¹. Esto hizo que en el mismo contexto de aprendizaje nos planteáramos la posibilidad de investigar al respecto del vocabulario que conocen nuestros alumnos y de si realmente resulta tan importante en su comunicación.

Esta importancia queda manifiesta en el *enfoque léxico* desarrollado por Lewis (1993), quien sitúa el vocabulario o léxico en una posición crucial para el aprendizaje de una lengua, o en el Marco Común Europeo de Referencia para las Lenguas (MCERL) del Consejo de Europa (2002) donde se entiende la competencia léxica como una de las competencias lingüísticas a tener en cuenta en la enseñanza-aprendizaje de las lenguas:

"La riqueza, el alcance y el control del vocabulario son parámetros importantes en la adquisición de la lengua y, por ello, de la evaluación del dominio de la lengua que tiene el alumno, y de la planificación del aprendizaje y de la enseñanza de lenguas." (Consejo de Europa, 2002: 86)

¹ Ten Have (1999:24) concede validez a estas observaciones que hace el profesor y que él denomina *general background observations*

² Daremos cuenta de estas investigaciones a lo largo del trabajo.

³ Véanse, por ejemplo, los monográficos publicados en las revistas de sendas universidades: V Jornadas de Lengua y Comunicación. "Léxico: Enseñanza e

En este mismo sentido, también se nos dice en el MCERL que "quienes elaboran materiales de exámenes y manuales están obligados a hacer una selección léxica" (Consejo de Europa, 2002: 149). Y en esta línea se manifiesta el Plan Curricular que ha desarrollado el Instituto Cervantes (PCIC) (Instituto Cervantes, 2006), o las investigaciones relacionadas con el léxico que se han publicado en los últimos años². Nos referimos, concretamente, a los estudios sobre *frecuencia léxica* y sobre *disponibilidad léxica* (DL) Estos últimos están pasando por un momento importante, tal y como han recogido los trabajos recopilatorios y expositivos de Moreno (2012), Paredes (2012, 2014), González (2014), López González (2014) o Jiménez Catalán (2014 y 2017). En general, la disponibilidad léxica, en cuanto a disciplina de investigación surgida en Francia a mediados del siglo XX, trata de dar cuenta del léxico que un sujeto puede emplear de forma inmediata y natural cuando es necesario (Micheá, 1953).

La investigación que presentamos aquí sigue esta línea, pues versa sobre el léxico que conocen los alumnos, el léxico disponible, y toma como referencia los numerosos trabajos que tienen como objeto de estudio el español como LE y que han sido llevados a cabo durante estos últimos veinte años en España³. En particular, tendremos presentes dos investigaciones, la pionera de Carcedo González (2000) sobre la disponibilidad léxica de un amplio grupo de

Investigación", coordinado por Martí Sánchez y Penadés Martínez y aparecido en *LinRed* (2015); el número 11 de la *Revista Nebrija de Lingüística Aplicada a la Enseñanza de las Lenguas* (2012); o el libro que dedica Jiménez Catalán (2014) a esta cuestión; así como las numerosas tesis doctorales defendidas estos últimos veinte años y a las que nos referiremos en el marco teórico.

estudiantes finlandeses de ELE, y la de Gallego Gallego (2014) sobre la disponibilidad léxica de los extranjeros residentes en Madrid. A estas dos referencias sobre DL en ELE hemos de añadir la recién defendida tesis doctoral de Hidalgo Gallardo (2017b), puesto que, además, tiene como objeto de estudio una amplia población de estudiantes sinohablantes chinos de ELE, lo que convierte la lengua materna de su población de estudio en denominador común de la nuestra.

En nuestro caso en particular y de acuerdo a las observaciones que hemos hecho en el aula, propondremos como OBJETIVO GENERAL descubrir cuál es el vocabulario disponible de los alumnos taiwaneses de ELE. Hemos querido comenzar a trabajar con un grupo de 53 alumnos de un nivel A1 porque nuestra intención es continuar con los siguientes niveles lingüísticos en investigaciones futuras y, si los resultados son satisfactorios, ampliarlo a otros grupos como objeto de estudio. Si bien, no hemos encontrado ningún estudio sobre la disponibilidad léxica en alumnos de un nivel A1, en cambio, sí contamos con tres investigadores en China: Lin (2006 y 2012) sobre la disponibilidad léxica de 263 informantes chinos; Jia Chen (2016), con su tesina de máster sobre la disponibilidad léxica de estudiantes chinos en comparación con estudiantes sudafricanos; e Hidalgo Gallardo (2017a y 2017b), cuya tesis doctoral se centra en la disponibilidad léxica de 440 informantes de seis universidades chinas y su comparación con la de los nativos de español, con la de otros estudiantes de ELE y con los resultados de la tesis de Lin (2006). Asimismo, este último investigador contrasta los resultados de su trabajo con el léxico aparecido en dos manuales de ELE (*Español Moderno 1* y *Bitácora 1*). Estos trabajos, si bien ofrecen algunos resultados muy interesantes, especialmente los de Hidalgo Gallardo, no llegan a describir el léxico disponible de los alumnos en su primera etapa de aprendizaje. A ello, hay que añadir que, en nuestro caso,

trabajamos con alumnos de Taiwán y, por lo tanto, se trata de una población con características culturales diferentes, como veremos llegado el momento de la comparación en trabajos futuros.

Por ello, vista la necesidad de realizar un posterior estudio paralelo a estas otras investigaciones y con la intención de comparar los resultados en el futuro, comenzamos planteando una PREGUNTA DE INVESTIGACIÓN para este primer trabajo:

¿Cuál es el léxico disponible de 52 alumnos taiwaneses de ELE en un nivel A1 y qué características presenta?

Para hallar respuesta a esta pregunta se precisa, en primer lugar, alcanzar ciertos objetivos durante el desarrollo de la investigación. Dichos OBJETIVOS, basándonos en los habitualmente presentados en los estudios sobre DL, serán planteados del siguiente modo:

1. Anotar la cantidad absoluta y relativa de las respuestas de los alumnos.
2. Describir cuantitativa y cualitativamente el rango que ocupan las respuestas.
3. Destacar las respuestas más prominentes según su frecuencia y su índice de disponibilidad.
4. Averiguar cuáles son los centros de interés más productivos según el número de respuestas dadas en cada uno de ellos.
5. Averiguar qué centros presentan una mayor riqueza léxica.

6. Cuantificar la compactibilidad de los centros de interés según el *Índice de cohesión* y el *Índice de densidad*.

7. Averiguar si la variable sexo incide de forma significativa en la disponibilidad léxica de los estudiantes.

Sobre el aprendizaje de vocabulario, parece ampliamente aceptado por la comunidad científica que lo lógico es enseñar las unidades léxicas más frecuentes a los alumnos de LE, ya que aparecen en una mayor cantidad de textos orales y escritos (Alvar, 2004; Davies, 2006, Almela *et al.*, 2005, entre otros), si bien, teniendo en cuenta el concepto de *utilidad*, tal y como explica Alonso (2012), no hay una relación proporcional entre utilidad y frecuencia. En futuros trabajos quizás sea necesario establecer un inventario de unidades léxicas más frecuentes y, al mismo tiempo, más apropiadas para los alumnos taiwaneses. Alvar Ezquerro (2004) presentó en el XV congreso ASELE un excelente trabajo al respecto de la utilidad de la frecuencia léxica en la enseñanza ELE. Volveremos a este asunto una vez tengamos una idea más exacta del conocimiento léxico que poseen los alumnos taiwaneses, así como del proceso de adaptación del mismo.

Con el fin de facilitar futuros estudios comparativos con otras comunidades de informantes, hemos adoptado los principales CRITERIOS METODOLÓGICOS del Proyecto panhispánico sobre disponibilidad léxica (PPDL). De todas formas, al igual que ha venido haciendo la mayoría de los estudios sobre estudiantes de ELE⁴, se han realizado algunas modificaciones debido a las singularidades que presenta la población de nuestro estudio, tal y como explicaremos en el apartado sobre metodología. Trabajaremos, por lo tanto, con la

⁴ Las dos publicaciones que con este fin hemos tomado como principales referentes son Samper Hernández (2002) y Gallego Gallego (2014).

terminología que emplean las principales investigaciones publicadas en España sobre producción léxica en el ámbito de la enseñanza de ELE.

Los resultados obtenidos en este trabajo permitirán dar una explicación detallada del léxico que conocen los alumnos taiwaneses de ELE en su primera etapa de aprendizaje. Esto resultará fundamental para futuras investigaciones que aborden la suficiencia o insuficiencia de este léxico o sobre la eficacia o ineficacia de los materiales de aprendizaje e, incluso, de la metodología de enseñanza seguida en el aula.

1. MARCO TEÓRICO

1.1. EL LÉXICO

Aunque ya en la Introducción aparecían los dos conceptos principales que vamos a manejar, el *léxico* y la *frecuencia léxica*, es necesario definir, en primer lugar, qué entendemos por *léxico*. Al hablar de léxico, obviamente, no nos referimos únicamente a *palabras*, sino más bien a *unidades léxicas*, ya que estas pueden estar formadas por un grupo de palabras inseparables que poseen un significado único (v.g. *de maravilla*, *traer sin cuidado*, *echar una mano*, etc.). Así, para Gómez Molina (2004: 497), la unidad léxica es "la unidad de significado en el lexicón mental⁵, que sirve como elemento vehiculador de la cultura del español y puede estar formada por una o más palabras (*cabeza*, *paraguas*, *dinero negro*, *a la chita callando*, *tomar el pelo*, *no hay más cera de la que arde*, etc.)". Cuando las

⁵ El *lexicón mental*, según el Instituto Cervantes en su *Diccionario de términos clave de ELE*, es "el conocimiento que un hablante tiene interiorizado del vocabulario".

unidades léxicas están formadas por una sola palabra, hablamos de *unidades léxicas simples* (*cabeza, paraguas...*). En cambio, si las unidades léxicas están formadas por dos o más palabras, hablaremos de *unidades léxicas complejas* (*tomar el pelo, no hay más cera de la que arde...*).

Si bien, en ciertas disciplinas lingüísticas pueden existir algunos matices que distinguen léxico de vocabulario, en cuanto a que el primero incluye al segundo, tal y como señala Gómez Molina (2004: 497), nosotros seguiremos lo apuntado por el Instituto Cervantes en su *Diccionario de términos clave de ELE*, donde se indica que "en la didáctica de la lengua, los términos vocabulario y léxico son equivalentes"⁶.

Tal y como hemos podido comprobar en los trabajos que se han consultado sobre las unidades léxicas, estas presentan una diversidad de características que permiten establecer distintas clasificaciones. En este caso, acudimos a la ofrecida por el Consejo de Europa (2002), pues nos parece la más apropiada al contexto en el que se desarrolla nuestra investigación. Así, según el Consejo de Europa en su obra *Marco Común Europeo de Referencia de las Lenguas* (MCERL) (2002: 108), los elementos léxicos comprenden: expresiones hechas (fórmulas fijas, modismos, estructuras fijas y otras frases hechas) y polisemia.

Si bien, en esta investigación no pretendemos clasificar las unidades léxicas extraídas del corpus de datos, sí resulta necesario mantener esta clasificación como referencia ante posibles alusiones terminológicas que vayan apareciendo.

⁶ Podemos encontrar esta cita en la versión electrónica del *Diccionario de términos clave de ELE* del Instituto Cervantes:

1.2. LA DISPONIBILIDAD LÉXICA

El concepto de disponibilidad léxica (DL) tiene su origen formal como objeto de estudio dentro del marco del *Centre d'étude du français élémentaire* a mediados del siglo XX, concretamente tras la publicación de Micheá en el año 1953 de *Mots fréquents et mots disponibles. Un aspect nouveau de la statistique du langage*. Los artífices de esta línea de investigación, Gougenheim, Michéa y Sauvageot (1956), trataban de facilitar el aprendizaje del francés en las colonias francesas de entonces gracias al conteo de las palabras que los nativos franceses empleaban con más frecuencia para así poder elaborar un listado que incluir en los materiales de enseñanza del francés como lengua extranjera.

En el mundo hispánico, los estudios de disponibilidad comenzaron en Hispanoamérica y el primero de ellos fue el de López Morales en 1973. En este primer trabajo, de corte sociolingüístico, López Morales trata de determinar las variables extralingüísticas que influyen en el proceso de adquisición del vocabulario de 63 estudiantes de la zona metropolitana de San Juan, Puerto Rico.

A partir de este primer trabajo, las investigaciones sobre DL se han ido expandiendo a ambos lados del Atlántico a través de las investigaciones incluidas en el Proyecto Panhispánico. Sus aplicaciones son numerosas y suelen destacar las comparaciones interdialectales de unas zonas geográficas con otras. También resultan de gran interés los resultados concernientes a la información

http://cvc.cervantes.es/ensenanza/biblioteca_ele/diccio_ele/diccionario/vocabulario.htm

sociolingüística extraída, como por ejemplo las posibles diferencias entre grupos de diferentes niveles socioeconómicos. Entre otras utilidades⁷, podemos mencionar la información que se extrae para el estudio de los procesos psicolingüísticos de producción.

De todas formas, las publicaciones que hemos ido recogiendo y que aparecen citadas a lo largo de este trabajo apuntan a que la aplicación más reconocida de los estudios de disposición es la mejora de la enseñanza de lenguas, tanto en lengua materna como en lengua extranjera. Nuestro estudio se sitúa en esta línea, que fue la que dio origen a los estudios de DL; concretamente, en la mejora de la enseñanza del español como LE. De este modo, el presente trabajo servirá para conocer con más detalle la competencia léxica de nuestros alumnos, entendida esta como "el conocimiento del vocabulario de una lengua y la capacidad para utilizarlo" (Consejo de Europa, 2002: 108), procederemos a un estudio de disponibilidad léxica.

En los últimos veinte años, han tenido cierta repercusión en el español como lengua extranjera (ELE) gracias a las investigaciones pioneras sobre informantes extranjeros de Carcedo González desde 1998⁸. Años más tarde, en 2002, se publica el trabajo de Samper Hernández sobre una muestra de 45 estudiantes de ELE, donde aparecen recogidos los criterios metodológicos y de edición de los materiales que más tarde han tomado como referencia otros investigadores de este mismo campo.

⁷ En el trabajo de Sánchez-Saus Laserna (2011) o en el de Gallego Gallego (2014) puede encontrarse una exposición muy completa de la trayectoria de los estudios de disponibilidad desde sus orígenes.

⁸ En su tesis doctoral, Gallego Gallego (2014) describe los trabajos más destacados de Carcedo González y de otros pioneros de la disponibilidad en ELE como Samper Hernández. Además, para tener una idea más amplia sobre las investigaciones de

Las investigaciones que han trabajado con el léxico disponible del español como LE han seguido objetivos similares⁹. Esto es: comparación con el léxico disponible de hablantes nativos o con manuales de ELE; descripción de la evolución de la DL según el nivel de español, según la lengua materna o según el método de enseñanza seguido; clasificación de las respuestas en función de su naturaleza gramatical o del tipo de error que representan, etc.

Nuestro trabajo también sigue esta línea, puesto que pretende aportar más información sobre la disponibilidad léxica de los estudiantes de ELE gracias al análisis de la producción de un grupo de alumnos taiwaneses de nivel A1 (MCER). Con los resultados obtenidos en esta primera aproximación al léxico disponible de los alumnos, podremos meternos de lleno en la comparación con los resultados de otros trabajos, en la introducción de variables como el nivel de español de los sujetos o de la utilización de diferentes manuales de aprendizaje.

Estas investigaciones, si bien ofrecen algunos resultados muy interesantes con respecto a los estudiantes chinos, especialmente los de Hidalgo Gallardo (2017a y 2017b), no llegan a describir el léxico disponible de los alumnos en su primera etapa de aprendizaje. A ello, hay que añadir que, en nuestro caso, trabajamos con alumnos de Taiwán y, al tratarse de una población con características socio-culturales diferentes, intuimos unos resultados diferentes con

DL aplicadas a la enseñanza ELE que se han llevado a cabo a lo largo de estos últimos años, destacamos los trabajos de Moreno (2012), Paredes (2012), González (2014) o López (2014).

⁹ Véase en Tomé Cornejo (2015: 338-339) un recopilatorio de muchos de estos estudios.

respecto a los chinos. Quizás en investigaciones futuras podamos abordar este interesante asunto; pero centrémonos por ahora en la cuestión que nos ocupa.

2. CARACTERIZACIÓN Y METODOLOGÍA DE LA INVESTIGACIÓN

Adelantábamos en la introducción que esta investigación sigue el mismo criterio metodológico que el Proyecto Panhispánico, tal y como han venido haciendo casi todos los trabajos que en nuestra lengua se han publicado a lo largo de estos últimos años y que en su mayor parte hemos reseñado en los apartados anteriores. De todos modos, cuestiones como la nominalización de los centros de interés han sido ligeramente adaptadas a las condiciones específicas de nuestra investigación. En los subapartados que aparecen a continuación daremos una explicación más detallada al respecto.

2.1. LA MUESTRA (INFORMANTES)

Las 52 encuestas de DL que forman el grupo de estudio son el resultado de la selección de una *población* de 61 estudiantes taiwaneses. El día que recogimos los datos, se presentaron 55 estudiantes. Luego, durante la edición de los materiales, fue necesario hacer una selección de individuos que garantizase la

¹⁰ Hemos calculado un nivel A1 ya que todos los alumnos, en el momento de la encuesta, habían trabajado ya con el manual ELE Actual A1, de la editorial SM, y estaban estudiando la unidad 4 del manual ELE Actual A2. Además, la mayoría de estos alumnos se presentaron en mayo al DELE A1. A pesar de que no poseemos cifras exactas al respecto, sabemos por comunicación personal con los examinados de otros años que todos los presentados a este examen lo aprueban.

homogeneidad de la muestra. Esta selección será la que constituya la *muestra productora de datos*. El criterio fundamental que seguimos fue que los individuos garantizaran un nivel lingüístico parejo en la lengua meta. En nuestro caso, se ha estimado un A1 según los parámetros establecidos por el MCERL¹⁰. Y para darle una consistencia aún mayor, hemos tenido en cuenta que las calificaciones obtenidas en el último semestre superen los 60 puntos¹¹. Esta nueva acotación nos llevó a desestimar 3 de los test que recogimos: el primero de los informantes es hablante nativo de español; el segundo de ellos ha estado en España durante 2 meses y estimamos un nivel cercano al B1; y el tercero no había llegado a los 60 puntos en el semestre pasado, por lo que consideramos que no alcanzaba un nivel A1.

Se trata, por lo tanto, de un *muestreo intencional*, ya que todos están cursando el tercer año de español en la Universidad Tunghai y poseen un nivel de español similar.

Una vez controlada la variable del nivel de español, a través de una *encuesta de datos personales*, establecimos otras características del grupo. De esta manera, pudimos averiguar la edad, el sexo, el lugar de nacimiento, la lengua materna, el conocimiento de otras lenguas y la estancia en países de habla española. El resultado de esta encuesta aparece resumido en el siguiente cuadro.

¹¹ En Taiwán, la nota máxima en las calificaciones es de 100 puntos, y el aprobado se sitúa en los 60 puntos. De esta forma, quedarían fuera del estudio aquellos estudiantes que no hayan superado los criterios mínimos para aprobar la asignatura el semestre pasado.

Tabla 1. Resumen de las características del grupo de estudio.

| VARIABLES | VARIANTES |
|---|--|
| Edad | 1. 21 años: 10 2. 22 años: 42 |
| Sexo | 1. Hombres: 6 2. Mujeres: 46 |
| Lugar de nacimiento | Taiwán |
| Lengua materna | Chino |
| Otras lenguas conocidas y nivel lingüístico | 1. Inglés: inicial (2), intermedio (34) y avanzado (17). 2. Japonés: inicial (8), intermedio (1) y avanzado (1). 3. Otros: francés inicial (1), alemán inicial (1), indonesio avanzado (1), malayo avanzado (2). |
| Estancia en países de habla española | 1. Nunca: 49 2. Turismo: 3 |
| Manual de aprendizaje utilizado | ELE Actual A1, SM y ELE Actual A2, SM |
| Nivel de español | A1 |

Tenemos, por lo tanto, un grupo de estudio formado por 52 individuos taiwaneses de L1 chino (6 hombres y 47 mujeres) y con un dominio satisfactorio del inglés. Además, alrededor de un 20% de ellos conoce al menos una tercera lengua extranjera. La edad de los sujetos es de 21 y 22 años, estudian en la universidad y todos han empleado el mismo material didáctico para aprender español, alcanzando un nivel de A1. Ninguno de ellos ha aprendido español en un contexto diferente de la universidad y, además, los 3 individuos que han estado en un país de habla española fue por turismo y no permanecieron allí más de 15 días.

¹² Véase en los Anexos la plantilla que se les entregó a los alumnos.

¹³ Tal y como señala Moreno (2012) la denominación *centros de interés* puede ser sustituida por alternativas como *temas del discurso*, *temas predilectos*, *dominios de experiencia*, *áreas temáticas*, *estímulos verbales*, *campos asociativos* o *situaciones*

Si bien, en el presente trabajo todas estas variables y sus posibles variantes serán empleadas como control, consideramos necesario dar cuenta de ellas por si en investigaciones futuras quisiéramos realizar estudios contrastivos con estos datos.

2.2. LA PRUEBA

La prueba que realizaron los alumnos consta de dos partes. En primer lugar, se les entregó una *hoja de datos* que deberían rellenar con la información personal que hemos comentado en el subapartado anterior. Una vez completada, se recogió y se entregó la segunda parte de la prueba, el *test de datos*¹².

El test de datos está formado por 5 páginas divididas en 3 columnas cada una. Cada columna corresponde a un *centro de interés o tema*¹³ y está dividida a su vez en 29 filas numeradas, que es donde se anotarán las unidades, para así facilitar el conteo de la posición de la unidad léxica dentro del total de respuestas de cada columna, aspecto que resulta fundamental para calcular la disponibilidad, tal y como señala Samper Hernández (2002: 21). Siguiendo lo apuntado por esta autora y por los trabajos realizados en el mundo hispánico, decidimos que las listas fueran abiertas para que el informante pudiera escribir todas las palabras que se le ocurrieran, aunque no estén seguros de cómo se escriben. De todas formas, limitamos el

comunicativas. En nuestra investigación, emplearemos *temas* como alternativa a *centros de interés*, puesto que es el que aparece en el PCIC, objeto de comparación en futuras investigaciones.

tiempo en 2 minutos para cada centro de interés¹⁴. De este modo, obtenemos qué palabras son las más disponibles y cuántas son.

El encabezamiento de cada columna consiste en un número y en un espacio en blanco donde debería ir el nombre del centro de interés. La decisión de no escribir con antelación el nombre de cada centro de interés fue que los alumnos agotasen los 2 minutos que tenían para escribir todas las unidades que les vinieran a la cabeza (verbos, nombres, adjetivos, expresiones...) relacionadas con el tema en cuestión. De este modo, tratamos de evitar que comenzaran a escribir en la siguiente columna antes del tiempo establecido.

La gestión del tiempo se llevó rigurosamente, empleándose un total de 30 minutos (2 minutos por cada centro de interés).

Las instrucciones fueron muy concretas y se dieron en su lengua materna. Antes de dar comienzo a los 2 minutos de rigor para cada centro de interés, se escribió el nombre de este en la pizarra y se tradujo al chino y al inglés. Una vez agotados los 2 minutos, se escribió el nombre del siguiente tema y así hasta completar los 15 temas.

Al comienzo de la prueba, se señalaron tres cuestiones relativas al test de datos: la primera, que no debían escribir nombres propios; la

¹⁴ Sobre el tiempo dado por centro de interés en los trabajos realizados con hablantes de ELE, Samper Hernández (2002: 21) apunta que: "dos minutos era un tiempo suficiente para escribir bastantes unidades e incluso para bucear en su memoria en busca de alguna más". Por este motivo, no creímos necesario dar más tiempo para que el alumno extrajese más unidades léxicas de su memoria.

¹⁵ Destacamos dos títulos muy concretos al respecto que fueron publicados en el mismo número de la Revista Nebrija de Lingüística Aplicada a la Enseñanza de Lenguas: *A vueltas con la selección de centros de interés en los estudios de disponibilidad léxica: para una propuesta renovadora a propósito de la disponibilidad léxica en ELE*, de Paredes García (2014), e *Idoneidad de los centros*

segunda, que si antes de terminar los 2 minutos, deseaban volver a algún tema anterior para corregirlo o añadir más unidades léxicas, podían hacerlo; por último, que podían escribir detrás de la hoja si les faltaba espacio.

En cuanto a los centros de interés que decidimos incluir en el estudio, veamos en el siguiente subapartado qué decisiones tomamos.

2.3. CENTROS DE INTERÉS

Con la intención de satisfacer el justificado interés de los trabajos de revisión en torno al número y los tipos de centros de interés que se consideran en los estudios de disponibilidad léxica¹⁵, pasamos a justificar la decisión que nosotros tomamos sobre ello.

En los comienzos de la DL, Gougenheim *et al.* (1956) propusieron 16 centros de interés que son los adoptados por el Proyecto Panhispánico de Disponibilidad Léxica (PPDL) y, por lo tanto, por la mayoría de las investigaciones de DL en español. Estos 16 centros de interés son los siguientes:

- 1) las partes del cuerpo

de interés clásicos en los estudios de disponibilidad léxica aplicados al español como lengua extranjera, de González Fernández (2014), si bien este último, tal y como reseña Bartol (2014) en el mismo número, no utiliza un método adecuado para cumplir con el objetivo que se extrae del título del artículo. Además de estos dos trabajos, encontramos también referencias detalladas en Bartol (2006), Samper Padilla y Samper Hernández (2006), Moreno (2012), Paredes (2012) y López González (2014), cuyo principal objeto es la revisión de muchas de las investigaciones publicadas sobre DL en español; o en Jiménez Catalán (2017), que incluye, además, algunas investigaciones sobre DL en inglés L1 y L2.

- 2) la ropa (sin importar que sea ropa de hombre o de mujer)
- 3) la casa (sin los muebles)
- 4) los muebles de la casa
- 5) los alimentos y las bebidas en las comidas (todas las comidas del día)
- 6) los objetos colocados sobre la mesa que se utilizan para todas las comidas del día
- 7) la cocina, sus muebles y los utensilios que en ella se encuentran
- 8) la escuela, sus muebles y su material escolar.
- 9) la calefacción y la iluminación
- 10) la ciudad
- 11) el pueblo
- 12) los medios de transporte
- 13) los trabajos del pueblo y del jardín
- 14) los animales

15) los juegos y las distracciones

16) los oficios (los diferentes oficios y no nombres que se refieren a un solo oficio).

El propósito de utilizar unos centros de interés en particular para extraer las palabras que el informante conoce es servir de estímulo asociativo gracias a unas áreas temáticas universales que cualquier individuo pueda abarcar, si bien estos criterios de coherencia y universalidad son discutibles, tal y como recuerda Paredes García (2014: 2). Y es que, si bien, trabajar con un mismo modelo de encuesta y unos mismos centros de interés ofrece la "ventaja de servir de criterio único para una multiplicidad de equipos" (Moreno, 2012: 2), estos han sido cuestionados desde las primeras investigaciones sobre DL aplicada a la enseñanza de ELE. Así, por ejemplo, encontramos algunos centros de interés que son más sensibles a los cambios cronológicos, de lugar y de estrato social, como la ropa o los juegos (Paredes, 2014). Estos factores hacen que el investigador no sólo tienda a adaptar los centros de interés tenidos en cuenta en la propuesta inicial de Gougenheim *et al.* (1956), sino que es de gran importancia incluir aquellos que el investigador crea necesarios para alcanzar los objetivos propuestos en su investigación, tal y como observa Herreros Marcilla (2015a), entre otros. Por este motivo, nosotros adaptaremos los centros de interés a nuestra conveniencia.

Así, aunque en los trabajos de algunos investigadores se han incluido otros temas o centros de interés¹⁶ que no son los utilizados

¹⁶ Tal y como señala Moreno (2012) la denominación *centros de interés* puede ser sustituida por alternativas como *temas del discurso*, *temas predilectos*, *dominios de experiencia*, *áreas temáticas*, *estímulos verbales*, *campos asociativos* o *situaciones*

comunicativas. En nuestra investigación, emplearemos *temas* o *centros de interés* indistintamente, que son los más empleados habitualmente.

tradicionalmente, nosotros hemos querido tomar como base los aparecidos en el *Marco Común Europeo de Referencia de las Lenguas*, y renovados por el *Plan Curricular del Instituto Cervantes* (PCIC) (IC, 2006) puesto el inventario de entradas recogido tiene el "objetivo de dar una idea del tipo de unidades léxicas que el hablante debe conocer en cada uno de los niveles" (IC, 2006) y es aquí donde hasta ahora mejor aparecen desarrollados los niveles de referencia para el español según las recomendaciones propuestas en el MCERL. Al mismo tiempo, nuestro objetivo en trabajos futuros será comparar la DL de nuestros estudiantes con la propuesta hecha por el IC (2006). Veamos, en primer lugar, cuáles son los temas o *nociones específicas* empleadas en el PCIC:

- 1) INDIVIDUO, DIMENSIÓN FÍSICA: partes del cuerpo, características físicas, etc.
- 2) INDIVIDUO, DIMENSIÓN PERCEPTIVA Y ANÍMICA: carácter y personalidad, sentimientos y estado de ánimo, etc.
- 3) IDENTIDAD PERSONAL: datos personales, documentación, etc.
- 4) RELACIONES PERSONALES: relaciones familiares, relaciones sociales, etc.
- 5) ALIMENTACIÓN: dieta y nutrición, bebida, etc.
- 6) EDUCACIÓN: centros e instituciones educativas, profesorado y alumnado, etc.
- 7) TRABAJO: profesiones y cargos, actividad laboral, etc.
- 8) OCIO: tiempo libre y entretenimiento, deportes, etc.

9) INFORMACIÓN Y MEDIOS DE COMUNICACIÓN: correspondencia escrita, teléfono, etc.

10) VIVIENDA: acciones relacionadas con la vivienda, características de la vivienda, etc.

11) SERVICIOS: servicio postal, servicio de transporte, etc.

12) COMPRAS, TIENDAS Y ESTABLECIMIENTOS: lugares, personas y actividades, pagos, etc.

13) SALUD E HIGIENE: salud y enfermedades, síntomas, etc.

14) VIAJES, ALOJAMIENTO Y TRANSPORTE: tipos de viajes, red de transportes, etc.

15) ECONOMÍA E INDUSTRIA: finanzas y bolsa, renta, etc.

17) CIENCIA Y TECNOLOGÍA: cuestiones generales, biología, etc.

18) GOBIERNO, POLÍTICA Y SOCIEDAD: vida en comunidad, ley y justicia, etc.

19) ACTIVIDADES ARTÍSTICAS: música y danza, literatura, etc.
RELIGIÓN Y FILOSOFÍA.

20) GEOGRAFÍA Y NATURALEZA: universo y espacio, fauna, etc.

Al comparar las nociones específicas con los dieciséis centros de interés del PPDL, encontramos que, si quisiéramos tener en cuenta las primeras en un trabajo sobre DL, estas "tienen la ventaja de tener

muchos puntos en común con las áreas temáticas del proyecto inicial de disponibilidad léxica" (Paredes, 2014: 4).

Por lo tanto, consideramos que emplear las nociones específicas del PCIC en nuestra investigación aportará coherencia a nuestro objetivo. De todas formas, no hemos querido perder la referencia de los temas propuestos por el Proyecto Panhispánico porque también tenemos la intención de comparar nuestros resultados con los que puedan obtener otros investigadores. De este modo, hemos intentado utilizar unos centros de interés válidos para comparar con las nociones específicas del PCIC y que, al mismo tiempo, puedan servirnos para comparar nuestros resultados con los de otras investigaciones. Veamos, en concreto, la siguiente tabla de correspondencias entre los centros de interés que vamos a utilizar en nuestra investigación, las nociones específicas del PCIC y los centros de interés del Proyecto Panhispánico.

Tabla 2. Comparativa entre los centros de interés tradicionalmente empleados, los empleados en nuestra investigación y las nociones específicas del PCIC.

| CENTROS DE INTERÉS DEL PROYECTO PANHISPÁNICO | CENTROS DE INTERÉS DE LA ENCUESTA | NOCIONES ESPECÍFICAS DEL PCIC |
|--|---------------------------------------|---|
| 1. Las partes del cuerpo. | 1. Descripción física y de carácter. | 1. Individuo, dimensión física 2. Individuo, dimensión perceptiva y anímica |
| | 2. Datos personales. | 3. 1. Datos personales 3. 2. Documentación |
| | 3. Acciones y actividades habituales. | 1. 3. Acciones y posiciones que se realizan con el cuerpo. 1. 4. Ciclo de la vida y reproducción. 18. Actividades artísticas. |

| | | |
|---|---|---|
| | 4. Relaciones personales (familia y amigos). | 4. Relaciones personales: relaciones familiares, relaciones sociales, etc. |
| 5. Comidas y bebidas. 6. Objetos colocados sobre la mesa para la comida. | 5. Alimentación (comida, bebida, cocina y restaurante). | 5. Alimentación: dieta y nutrición, bebida, etc. |
| 8. La escuela (muebles y materiales). | 6. Educación (universidad, clase, estudios y exámenes). | 6. Educación: centros e instituciones educativas, profesorado y alumnado, etc. |
| 16. Profesiones y oficios. | 7. Trabajo y profesiones. | 7. Trabajo: profesiones y cargos, actividad laboral, etc. |
| 15. Juegos y distracciones. | 8. Ocio (deporte, espectáculos y aficiones). | 8. Ocio: tiempo libre y entretenimiento, deportes, etc. 18. Actividades artísticas. |
| 3. La casa (sin los muebles). 4. Los muebles de la casa. 7. La cocina (muebles y utensilios). 9. Iluminación, calefacción y ventilación. | 9. Vivienda (acciones, características y objetos). | 10. Vivienda: acciones relacionadas con la vivienda, características de la vivienda, etc. |
| 2. La ropa. | 10. Compras (dinero, colores, ropa y alimentación). | 12. Compras, tiendas y establecimientos: lugares, personas y actividades, pagos, etc. |
| 12. Medios de transporte. | 11. Vacaciones (viaje, alojamiento y transporte). | 14. Viajes, alojamiento y transporte: tipos de viajes, red de transportes, etc. |
| 10. La ciudad. | 12. Ciudad | 20. 3. 1. Ciudad |

| | | |
|---|--|---|
| 11. El campo. 13. Trabajos del campo y del jardín. | 13. Campo. | 20.3.2. Campo. |
| 14. Los animales. | 14. Naturaleza (clima, animales y plantas). | 20. Geografía y naturaleza: universo y espacio, fauna, etc. 2.3. Sensaciones y percepciones físicas |
| | 15. Medios de comunicación (internet, prensa, radio, televisión y teléfono). | 9. Información y medios de comunicación: correspondencia escrita, teléfono, etc. 16.4. Informática y nuevas tecnologías. |

Vemos que cuatro de los 15 temas propuestos coinciden en los tres casos, si bien puede variar su denominación. Estos cuatro temas son, en la denominación que nosotros utilizamos, *7. Trabajo y profesiones*, *8. Ocio (deporte, espectáculo y aficiones)*, *12. Ciudad* y *13. Campo*.

Vemos también que hay varios centros de interés del Proyecto Panhispánico que se han agrupado en una denominación más general para así hacerlo coincidir con el PCIC. Tal es el caso de *3. La casa (sin los muebles)*, *4. Los muebles de la casa*, *7. La cocina (muebles y utensilios)* y *9. Iluminación, calefacción y ventilación*. Estas tres áreas temáticas se agrupan en *9. Vivienda (acciones, características y objetos)*. Lo mismo sucede con *11. El campo* y *13. Trabajos del campo y del jardín*, que pasan a llamarse *13. Campo*. Si bien coincidimos con algunos autores que opinan que *13. Trabajos del campo y del jardín* o *9. Iluminación, calefacción y ventilación* apenas producirían resultado alguno, y menos aún en las primeras etapas de aprendizaje

¹⁷ Para seguir con más detalle los centros elegidos por esta investigadora y los motivos que le llevó a ello, véase Sánchez-Saus (2009:145 y 146). Y para ver de un

de la lengua, consideramos útil dejar la opción de incluirlas en nuestro estudio bajo otra denominación.

En cuanto a *5. Comidas y bebidas* y *6. Objetos colocados sobre la mesa para la comida*, quedan agrupados en *5. Alimentación (comida, bebida, cocina y restaurante)*.

Paredes (2015: 11) señala en su monográfico sobre DL, que algunos de los centros de interés que ciertos autores proponen añadir a la lista de los tradicionales ya se recogen en el PCIC (Áyora Esteban, 2006, sobre los *defectos y cualidades físicas y morales*; o Benítez Pérez, 1992, sobre los *colores*; entre otros). Otra de las razones que nos llevan a adaptar los centros de interés en nuestra encuesta coincide con lo apuntado por Sánchez-Saus (2009)¹⁷, y es que existen temas como *Iluminación, calefacción y medios de airear un recinto* y *Trabajos del campo y el jardín* que poseen pocas unidades léxicas y no nos sirven para nuestro objetivo. Además, hay otros que resultan demasiado concretos, como *Partes de la casa (sin muebles)*, *Los muebles de la casa*, *Objetos colocados en la mesa para la comida* y *La escuela: muebles y materiales*.

Hemos de añadir, al mismo tiempo, que en el MCERL aparecen temas que no son recogidos en los centros de interés del PPDL y, sin embargo, son parte fundamental del repertorio léxico de cualquier alumno de ELE. Por ejemplo, no aparecerían cuestiones relativas a la dimensión perceptiva y anímica del individuo (simpático, inteligente, tener miedo, odiar, etc.). Tampoco tendrían cabida los *adjetivos* (claros, oscuros, azules, rubio, bajo, etc.) relativos a la descripción de un individuo, tema que autores como Gallego Gallego (2014), entre

modo más amplio los temas o centros de interés elegidos por otros investigadores, véase López (2009: 416).

otros, han considerado necesario incluir en sus investigaciones. De este modo, vemos que, tal y como recomiendan autores como Paredes (2014), en los centros de interés que hemos empleado, se incluyen ciertas categorías gramaticales (adjetivos y verbos) que tradicionalmente no tienen representatividad en los diccionarios léxicos.

También tenemos áreas temáticas como los datos personales (domicilio, sexo, edad, número de teléfono, etc.), las acciones y actividades habituales (levantarse, desayunar, ir a clase, etc.) o las relaciones personales (amigo, padre, madre, etc.), propias de las primeras etapas de aprendizaje de la lengua y que, sin embargo, no tienen representación en los centros de interés propuestos por el PPDL.

Es cierto que hay otros temas del PPDL como *12. Medios de transporte* que pudieran no verse tan profusamente representados en nuestro centro de interés *11. Vacaciones (viaje, alojamiento y transporte)*. Lo mismo sucedería con el tema del PPDL *8. La escuela (muebles y materiales)*, que tal vez en *6. Educación (universidad, clase, estudios y exámenes)* no quedaría tan claramente representado, pero consideramos que hay suficiente relación entre ellos como para que los centros del PPDL mencionados queden incluidos en los nuestros. También hay otros centros de interés que, si bien se denominan de modo distinto, vienen a evocar los mismos conceptos. Por ejemplo, el tradicional *15. Juegos y distracciones*, que en nuestro estudio pasa a ser *8. Ocio (deporte, espectáculos y aficiones)*.

Hemos de añadir que algunas nociones específicas del PCIC no tienen ninguna entrada en el nivel A1, así que se quedan fuera de esta clasificación. Nos referimos a los siguientes temas: *11. Servicios*, *13.*

Salud e higiene, *15. Economía e industria*, *16. Ciencia y tecnología*, *17. Gobierno, política y sociedad*, *19. Religión y filosofía*. De todas formas, de cara a la investigación con sujetos de niveles superiores, creemos que algunos de los términos que tienen cabida dentro de estas nociones específicas (científico y tecnología, para el nivel A2 en *16. Ciencia y tecnología*) podrían incluirse en algunos de los centros de interés que hemos creado para este test. Por ejemplo, *Servicios y Economía e industria* podrían incluirse en *7. Trabajo y profesiones*, *16. Ciencia y tecnología* (términos como científico, tecnología, ciencia, etc.) o *Economía e industria* podrían repartirse entre *7. Trabajo y profesiones* y *15. Medios de comunicación*.

No obstante, y a pesar de los motivos que hemos ofrecido, somos conscientes de que nuestra clasificación puede no ser tan exhaustiva, pero hemos considerado que emplear las 20 nociones específicas que aparecen en el PCIC afectaría a la concentración que requiere completar una encuesta de este tipo (se necesitarían 40 minutos) y, por lo tanto, perjudicaría en gran medida la obtención de las respuestas.

Al mismo tiempo, entendemos que la diferencia en la denominación y, a veces, en la naturaleza misma entre los centros de interés clásicos y los utilizados aquí hace que el estudio comparativo de nuestros resultados con los obtenidos en otros trabajos de DL más tradicionales resulte lo más adecuado en conjunción con nuestro objetivo final, que es llegar a saber si el alumno taiwanés está aprendiendo el léxico que una obra de referencia como el PCIC aconseja para cada uno de los seis niveles de ELE.

2.4. EDICIÓN DE LOS MATERIALES

Además de la selección de los centros de interés, otra de cuestiones que más preocupan y que resultan más complejas en las investigaciones sobre DL es la edición de los materiales, ya que resulta fundamental seguir unos criterios de edición comunes que permitan la comparación entre diferentes investigaciones. Así lo indicaba Samper Padilla (1998) en un artículo que resultó ser pionero en la unificación de estos criterios. De hecho, la mayoría de los investigadores de disponibilidad en el mundo hispánico siguen estas mismas sugerencias, si bien suelen detallarse en función de las necesidades específicas de cada investigación. Nosotros, por nuestra parte, también utilizaremos estos criterios generales, que consisten en:

1) ELIMINACIÓN DE TÉRMINOS REPETIDOS. En el caso de que aparezca un término repetido dos veces en el mismo centro de interés, se mantendrá el que aparezca en primer lugar para, así, respetar su posición.

2) CORRECCIÓN ORTOGRÁFICA. El grado de faltas cometidas es muy elevado. Parte de ello puede explicarse porque durante la encuesta animamos a los alumnos a escribir todas las ocurrencias que tuvieran, aunque no estuvieran seguros de cómo escribirlas. De todas formas, los errores que hemos encontrado no son propiamente ortográficos, sino de grafía, posiblemente debidos a la deficiencia en el aprendizaje o a la interferencia de otras lenguas, generalmente el inglés.

Para solucionar esta cuestión, hemos acudido a las recomendaciones de la Real Academia Española en la última edición de su Diccionario de la lengua española (DRAE). En el caso

de que apareciesen palabras de uso habitual, pero no incluido en el DRAE, utilizaríamos el Corpus de referencia del español actual (CREA).

3) UNIFICACIÓN ORTOGRÁFICA. En nuestro trabajo no han aparecido palabras que permitan una doble ortografía, como *video / vídeo*, *cacto / cactus*, *psicología / sicología*. De todas formas, conviene tener en cuenta este principio, pues podría aparecer en encuestas a sujetos con un nivel más alto de español.

4) UNIFICACIÓN DE VARIANTES MERAMENTE FLEXIVAS. Este criterio consiste en utilizar la forma que suele aparecer en los diccionarios, siempre y cuando no suponga un cambio en el significado léxico. Es decir, el masculino singular en el caso de sustantivos y adjetivos y el infinitivo para los verbos.

En el caso de aquellas palabras que admitan el singular y el plural (*pantalón - pantalones*), seguiremos el mismo criterio que Bartol Hernández (2004) y otros muchos autores para poder facilitar las comparaciones. Por lo tanto, mantendremos lo escrito por el sujeto cuando la forma aparece en plural (*pantalones*), pero escribiremos la unificación bajo la forma *pantalón(es)* si aparece tanto en plural como en singular. Y si nos encontramos con la forma en singular, mantendremos esta (*pantalón*).

En el caso de que aparezca la forma en plural cuando existe una acepción diferente (*padre - padres*, *rebaja - rebajas*, *deber - deberes*), mantendremos lo escrito. Por ejemplo, *padres*, *rebajas*, *deberes*, etc.

Si encontramos heterónimos como *padre - madre*, *rey - reina*, *hombre - mujer*, etc. se mantendrá el lexema que aparezca.

5) UNIFICACIÓN EN EL CASO DE DERIVACIÓN QUE NO SUPONGA ALTERACIÓN DEL SIGNIFICADO LÉXICO. Se mantendrán separados los diminutivos y aumentativos que, según el DRAE o el CREA así lo especifiquen. Por ejemplo: *cuchara - cucharilla, mesa - mesilla*, etc.

6) UNIFICACIÓN DE LOS ACORTAMIENTOS. Todos los casos de palabras que admitan las dos formas serán recogidos de la manera *bici(cleta), tele(visión), (auto)bús*, etc.

7) TRATAMIENTO DE LOS SINTAGMAS Y DE AQUELLAS ELIPSIS ORIGINADAS EN COMBINACIÓN LÉXICA. Cuando nos encontremos con unidades fraseológicas como *ir de copas, ir de excursión, empezar a jugar*, etc. editaremos la respuesta del siguiente modo: *ir, copa(s), excursión, empezar, jugar*. De todas formas, hemos decidido respetar aquellos sintagmas que poseen unicidad semántica y referencial y cuya frecuencia de uso es alta. Es decir, si se trata de locuciones, como, por ejemplo, *caer bien, ponerse morado, ama de casa, ni fu ni fa, a diario*, etc., mantendremos el original. Para decidir si un sintagma debe mantenerse o no, acudimos a los diccionarios de locuciones de Inmaculada Penadés (2002, 2005, 2006 y 2008) y, en caso de que siguiésemos con dudas al respecto, utilizaremos el DRAE o el CREA para decidir, dependiendo de los casos.

8) TRATAMIENTO DE LAS MARCAS COMERCIALES Y LOS NOMBRES DE ESTABLECIMIENTOS Y EMPRESAS. Al igual que muchas otras investigaciones, admitiremos todas las marcas que aparezcan en el cuestionario, y no únicamente las marcas que estén lexicalizadas, como es el caso de *coca-cola, clínex, pimpón*, etc.

Además de estos criterios, pensados aplicar en estudios sobre la

lengua materna, tenemos la suerte de contar con los principios que, dos años más tarde, Carcedo-González (2000) propuso a partir de la ampliación y adaptación de los criterios de Samper Padilla (1998) a las necesidades propias de las investigaciones de DL aplicadas al español como lengua extranjera. Por su parte, Samper Hernández (2002) revisó estos criterios de Carcedo González (2000) y dejó constancia de una base que ha servido para la mayoría de los trabajos realizados en este campo. En nuestra investigación, utilizaremos estas indicaciones del siguiente modo:

1) TRATAMIENTO DE LOS ERRORES. Todas las palabras que aparezcan mal escritas serán corregidas e incluidas de modo correcto en el listado de disponibilidad, como, por ejemplo: *delgata** - *delgada, sencerio** - *sincero*, etc.

2) TRATAMIENTO DE LOS EXTRANJERISMOS. Para el caso de las importaciones léxicas que sean admitidas por el DRAE (*voleibol, vodka*, etc.) y aquellas que están muy extendidas en español, serán mantenidas sin modificaciones en las respuestas que analicemos (*facebook, wifi*, etc.). Aquellas que claramente pertenezcan a una lengua que no sea el español, serán eliminadas del estudio. Por ejemplo: *actress**, *movie**, etc.

3) TRATAMIENTO DE LAS INTERFERENCIAS DE OTRAS LENGUAS. Todas las interferencias semánticas y estructurales que no se usan en español han sido eliminadas del listado. Sin embargo, aunque la palabra aparezca en otra lengua, si su escritura es muy similar al español, será tratada como un error gráfico y optaremos por su corrección. Por ejemplo, *professor** - *profesor, cathedral** - *catedral*, etc. Cuando este último grupo de palabras presente dudas sobre si se trata o no de otra lengua, adoptaremos una posición flexible y las corregiremos. Por ejemplo: *football** (*football*)

en inglés) - *fútbol, patato** (*potato* en inglés) - *patata*, etc.

4) PRESTADORES Y MODIFICADORES. Cuando las palabras aparecen acompañadas de un artículo, hemos optado por eliminar este. En cambio, si se trata de un adjetivo o un adverbio que sirven para modificar el significado del núcleo del sintagma, hemos decidido aplicar los mismos criterios que explicábamos en el *Tratamiento de los sintagmas y de aquellas elipsis originadas en combinación léxica*, manteniendo así aquellos que poseen unicidad semántica y referencial

5) AMPLITUD DE LAS RELACIONES ASOCIATIVAS. Para nuestra investigación, hemos tratado de mantener todas las palabras que el informante ha incluido en cada uno de los centros de interés, excepto si no tienen una relación explicable con el tema.

En general, los cambios que hemos aplicado en la edición de los materiales obedecen a un enfoque pedagógico, con lo que no consideramos necesario mantener las variables morfológicas, los extranjerismos, etc. que puedan aparecer en el test, tal y como señala Tomé Cornejo en su tesis doctoral (2015).

2.5. TRATAMIENTO INFORMÁTICO DE LOS DATOS

¹⁸ Respecto a las fórmulas matemáticas que se han venido utilizando a lo largo de los años, Callealta Barroso y Gallego Gallego (2016) llevan a cabo un examen minucioso, centrándose en las recientemente más utilizadas, la de López Chávez y Strassburger Frías (1987, 1991) y la de Ávila Muñoz y Sánchez Sáez (2010). En su trabajo, plantean también unas fórmulas alternativas que buscan corregir cuestiones como la influencia de "factores como el número de informantes o la longitud de la lista de palabras en el centro" (Callealta Barroso y Gallego Gallego,

Para la cuantificación y edición de los datos utilizamos *Dispogen II* (Eveverría *et al.*, 2005), un programa informático desarrollado en MatLab que nos da el índice de disponibilidad de las unidades léxicas gracias a la utilización de la fórmula de López Chávez y Strassburger (1987 y 1991):

$$D(P_j) = \sum_{i=1}^n e^{-2.3 \left(\frac{i-1}{n-1} \right)} \frac{f_{ji}}{I_1}$$

Donde:

$D(P_j)$ = disponibilidad de la palabra j

n = máxima posición alcanzada en el centro de interés de la encuesta.

i = número de la posición en cuestión.

j = índice de la palabra tratada.

e = número de Euler, o constante de Napier (2,718281828459045...).

f_{ji} = frecuencia absoluta de la palabra j en la posición i .

I_1 = número de informantes que participaron en la encuesta.

Aunque haya autores que prefieran utilizar otras fórmulas y otros programas informáticos más adecuados a sus necesidades¹⁸, *Dispogen* presenta algunas ventajas que hemos querido aprovechar en nuestro trabajo. Permite, por ejemplo, realizar cálculos por centro de interés. Además, nos ofrece un *Índice de disponibilidad léxica individual* (relativo a los individuos que participan en el estudio).

2016: 90). Nosotros hemos preferido utilizar la de López Chávez y Strassburger Frías (1987, 1991) por la simplicidad de su uso gracias al programa *Dispogen* y porque es, hasta ahora, la más comúnmente empleada, por lo que nos facilitará los estudios comparativos que llevemos a cabo en el futuro.

Y en cuanto a los programas informáticos que se emplean en DL, de entre los trabajos consultados, destacamos el de Gallego Gallego (2014), pues hace una descripción general de los que se fueron utilizando hasta 2014.

También nos ofrece datos sobre los *clústeres* o asociaciones de diferentes unidades léxicas y otros índices como el número de respuestas por sujeto, el total de palabras, el total de vocablos, el promedio de palabras o el *Índice de cohesión* de cada centro de interés.

Este programa permite utilizar hasta cinco variables que poder seleccionar para crear diferentes grupos y compararlos entre sí. En nuestro caso, hemos incluido: sexo (hombre / mujer), lengua materna (chino), estancia en España (nunca/de turismo), material utilizado (ELE actual, SM) y nivel lingüístico (A1). La definición de estas variables nos servirá para poder hacer comparaciones con otros grupos de estudiantes de, por ejemplo, otra lengua materna u otro nivel lingüístico en ELE.

3. ANÁLISIS DE LOS MATERIALES

3.1. ANÁLISIS CUANTITATIVO

Si bien es cierto que, dada la naturaleza de nuestra investigación, no vamos a obtener resultados comparativos con respecto a otros grupos de informantes, los datos que presentaremos nos darán una descripción detallada del léxico disponible de 52 estudiantes taiwaneses con un nivel lingüístico de A1.

En este apartado presentaremos los índices obtenidos en nuestra investigación tal y como han ido apareciendo en otros estudios similares sobre DL en el mundo hispano, si bien, a lo largo de los años tanto la fundamentación teórica como la metodológica han ido evolucionando gracias a la publicación de nuevas investigaciones, tal y como detalla Paredes García (2012) en su artículo.

Comenzaremos con la presentación de los resultados generales que muestran los datos recogidos y que atañen a las unidades léxicas totales y diferentes. Continuaremos con un análisis más detallado sobre la productividad léxica de los informantes y de la riqueza léxica encontrada y finalizaremos con el índice de cohesión y la densidad léxica de cada uno de los centros de interés.

3.1.1. Resultados generales

Antes de comenzar a ver el análisis pormenorizado de los datos, es importantísimo recordar el significado de los términos *unidades léxicas totales* y *unidades léxicas diferentes*. Las *unidades léxicas totales* (ULT), que comúnmente se conocen como *palabras*, son todas las unidades con significado único, simples o complejas, que aparecen recogidas en las respuestas de los informantes y que pueden estar repetidas o no, lo que en algunas investigaciones se ha querido llamar *palabras*. En cambio, las *unidades léxicas diferentes* (ULD), o lo que en otras investigaciones se denomina *vocablos*, son aquellas que aparecen recogidas una sola vez en los diccionarios de DL y que puede encontrarse en otros trabajos como *vocablos*. Una vez aclarados ambos conceptos, veamos en la siguiente tabla cuál es el número de unidades léxicas totales encontradas y cuál es su promedio por centro de interés y por informante.

Tabla 3. Número de unidades léxicas totales y promedio por centro de interés y por informante.

| Número de unidades léxicas totales | Promedio de unidades léxicas totales (ULT) por centro de interés | Promedio de unidades léxicas totales (ULT) por informante |
|------------------------------------|--|---|
| 7028 | 468,53 | 132,60 |

En total, recogimos 7028 ULT en las 52 encuestas obtenidas, 469,53 ULT por centro de interés y 132,60 ULT por informante. En cuanto al

número de unidades léxicas diferentes y su promedio por centro de interés y por informante, los resultados se registran en la siguiente tabla.

Tabla 4. Número de unidades léxicas diferentes y promedio por centro de interés y por informante.

| Número de unidades léxicas diferentes | Promedio de de unidades léxicas diferentes (ULD) por centro de interés | Promedio de de unidades léxicas diferentes (ULD) por informante |
|---------------------------------------|--|---|
| 2008 | 133,87 | 37,89 |

Como es lógico, el número de ULD es menor que el de ULT; obtuvimos 2008 ULD en total, un promedio de 133,87 ULD por centro de interés y un promedio de 37,89 ULD por informante.

Las cifras aparecidas en estas dos tablas, más propiamente los promedios por centro de interés, servirán para determinar en los subapartados que vienen a continuación qué centros de interés están por encima de estos promedios y cuáles por debajo. De este modo, averiguaremos el valor de la productividad y la riqueza léxicas.

3.1.2. Productividad léxica. Número de unidades léxicas totales y promedio

La edición de los datos nos ha proporcionado cifras relativas al número de *unidades léxicas totales* (ULT) en cada uno de los centros de interés y el promedio de ULT por informante. No olvidemos que las ULY son todas las unidades léxicas que han sido recogidas en las encuestas. Por ejemplo, si la unidad léxica *feo* aparece 28 veces en las 53 encuestas, se contará como 28. Esto nos dará una idea de cuál es la *productividad léxica* en cada centro, tal y como aparece en la siguiente tabla.

Tabla 5. Número de unidades léxicas totales y promedio por informante en cada centro de interés (productividad léxica).

| Centro de interés | Número de unidades léxicas totales (ULT) | Promedio de unidades léxicas totales (ULT) por informante |
|--------------------------------------|--|---|
| 1. Descripción física y de carácter | 429 | 8,25 |
| 2. Datos personales | 454 | 8,73 |
| 3. Acciones y actividades habituales | 635 | 12,21 |
| 4. Relaciones personales | 423 | 8,13 |
| 5. Alimentación | 621 | 11,94 |
| 6. Educación | 493 | 9,48 |
| 7. Trabajo y profesiones | 341 | 6,55 |
| 8. Ocio | 500 | 9,61 |
| 9. Vivienda | 571 | 10,98 |
| 10. Compras | 535 | 10,28 |
| 11. Vacaciones | 533 | 10,25 |
| 12. Ciudad | 479 | 9,21 |
| 13. Campo | 345 | 6,63 |
| 14. Naturaleza | 318 | 6,11 |
| 15. Medios de comunicación | 351 | 6,75 |

Los resultados parecen mostrar unas cifras similares tanto en el número de ULT en cada centro de interés, como en el promedio de ULT por informante. Para poder apreciar con más claridad cuáles son los centros de interés más y menos productivos, hemos de establecer un rango a cada uno en función del número de ULT que contienen. Además, si dividimos el número total de unidades léxicas por los 15 centros de interés, obtendremos que la media es de 468,53 ULT por cada centro de interés (véase la Tabla 3), por lo que podremos establecer qué centros están por encima y por debajo de esa media. Veamos esta ordenación en la siguiente tabla.

Tabla 6. Rango de cada centro de interés según las unidades léxicas totales (productividad léxica).

| Centro de interés | RANGO | Número de unidades léxicas totales (ULT) |
|---|-------|--|
| 3. Acciones y actividades habituales | 1 | 635 |
| 5. Alimentación (comida, bebida, cocina y restaurante) | 2 | 621 |
| 9. Vivienda (acciones, características y objetos) | 3 | 571 |
| 10. Compras (dinero, colores, ropa y alimentación) | 4 | 535 |
| 11. Vacaciones (viaje, alojamiento y transporte) | 5 | 533 |
| 8. Ocio (deporte, espectáculos y aficiones) | 6 | 500 |
| 6. Educación (universidad, clase, estudios y exámenes) | 7 | 493 |
| 12. Ciudad | 8 | 479 |
| 2. Datos personales | 9 | 454 |
| 1. Descripción física y de carácter | 10 | 429 |
| 4. Relaciones personales (familia y amigos) | 11 | 423 |
| 15. Medios de comunicación (internet, prensa, radio, televisión y teléfono) | 12 | 351 |
| 13. Campo | 13 | 345 |
| 7. Trabajo y profesiones | 14 | 341 |
| 14. Naturaleza (clima, animales y plantas) | 15 | 318 |

Según el rango que hemos representado en esta tabla, los centros más productivos son *3. Acciones y actividades habituales*, con 635 ULT y 12,21 ULT por informantes) *5. Alimentación* (621 ULT y 11,94 ULT por informante) y *9. Vivienda* (571 ULT y 10,98 ULT por informante). En cambio, los centros de interés que menos unidades

¹⁹ Recordemos que este grupo de estudiantes ha trabajado durante algo más de dos años con el libro de texto ELE Actual A1 y en el momento de recoger los datos se encontraban en la Unidad 8 del libro ELE Actual A2, de la editorial SM.

léxicas totales han recogido son *14. Naturaleza* (318 ULT y 6,11 ULT por informante), *7. Trabajo y profesiones* (341 ULT y 6,55 ULT por informante) y *13. Campo* (345 ULT y 6,63 ULT por informante). La mayor o menor productividad léxica está relacionada con la naturaleza misma de cada centro de interés y podría tener explicación en los materiales de trabajo de la clase. Sabemos que hay temas que apenas aparecen en el libro de texto¹⁹ (*14. Naturaleza* o *15. Medios de comunicación*) y esto podría explicar su bajo rango. Sin embargo, hay otros temas, como *1. Descripción física y de carácter* y *2. Datos personales*, que aparecen ampliamente explicados los primeros temas del libro de texto *ELE Actual A1*, con lo que hace más de un año y medio que los estudiaron. Esta "falta de frescura" con respecto a la temática podría ser el motivo de su bajo rango. Con ello, no queremos dar una explicación definitiva al respecto, sino manifestar una intuición que deberá ser probada en una investigación específica sobre la influencia de los materiales de estudio en los resultados de la productividad léxica.

Otro de los motivos de la diferencia en el rango de cada centro de interés podríamos encontrarla en la naturaleza misma de cada uno, y es que algunos de los temas son más cerrados que otros y, por lo tanto, las posibilidades de encontrar unidades léxicas asociadas pueden ser menor.

Una cuestión que destaca en la tabla es que tenemos 8 centros de interés por encima de la media y 7 por debajo, lo cual indica que hay un reparto equilibrado entre los 15 centros de interés, si bien *3. Acciones y actividades habituales* y *5. Alimentación* se produce una

diferencia significativa con respecto al resto de centros.

3. 1. 3. Riqueza léxica. Número de unidades léxicas diferentes

Si lo que queremos saber es cuáles son los centros con más y con menos *riqueza léxica*, es necesario presentar los resultados correspondientes al número de *unidades léxicas diferentes* (ULD) en cada centro de interés. Recordemos que las ULD, a diferencia de las ULT, se refieren, tal y como se recoge en la mayoría de los trabajos de DL, a aquellas que aparecen una sola vez como entrada en los diccionarios de DL. Si tomamos el ejemplo de la unidad léxica *feo*, que aparece 28 veces a lo largo de las 53 encuestas, esta se contaría como una sola ULD.

Veamos cuáles son las cifras arrojadas por el análisis de los datos.

Tabla 7. Número de unidades léxicas diferentes y promedio por informante y por centro de interés (riqueza léxica).

| Centro de interés | Número de unidades léxicas diferentes (ULD) |
|--|---|
| 1. Descripción física y de carácter | 118 |
| 2. Datos personales | 115 |
| 3. Acciones y actividades habituales | 147 |
| 4. Relaciones personales (familia y amigos) | 90 |
| 5. Alimentación (comida, bebida, cocina y restaurante) | 149 |
| 6. Educación (universidad, clase, estudios y exámenes) | 133 |
| 7. Trabajo y profesiones | 144 |
| 8. Ocio (deporte, espectáculos y aficiones) | 134 |
| 9. Vivienda (acciones, características y objetos) | 181 |
| 10. Compras (dinero, colores, ropa y alimentación) | 160 |
| 11. Vacaciones (viaje, alojamiento y transporte) | 158 |
| 12. Ciudad | 145 |
| 13. Campo | 125 |

| | |
|---|-----|
| 14. Naturaleza (clima, animales y plantas) | 99 |
| 15. Medios de comunicación (internet, prensa, radio, televisión y teléfono) | 110 |

En este caso, tampoco parecen existir grandes diferencias entre unos centros y otros. Sin embargo, si tenemos en cuenta que el promedio del número de ULD por centro es de 133,87, podemos establecer que centros de interés están por encima y por debajo de este promedio. Veámoslo en la siguiente tabla.

Tabla 8. Rango de cada centro de interés según las unidades léxicas diferentes (riqueza léxica).

| Centro de interés | RANGO | Número de unidades léxicas diferentes (ULD) |
|---|-------|---|
| 9. Vivienda (acciones, características y objetos) | 1 | 181 |
| 10. Compras (dinero, colores, ropa y alimentación) | 2 | 160 |
| 11. Vacaciones (viaje, alojamiento y transporte) | 3 | 158 |
| 5. Alimentación (comida, bebida, cocina y restaurante) | 4 | 149 |
| 3. Acciones y actividades habituales | 5 | 147 |
| 12. Ciudad | 6 | 145 |
| 7. Trabajo y profesiones | 7 | 144 |
| 8. Ocio (deporte, espectáculos y aficiones) | 8 | 134 |
| 6. Educación (universidad, clase, estudios y exámenes) | 9 | 133 |
| 13. Campo | 10 | 125 |
| 1. Descripción física y de carácter | 11 | 118 |
| 2. Datos personales | 12 | 115 |
| 15. Medios de comunicación (internet, prensa, radio, televisión y teléfono) | 13 | 110 |
| 14. Naturaleza (clima, animales y plantas) | 14 | 99 |
| 4. Relaciones personales (familia y amigos) | 15 | 90 |

En primer lugar, al igual que ocurre con la productividad léxica, podemos observar que hay 8 centros de interés por encima de la media y 7 por debajo, por lo que el reparto de ULD entre los 15 centros de interés es equilibrado.

Además, distinguimos que los centros de interés más ricos son *9. Vivienda*, *10. Compras* y *11. Vacaciones*, mientras que los menos ricos son *4. Relaciones personales*, *14. Naturaleza* y *15. Medios de comunicación*. Conocer el motivo de estos resultados resulta complejo. Podemos entender que estos son aquellos centros cuya temática no solo es cercana a los estudiantes, sino que presentan una variedad grande de unidades léxicas que los contienen. En cambio, *14. Naturaleza* se trata de un centro más alejado de nuestros alumnos y, por lo tanto, ocupa un rango muy bajo. En cuanto a *4. Relaciones personales* y *15. Medios de comunicación*, creemos que el principal motivo de esta baja riqueza léxica es más bien propia de estos centros de interés, ya que encontramos un número menor de ULD que pueden asociarse a ellos. De todas formas, tal y como decíamos en el apartado relacionado con la productividad léxica, las características del libro de texto que utilizan también puede ser un motivo fundamental que afecta a estas diferencias en el rango, con lo que dejamos anotamos este motivo para indagar en él en trabajos posteriores.

3.1.4. Riqueza y productividad léxica

Tras lo visto con respecto a la riqueza y a la productividad léxica, podríamos pensar que existe un paralelismo entre los centros de interés más productivos y los más ricos léxicamente. Veamos una comparativa entre ambos índices en la siguiente tabla.

Tabla 9. Rango de los centros de interés según su riqueza y su productividad léxica.

| Número de ULD (riqueza léxica) | Centro de interés | RANGO | Centro de interés | Número de ULT (productividad léxica) |
|--------------------------------|--------------------------------------|-------|--------------------------------------|--------------------------------------|
| 181 | 9. Vivienda | 1 | 3. Acciones y actividades habituales | 635 |
| 160 | 10. Compras | 2 | 5. Alimentación | 621 |
| 158 | 11. Vacaciones | 3 | 9. Vivienda | 571 |
| 149 | 5. Alimentación | 4 | 10. Compras | 535 |
| 147 | 3. Acciones y actividades habituales | 5 | 11. Vacaciones | 533 |
| 145 | 12. Ciudad | 6 | 8. Ocio | 500 |
| 144 | 7. Trabajo y profesiones | 7 | 6. Educación | 493 |
| 134 | 8. Ocio | 8 | 12. Ciudad | 479 |
| 133 | 6. Educación | 9 | 2. Datos personales | 454 |
| 125 | 13. Campo | 10 | 1. Descripción física y de carácter | 429 |
| 118 | 1. Descripción física y de carácter | 11 | 4. Relaciones personales | 423 |
| 115 | 2. Datos personales | 12 | 15. Medios de comunicación | 351 |
| 110 | 15. Medios de comunicación | 13 | 13. Campo | 345 |
| 99 | 14. Naturaleza | 14 | 7. Trabajo y profesiones | 341 |
| 90 | 4. Relaciones personales | 15 | 14. Naturaleza | 318 |

La comparativa, a simple vista, nos ofrece una perspectiva distinta a lo que podríamos esperar. Ninguno de los centros de interés coincide en rango con respecto a su riqueza y a su productividad léxica. De hecho, tal y como señala Gallego Gallego (2014), la mayoría de los estudios de disponibilidad léxica señalan esta situación, que tiene que ver con el carácter abierto o cerrado de cada centro de interés. Aun habiendo diferencias en el rango de la riqueza y de la productividad léxica, la mayoría de los centros de interés se establecen según estos dos índices por encima o por debajo de la media. De este modo, si bien *9. Vivienda* ocupa el primer lugar en riqueza léxica y el tercer lugar en cuanto a productividad léxica, ambos rangos se sitúan por encima de la media. Los únicos centros de interés que distan en cuanto al rango que ocupan en uno u otro índice son: *5. Alimentación*, que ocupa el primer lugar en productividad léxica y baja a un quinto puesto en riqueza léxica; y *7. Trabajo y profesiones*, que se sitúa en el último lugar de productividad, pero está en un séptimo rango de riqueza léxica.

Volviendo al carácter abierto o cerrado de cada centro de interés, al que aludíamos en el párrafo anterior, es obvio que aquellos centros que tengan menos posibilidades de asociación de unidades léxicas ocupen posiciones bajas en cuanto a riqueza léxica. En el caso de *7. Trabajo y profesiones*, resulta que, aunque presenta una riqueza léxica por encima de la media, tiene uno de los rangos más bajos en cuanto a productividad léxica. Esto podría ser debido a que se trata de un tema que en el nivel A1 suele trabajarse someramente y dentro de temas como la información personal o la familia.

También creemos que el centro de interés *14. Naturaleza* no se ha tratado en profundidad en un primer nivel de aprendizaje, a excepción del clima, que sí suele aparecer en los manuales ELE de nivel A1. Todo esto, tal y como venimos diciendo, deberá ser tratado

con rigor en un trabajo sobre los manuales empleados en la clase ELE.

En cuanto a *4. Relaciones personales*, es el centro de interés más pobre léxicamente y uno de los menos productivos. Creemos que estas posiciones tan bajas en el rango pueden ser debidas a que se trata de un tema que no admite una variedad léxica abundante.

3.1.5. Índice de cohesión y densidad léxica

Tanto el índice de cohesión como la densidad léxica nos permiten saber si las respuestas dadas por cada individuo son iguales a las del resto de individuos (centro *compacto*) o difieren mucho (centro *difuso*).

El *índice de cohesión* de cada centro de interés se obtiene al dividir la frecuencia media de respuestas por individuo en cada centro de interés por el número de unidades léxicas diferentes, lo que nos da un valor comprendido entre 0 y 1. Cuanto más cercano a 1 sea el valor del índice de cohesión, más compacto será el centro. La *densidad léxica*, por su parte, se calcula dividiendo el número de unidades léxicas totales en cada centro por el número de unidades léxicas diferentes, lo que nos da el promedio de repeticiones de las unidades léxicas.

Sin embargo, si bien el índice de cohesión y la densidad léxica miden el nivel de coincidencia en las respuestas de los informantes, no concuerdan necesariamente. Esto es debido a que, como ya señalaba Hernández Muñoz (2004: 57), la densidad léxica “carece de límites referenciales que faciliten la comparación de los datos”.

La repercusión de los valores conseguidos por el índice de cohesión en cada centro de interés será comentada con más detalle en el análisis cualitativo, con lo que podremos analizar cuáles son las

unidades léxicas que los informantes han escrito en el test.

Veamos antes una tabla comparativa entre el índice de cohesión y la densidad léxica.

Tabla 10. Rango en función del índice de cohesión y la densidad léxica en cada centro.

| Índice de cohesión | Centro de interés | RANGO | Centro de interés | Densidad léxica |
|--------------------|--------------------------------------|-------|--------------------------------------|-----------------|
| 0,090 | 11. Vacaciones | 1 | 4. Relaciones personales | 4,70 |
| 0,083 | 10. Compras | 2 | 3. Acciones y actividades habituales | 4,32 |
| 0,080 | 5. Alimentación | 3 | 5. Alimentación | 4,17 |
| 0,075 | 9. Vivienda | 4 | 2. Datos personales | 3,95 |
| 0,071 | 7. Trabajo y profesiones | 5 | 8. Ocio | 3,73 |
| 0,071 | 3. Acciones y actividades habituales | 6 | 6. Educación | 3,71 |
| 0,069 | 1. Descripción física y de carácter | 7 | 1. Descripción física y de carácter | 3,64 |
| 0,064 | 13. Campo. | 8 | 11. Vacaciones | 3,37 |
| 0,064 | 6. Educación | 9 | 10. Compras | 3,34 |
| 0,063 | 2. Datos personales. | 10 | 12. Ciudad | 3,30 |
| 0,061 | 14. Naturaleza | 11 | 14. Naturaleza | 3,21 |
| 0,061 | 4. Relaciones personales | 12 | 15. Medios de comunicación | 3,19 |
| 0,060 | 8. Ocio | 13 | 9. Vivienda | 3,15 |
| 0,053 | 15. Medios de comunicación | 14 | 13. Campo | 2,76 |

| | | | | |
|-------|------------|----|--------------------------|------|
| 0,045 | 12. Ciudad | 15 | 7. Trabajo y profesiones | 2,37 |
|-------|------------|----|--------------------------|------|

En la tabla anterior vemos cómo, según tomemos el índice de cohesión o la densidad léxica como referencia, la posición que ocupa cada uno de los centros de interés no es la misma. Esta diferencia significa que los centros de interés que hemos utilizado tienen un comportamiento interno diferente que será comentado con detalle en el análisis cualitativo. Aunque en ambos índices existe una relación directa entre este y el número de participantes (cuanto mayor es el número de participantes, mayor es la densidad léxica), desarrollaremos con más detalle el índice de cohesión, puesto que es el que más se utiliza en los estudios de disponibilidad léxica, lo cual nos permitirá establecer una comparación más adecuada con los resultados de otros trabajos. Así, los centros de interés que más destacan teniendo en cuenta su índice de cohesión son:

- Los centros de interés más compactos, los que tienen una asociación conceptual mayor, son *11. Vacaciones, 10. Compras y 5. Alimentación.*
- Los centros de interés más dispersos, los participantes ofrecieron respuestas diferentes unos de otros son *8. Ocio, 15. Medios de comunicación y 12. Ciudad.*

3.1.6. Resultados por variables

En la mayoría de los estudios suelen tenerse en cuenta algunas variables que también hemos incluido en la recogida de datos, tal y como explicamos en el apartado *5.1. La muestra.* De todas estas variables, solamente dos pueden ser objeto de comparación en nuestro estudio: el sexo (6 hombres y 46 mujeres) y la estancia en países de habla española (49 nunca han estado en España y 3 han

estado de turismo). El resto de variables se convierten en *variables control*, ya que todos los sujetos de la muestra comparten las mismas características en cuanto a manual de aprendizaje utilizado (*ELE Actual A1 y A2*, de SM), nivel de español (A1), edad (entre 21 y 22 años), lugar de nacimiento (Taiwán) y lengua materna (chino). En la recogida de datos también incluimos el conocimiento de otras lenguas y su nivel lingüístico, pero entendemos que nuestro estudio se complicaría y la atención se apartaría de nuestro propósito inicial.

Dado que la relación entre individuos que han estado de turismo en España (3) e individuos que no han estado nunca (49) resulta desproporcionada, no hemos querido incluirla en este trabajo.

Sin embargo, la variable sexo (hombre / mujer) sí será tratada en nuestra investigación puesto que, incluso cuando la relación entre hombres y mujeres es dispar (6 hombres y 46 mujeres) y consideramos que los resultados no nos servirán para extraer conclusiones definitivas, creemos que los datos que tenemos podrán servirnos para hacer comparaciones futuras con los resultados obtenidos en otras comunidades. Así, teniendo en cuenta que hemos recogido las respuestas de 6 hombres y 46 mujeres, los índices generales obtenidos en función del sexo aparecen en la siguiente tabla.

Tabla 11. Índices generales en función del sexo.

| Índices | Hombres (6) | Mujeres (46) |
|--|-------------|--------------|
| Número unidades léxicas totales (ULT) | 970 | 60258 |
| Promedio de ULT por centro de interés | 64,66 | 403,86 |
| Promedio de ULT por informante | 161,66 | 131,70 |
| Número unidades léxicas diferentes (ULD) | 678 | 1805 |
| Promedio de ULD por centro de interés | 45,2 | 120,33 |
| Promedio de ULD por informante | 113 | 39,24 |

En la tabla, puede verse a simple vista que existe una diferencia enorme entre los resultados extraídos del grupo de hombres y el grupo de mujeres. Si bien, resulta lógico que tanto el número de ULT y el de ULD sea mayor en el caso de las mujeres, cuando atendemos al promedio de ambos índices por informante, observamos que el valor del número de ULT en el grupo de los hombres es mucho mayor: 161,66 ULT frente a 131,70 ULT en el grupo de las mujeres. Esta diferencia se acentúa en la comparación del promedio de ULD por informante: 113 ULD para el grupo de hombres y 39,24 ULD para el de mujeres.

En principio, no alcanzamos a justificar los motivos de estas diferencias, si no es la mera casualidad de que los hombres de nuestro grupo de estudio sean más aplicados en el estudio. Entendemos, por lo tanto, que será necesario ampliar la muestra poblacional o, incluso, comparar estos resultados con los obtenidos en otros trabajos.

3.2. ANÁLISIS CUALITATIVO

En este apartado, analizaremos cada uno de los centros de interés de la encuesta y describiremos el comportamiento de las diez unidades léxicas con mayor índice de disponibilidad, así como de los individuos que completaron la encuesta. Para ello, acudiremos a los siguientes valores: el ÍNDICE DE FRECUENCIA, el PORCENTAJE DE APARICIÓN y la FRECUENCIA ACUMULADA. La operación con estos valores nos dará como resultado el ÍNDICE DE DISPONIBILIDAD LÉXICA, tanto de las unidades léxicas, como de los individuos que participan.

Además, teniendo en cuenta que los centros de interés se forman a partir de la asociación de unas palabras con otras, lo que Coseriu (1987) denominó *configuraciones asociativas*²⁰, extraeremos los pares léxicos más repetidos en cada centro de interés para poder describir las relaciones léxicas y significativas de la encuesta.

Veamos, en primer lugar, cuáles son los resultados de la encuesta por centros de interés.

3.2.1. Índice de disponibilidad léxica

El ÍNDICE DE DISPONIBILIDAD LÉXICA DE UNA UNIDAD LÉXICA se refiere a un valor numérico comprendido entre 0 (disponibilidad mínima) y 1 (disponibilidad máxima) que señala "el grado de accesibilidad de un vocablo en el lexicón cuando el tema de la conversación lo requiere [y que] resulta de la combinación entre el número de veces que aparece un término y la posición que ocupa en el proceso de actualización" (Paredes García, 2012). El programa informático que hemos utilizado

para calcularlo, *Dispogen II*, emplea la fórmula establecida por López Chávez y Strassburger (1987, 1991). Este índice suele emplearse para determinar qué unidades léxicas pasarían a formar parte de lo que entenderíamos por léxico relevante. En nuestra investigación, este índice nos servirá para determinar cuáles son las unidades léxicas más recurrentes en el lexicón de nuestros alumnos de nivel A1 y, de esta manera, determinar si, una vez contrastados los resultados con el PCIC, es necesario hacer más énfasis en aquellas unidades léxicas con un bajo índice de disponibilidad y que, sin embargo, aparecen listadas en el PCIC. Además, el programa *Dispogen II* nos ofrece información relativa a los siguientes valores:

- 1) El ÍNDICE DE DISPONIBILIDAD LÉXICA INDIVIDUAL trata de calcular el grado de aportación de un informante en el listado general de disponibilidad.
- 2) El ÍNDICE DE FRECUENCIA señala el número de veces que se actualiza cada unidad léxica sobre el número total de unidades léxicas disponibles.
- 3) Por PORCENTAJE DE APARICIÓN entendemos el porcentaje de informantes que han mencionado la unidad léxica en cuestión.
- 4) La FRECUENCIA ACUMULADA es la suma de las diferentes frecuencias relativas desde la unidad léxica en cuestión hasta la que ocupa el primer lugar en cuanto a frecuencia absoluta. También se entiende como la contribución cuantitativa de una unidad léxica al conjunto del centro de interés correspondiente.

²⁰ En su tesis doctoral, Sánchez-Saus Laserna (2011: 323-436) presenta un apartado extenso y profundamente desarrollado sobre las "relaciones léxicas y relaciones *significativas* en los centros de interés".

Veamos a continuación cuál es el comportamiento de las unidades léxicas con un índice mayor en cada uno de los centros de interés aparecidos en el estudio.

3.2.1.1. DESCRIPCIÓN FÍSICA Y DE CARÁCTER

En este primer centro de interés, las cinco unidades léxicas con un mayor IDL son: ojo (0,5), guapo (0,49), feo (0,41), alto (0,37) y bonito (0,29).

Como vemos, cuatro de las cinco léxías son adjetivos que se emplean como descriptores generales y, además, del físico. Aparecen en más de la mitad de las encuestas, y las dos primeras (ojo y guapo) son mencionadas por un 65% de los entrevistados.

Todas estas unidades léxicas, tal y como esperábamos en este nivel, son simples (están formadas por una sola palabra) y no nos encontramos con fórmulas fijas, locuciones, modismos u otras expresiones hechas.

3.2.1.2. DATOS PERSONALES

Dentro de Datos personales, encontramos que las unidades léxicas con mayor accesibilidad por parte de los sujetos del estudio son: nombre (0,81), año (0,32), cumpleaños (0,21), teléfono (0,20) y número (0,17). De estas cinco palabras, todas ellas sustantivos, destaca *nombre*, que, además, aparece en un 92% de las respuestas.

Este tema no tiene representación en los tradicionales centros de interés, tal y como veíamos en el apartado 5.3. *Centros de interés*. Sin embargo, el PCIC, contempla dos nociones específicas referidas a los datos personales. Si mantenemos la numeración original del PCIC, se trata de 3.1. *Datos personales* y 3.2. *Documentación*.

La relación entre estas cinco palabras se limita a su pertenencia al mismo tema. De todas formas, hay una obvia relación sintáctica entre el sustantivo *número* y el complemento *de teléfono* que lo suele acompañar para formar *número de teléfono*. De este tipo de relaciones nos ocuparemos en el siguiente apartado.

3.2.1.3. ACCIONES Y ACTIVIDADES HABITUALES

El IDL más alto corresponde a las siguientes unidades léxicas: ir (0,43), jugar (0,33), cine (0,29), comer (0,28) y leer (0,27).

Se trata, tal y como el nombre de este centro de interés adelanta, de cuatro verbos que designan acciones y actividades habituales. Solamente encontramos un sustantivo, *cine*, propio de este tema.

Resulta lógico que *ir*, con un 61,5% de apariciones, sea la palabra que tiene un mayor IDL en este centro ya que los informantes la emplearon junto a otras como *cine*, *paseo*, *camping*, etc. para formar construcciones idiomáticas como *ir al cine*, *ir de paseo* e *ir de camping*. Ninguna de ellas es una locución verbal, como sí ocurre con *ir al grano* o *ir tirando*, que tienen un significado idiomático y, además, son inseparables y mantienen un significado único²¹. Por

²¹ Los trabajos de Penadés Martínez (2002, 2005, 2006, 2008 y 2012) son muy esclarecedores al respecto del estudio en profundidad de las locuciones en nuestro idioma y de su introducción en la clase ELE.

esta razón, cuando encontrábamos respuestas como *ir al cine*, las anotamos como dos unidades léxicas, *iry cine*.

En cuanto a su equivalente con los centros de interés tradicionales del PPDL, a efectos de futuras comparaciones, diremos que el tema *Acciones y actividades habituales* no aparece representado en estos.

3.2.1.4. RELACIONES PERSONALES (FAMILIA Y AMIGOS)

Las unidades léxicas con un mayor IDL son: padre (0,58), madre (0,55), hermano (0,46), amigo (0,41) y abuelo (0,32).

Las palabras que aparecen en primer lugar en las respuestas tienen que ver con la realidad inmediata de los sujetos en este caso concreto. Como era de esperar en una sociedad como la china, donde la familia es un pilar fundamental de la vida, los términos que designan relaciones de parentesco más cercanas aparecen en primer lugar. De hecho, *padre*, *madre* y *hermano* se registran en 40 de las 52 encuestas.

Destaquemos también que la palabra *amigo* aparece en cuarto lugar, debido tal vez a que los amigos son personajes clave en la vida de los jóvenes o, quizás, a que la palabra *amigo* forma parte del título del centro.

Tampoco el tema *Relaciones personales (familia y amigos)* tiene su centro de interés parejo en el PPDL, aunque sí en el PCIC, concretamente en las nociones específicas de *4. Relaciones personales*.

3.2.1.5. ALIMENTACIÓN (COMIDA, BEBIDA, COCINA Y RESTAURANTE)

En el centro de interés *Alimentación*, las unidades léxicas con un IDL mayor son: vino (0,69), agua (0,59), leche (0,51), café (0,39) y chocolate (0,32).

Estas cinco lexías son sustantivos que designan alimentos y bebidas, pero no encontramos ningún objeto colocado sobre la mesa para la comida. Las palabras que aparecen con relativa frecuencia en las respuestas y no son alimentos o bebidas son *vaso*, que se encuentra en 6 encuestas, *mesa*, con 5 apariciones, *plato*, con 4, y *menú*, que se escribe en 3 encuestas.

Cabe destacar que de entre las palabras más disponibles no se encuentren alimentos básicos como el arroz, la carne o el pescado. En cambio, la palabra más disponible es *vino*, que alcanza un 92% de aparición, lo que señala la estrecha relación que los informantes han establecido entre la cultura española y el vino.

3.2.1.6. EDUCACIÓN (UNIVERSIDAD, CLASE, ESTUDIOS Y EXÁMENES)

Las unidades léxicas más accesibles, atendiendo a su IDL, son: profesor (0,62), universidad (0,37), estudiante (0,35), clase (0,34) y español (0,32).

Todas estas palabras pertenecen al entorno personal de los sujetos de estudio, ya que se trata de estudiantes de universidad, pero llama la atención que no exista un porcentaje de aparición tan alto como el que encontrábamos en *vino* (92%) o *nombre* (92%). Así, *profesor* tiene un 82% de aparición y *estudiante* un 50%.

3.2.1.7. TRABAJO Y PROFESIONES

Dentro de *Trabajo y profesiones*, encontramos que las cinco unidades léxicas con un IDL más alto son: *profesor* (0,81), *estudiante* (0,38), *médico* (0,14), *actor* (0,13) y *periódico* (0,11).

Se entiende que *profesor* aparezca en primer lugar ya que es un término perteneciente a su realidad inmediata. Además, se recoge en 47 de las 53 encuestas analizadas.

Resulta curioso cómo se ha incluido *estudiante* dentro de este centro, y así lo hemos querido mantener ya que los informantes están habituados a completar formularios en los que, al preguntarles por la profesión que ejercen, anotan su condición de estudiantes.

El resto de términos referidos a trabajo y profesiones aparece ocasionalmente en una u otra encuesta ya que, tal y como vemos, tres de las cinco palabras más disponibles son recogidas en un porcentaje inferior al 25%.

3.2.1.8. OCIO (DEPORTE, ESPECTÁCULOS Y AFICIONES)

Veamos a continuación cuáles son las unidades léxicas con un IDL más elevado: *bailar* (0,30), *jugar* (0,29), *fútbol* (0,28), *ir* (0,25) y *cine* (0,25).

En esta ocasión, nos encontramos con 3 verbos y 2 sustantivos que en las respuestas aparecieron juntos (*jugar al fútbol* e *ir al cine*). Cuando nos encontramos con casos de este tipo, para unificar criterios, optamos por dejar en la lista tanto el verbo como el sustantivo (*jugar, fútbol, ir, cine*). Los verbos *jugar* e *ir* sirven para designar otras actividades que han aparecido en las respuestas, como *jugar al tenis, ir de paseo*, etc., y que fueron anotadas como *jugar, tenis, ir, paseo*, motivo por el que se encuentran entre los más disponibles.

3.2.1.9. VIVIENDA (ACCIONES, CARACTERÍSTICAS Y OBJETOS)

Las unidades léxicas con mayor IDL son: *mesa* (0,37), *casa* (0,31), *tele(visión)* (0,27), *autobús* (0,24) y *baño* (0,23).

Tenemos cinco sustantivos, tres de los cuales se refieren a objetos de la vivienda y uno a un tipo de habitación. Sorprende que *autobús* sea una de las palabras más disponibles en este centro, ya que, aparentemente, no parece tener relación alguna con la vivienda. Sin embargo, aparece en 9 de las encuestas analizadas.

Tal vez debido a las variadas respuestas que admite este centro de interés, el índice de disponibilidad es más bajo, en general, que el de otros centros. Es decir, las respuestas de los alumnos no coinciden (ocupa el rango 13 en cuanto a densidad léxica se refiere) a pesar de ser el centro con el mayor número de unidades léxicas diferentes.

3.2.1.10. COMPRAS (DINERO, COLORES, ROPA Y ALIMENTACIÓN)

Las unidades léxicas pertenecientes al centro de interés Compras que mayor IDL poseen son: *azul* (0,38), *blanco* (0,37), *rojo* (0,34), *euro* (0,28) y *pantalón* (0,22).

Tal y como está presentado, este tema es, quizás, el más abierto de todos, ya que permite léximas pertenecientes a diversos campos como el textil, el alimentario o el de las transacciones comerciales. La decisión que nos llevó a esta designación fue su equivalente en el PCIC (*12. Compras, tiendas y establecimientos*). Somos conscientes de que tenemos una superposición de este tema con *5. Alimentos*, pero creímos que si denomináramos *Compras* a este centro de interés (fundamental en las clases de ELE), resultaría inevitable pensar en alimentos.

Las respuestas que han dado los informantes pueden resultarnos muy útiles ya que, a pesar del carácter abierto de este centro de interés, las palabras que más aparecen son colores, lo que indica cierta facilidad a la hora de evocar términos relacionados con los colores.

3.2.1.11. VACACIONES (VIAJE, ALOJAMIENTO Y TRANSPORTE)

Dentro del centro de interés *Vacaciones*, encontramos las siguientes unidades léxicas con un IDL más elevado: *tren* (0,58), *autobús* (0,55), *hotel* (0,52), *avión* (0,47) y *coche* (0,35).

Aunque la naturaleza de este tema admite subtemas muy diferentes, como los viajes, el alojamiento o los medios de transporte, las respuestas se han centrado en cuatro medios de transporte, destacando el tren, cuyo porcentaje de aparición es del 75%.

3.2.1.12. CIUDAD

Estas son las unidades léxicas con un IDL más alto dentro de *Ciudad*: *calle* (0,54), *autobús* (0,52), *tren* (0,27), *coche* (0,26) y *banco* (0,22).

Una vez más, al igual que sucedía con el centro *Vacaciones*, la mayoría de las palabras más disponibles son medios de transporte. Solamente *calle*, con un porcentaje de aparición del 60% y *banco*, con un 27% son sustantivos que denominan una zona de la ciudad y un establecimiento respectivamente.

3.2.1.13. CAMPO

En cuanto al centro *Campo*, las unidades léxicas con más disponibilidad por parte de los informantes son: *perro* (0,30), *bici(cleta)* (0,20), *tranquilo* (0,19), *casa* (0,17) y *animal* (0,16).

Sabemos que este tema cabe dentro de *Naturaleza*, como así ocurre en el PCIC, pero creímos adecuado marcar esa diferencia para poder

establecer comparaciones futuras con estudios en los que estuvieran presentes *El campo*, *Trabajos del campo y del jardín* y *Los animales*, que son centros de interés del PPDL.

Las cinco palabras más disponibles de este centro pertenecen a categorías diferentes: fauna (*perro* y *animal*), medio de transporte (*bici(cleta)*), sensaciones (*tranquilo*) y vivienda (*casa*). Además, excepto *perro*, ninguna supera un 30% de aparición, lo que convierte este centro en el penúltimo en cuanto a densidad léxica.

3.2.1.14. NATURALEZA (CLIMA, ANIMALES Y PLANTAS)

Encontramos, en este centro, que las unidades léxicas con un mayor IDL son: *perro* (0,47), *frío* (0,45), *calor* (0,27), *sol* (0,24) y *gato* (0,21).

Tres palabras (*frío*, *calor* y *sol*) pertenecen a la categoría *Clima*, apuntada como subtema dentro de *Naturaleza*, mientras que las otras dos pertenecen a la categoría *Animales*. Entendemos, además, que son unidades léxicas vinculadas a la realidad de los informantes.

3.2.1.15. MEDIOS DE COMUNICACIÓN (INTERNET, PRENSA, RADIO, TELEVISIÓN Y TELÉFONO)

El centro *Medios de comunicación* tiene las siguientes unidades léxicas con un IDL mayor que el resto: *teléfono* (0,38), *tele(visión)* (0,35), *periódico* (0,29), *internet* (0,22) y *radio* (0,22).

Las unidades léxicas más disponibles abarcan un amplio espectro de medios de comunicación y no se centran en uno solo, como cabría esperar si atendemos la realidad dominante de los sujetos de estudio. Y es que, aunque internet tiene el protagonismo en detrimento de otros medios como la televisión, la radio o el periódico, el sustantivo *internet* aparece solamente en un 30% de las respuestas por debajo

también de *radio*, que posee un IDL mayor, pero su porcentaje de aparición es del 40%.

3.2.2. Pares léxicos

De cara a la enseñanza del léxico en la clase ELE, puede resultar muy útil saber qué unidades léxicas comparten más rasgos entre sí debido a que una fuerte relación entre estas podría facilitar su aprendizaje, tal y como señala Baralo (2005 y 2007). El modo en que las unidades léxicas forman relaciones puede cuantificarse gracias a una lista de asociaciones léxicas que proporciona el programa *Dispogen*²². Estas asociaciones, también llamadas *clústeres* o *pares léxicos*, son agrupaciones de unidades léxicas que diferentes individuos escriben en sus respuestas. Por ejemplo, dentro del centro de interés *Descripción física y de carácter* de nuestro corpus, el par léxico *guapo-feo* es el más frecuente, ya que aparece entre las respuestas de 14 informantes.

Estas asociaciones dependerán de la idiosincrasia de los centros de interés, así como de su organización interna (del Barrio, 2016). Según lo apuntado por Sánchez-Saus (2011), la tipología de las asociaciones léxicas sigue siendo objeto de discusión, pero hemos querido acogernos al esquema adaptado por esta autora (2011: 341) a partir del de Casas (2011: 102). Sin necesidad de entrar en el desarrollo minucioso de esta tipología, las asociaciones léxicas pueden establecerse *semánticamente*, según el nivel lingüístico en el que se opere (de carácter morfológico, léxico, suboracional, oracional y

²² En este programa aparecen los clústeres que más informantes escriben en cada uno de los centros de interés y tenemos la opción de obtenerlos en el mismo orden que aparecen. Es decir, para el CI 4.- Relaciones personales, el clúster *padre-madre* aparece 37 veces. Si seleccionamos la opción de que aparezcan en orden, obtendremos 24 veces para el clúster *padre-madre* y 13 para *madre-padre*.

textual) o *significativamente*, según el/los componentes del signo (significante, significado y referencia) que adoptemos como punto de partida.

Tal y como señala Saus-Laserna (2011), las conexiones que se establecen entre unas unidades léxicas y otras se basan en relaciones de tipo *asociativo*, es decir, se circunscriben a la realidad cercana de los informantes. De todas formas, continúa la autora (2011: 342), esto no impide que coexistan otros vínculos que vayan más allá de la asociación.

Para este primer trabajo de aproximación, veamos, entonces, cuáles son los diez pares léxicos que han aparecido con más frecuencia en las respuestas de los estudiantes y tratemos de establecer una relación entre ellos.

Sin embargo, hemos de señalar que Sánchez-Saus (2011) utiliza DispoGrafo, un programa desarrollado por Max Echevarría (2008) en la Universidad de Concepción (Chile) y que tiene en cuenta, además, que las unidades léxicas que forman el clúster aparezcan una a continuación de la otra. Si bien, en este trabajo no deseamos profundizar en este aspecto psicolingüístico de la disponibilidad léxica, encontramos de sumo interés su utilización.

Tabla 12. Pares léxicos formados por dos unidades léxicas según el número de apariciones.

| Rango | Par léxico (2 unidades léxicas) | Número de apariciones | Centro léxico al que pertenece |
|-------|---------------------------------|-----------------------|--------------------------------|
| 1 | madre-padre | 37 | 4 |
| 2 | café-leche | 21 | 5 |
| 3 | agua-vino | 16 | 5 |
| | ir-cine | 16 | 3 |
| 4 | frío-calor | 15 | 14 |
| 5 | perro-gato | 14 | 14 |
| | estudiante-profesor | 14 | 7 |
| | guapo-feo | 14 | 1 |
| 6 | autobús-tren | 13 | 11 |

En la tabla que hemos elaborado, destacan dos centros de interés: 5. *Alimentación*, con dos pares léxicos situados en una segunda (*café-leche*) y una tercera posición (*agua-vino*); y 14. *Naturaleza*, con el par léxico *frío-calor* en una cuarta posición y el par *perro-gato* en la quinta.

Si bien, podríamos creer que la relación *café-leche* es debida a la colocación *café con leche*, hemos de decir que tanto *café* como *leche* han aparecido de forma separada. No obstante, *café con leche* queda registrada en nuestra lista (aparece en 9 cuestionarios) ya que se trata de una unidad compleja que, si bien no es un compuesto sintagmático, se trata del nombre de una bebida que suele aparecer en la clase de ELE.

En cuanto al par *frío-calor*, se trata claramente de una relación léxica de antonimia, la misma que se da con el par léxico *guapo-feo*, en quinta posición. Sin embargo, los pares *agua-vino* y *perro-gato* no tendrían por qué darse como antónimos, pero existe una relación de prototipicidad en la que la antonimia viene por su relación

significativa designativa, siguiendo la terminología propuesta por Saus-Laserna (2011).

El par léxico que más aparece es el de *padre-madre*, dentro del centro de interés 4. *Relaciones personales (familia y amigos)*. Ambas lexías son cohipónimos, ya que pertenecen a un mismo hiperónimo, que es *familia*. Y también son cohipónimos *autobús-tren*, que se encuentran bajo el hiperónimo *medios de transporte*.

El par *ir-cine* obedece a una relación habitual dentro del tema 3. Acciones y actividades habituales, aunque también tendría sentido encontrarlo en 8. Ocio. En este caso, cuando los informantes anotaban *ir al cine* como una acción de ocio, optamos por mantener ambas palabras, pero de forma separada (*ir, cine*). Lo mismo hicimos con otras acciones como *ir de camping (ir, camping)*, *ir de compras (ir, compras)*, etc., de ahí, quizás, la razón de que *ir* aparezca junto a *cine* y formen un par léxico frecuente.

Estudiante-profesor se trata de una relación de tipo asociativo, ya que son los propios sujetos de la investigación los que, debido a su experiencia personal, relacionan ambos conceptos.

4. CONCLUSIONES

Este trabajo se ha orientado hacia el estudio del nivel de producción léxica de un grupo de alumnos taiwaneses, concretamente, en una etapa del aprendizaje que servirá de partida para estudios posteriores sobre la corrección del vocabulario empleado y sobre el proceso de adquisición a lo largo de las etapas de aprendizaje. En este primer acercamiento al estudio del léxico, tal y como anunciamos en la introducción, trabajamos con la metodología utilizada en el Proyecto

panhispánico de disponibilidad léxica, puesto que se ha convertido en un referente para los estudios de DL en nuestro idioma.

De acuerdo con los objetivos planteados, se han ido presentando los resultados y las reflexiones más importantes en el apartado dedicado al análisis de los materiales. Este proceso es el que nos ha llevado a desarrollar la respuesta a nuestra pregunta de investigación: *¿Cuáles el léxico disponible de 53 alumnos taiwaneses de ELE en un nivel A1 y qué características tiene?* Veamos a continuación cuáles son las conclusiones finales aportadas por el análisis de los datos.

Con respecto a la *productividad léxica*, si tenemos en cuenta cuáles son los centros de interés que menos unidades léxicas totales (ULT) presentan²³, podríamos pensar que los centros *14. Naturaleza*, *7. Trabajo y profesiones* y *13. Campo* tienen una temática que se aleja de la realidad de nuestros alumnos debido a que estos viven en un contexto urbano. Sin embargo, los campos *2. Datos personales*, *3. Descripción física y de carácter*, *4. Relaciones personales* y *15. Medios de comunicación*, que también están por debajo de la media, sí parecen ser más cercanos al contexto en el que viven. Creemos que la respuesta puede estar en el breve tratamiento que el manual utilizado en la clase da a estos temas, ya que los informantes apenas han sido expuestos a otro tipo de estímulos lingüísticos.

En términos de *riqueza léxica*, los centros de interés más ricos, es decir, los que presentan más unidades léxicas diferentes, son *9. Vivienda (acciones, características y objetos)*, *10. Compras (dinero, colores, ropa y alimentación)* y *11. Vacaciones (viaje, alojamiento y transporte)*. Estos tres centros de interés pertenecen a la realidad

²³ Véase la Tabla 6: Rango de cada centro de interés según las unidades léxicas totales (productividad léxica).

cercana de los estudiantes y, además, son abiertos, pues contienen varios subtemas que facilitan la evocación de diferentes unidades léxicas. Al contrario, los centros *4. Relaciones personales (familia y amigos)* y *15. Medios de comunicación (internet, prensa, radio, televisión y teléfono)* son de carácter cerrado, no admiten una variedad léxica tan alta como los anteriormente citados. En cuanto a *14. Naturaleza (clima, animales y plantas)*, a pesar de ser un tema muy rico léxicamente, se trata de una realidad no tan cercana a los estudiantes y, además, no se trabaja específicamente (a excepción del clima) en el manual que emplean los alumnos.

Cuando queremos saber si las respuestas dadas por cada individuo son iguales a las del resto de participantes en el estudio y así determinar que un centro es compacto o difuso, acudimos a dos índices: el índice de cohesión y la densidad léxica. A pesar de que ambos índices no coinciden necesariamente, hemos querido mantenerlos porque nos resultarán útiles en futuras comparaciones con otros trabajos que se refieran a uno u otro. En este primer caso, teniendo en cuenta que la mayoría de los estudios trabajan con el índice de cohesión, las conclusiones a las que llegamos son que los centros de interés más compactos son *10. Compras* y *5. Alimentación*, mientras que los más dispersos son *8. Ocio*, *15. Medios de comunicación* y *12. Ciudad*.

Cabe pensar que una mayor asociación conceptual (mayor índice de cohesión) es debida a que el centro en cuestión no admite una variedad léxica alta. Sin embargo, tras analizar una tabla comparativa con estos cuatro valores²⁴ (riqueza léxica, productividad léxica, índice de cohesión y densidad léxica), los resultados señalan que no hay

²⁴ Véase la Tabla 10: Rango en función del número de unidades léxicas diferentes y totales, del índice de cohesión y de la densidad léxica de cada centro.

relación entre ellos. Esto nos lleva a especular que el motivo se encuentra en el comportamiento de las respuestas de los informantes para cada uno de los centros de interés.

Analizamos también los resultados entre hombres y mujeres y no hemos extraído otra conclusión más que la mera casualidad de que la enorme diferencia entre las ULD presentadas por los hombres (161,66 ULD) y por las mujeres (131,70 ULD) sea debida a que los hombres de nuestro grupo de estudio sean más aplicados en las clases.

Y centrándonos en la accesibilidad o disponibilidad de las unidades léxicas dentro de cada centro de interés, podemos concluir que las doce unidades léxicas con un mayor índice de disponibilidad léxica son:

Tabla 13. Índice de disponibilidad y porcentaje de aparición de las unidades léxicas más disponibles.

| Rango | Unidad léxica | Índice de disponibilidad | Porcentaje de aparición | Centro de interés al que pertenece |
|-------|---------------|--------------------------|-------------------------|------------------------------------|
| 1 | nombre | 0,81 | 92,30% | 2. Datos personales |
| | profesor | 0,81 | 90,38% | 7. Trabajo y profesiones |
| 2 | vino | 0,69 | 92,30% | 5. Alimentación |
| 3 | profesor | 0,62 | 82,69% | 6. Educación |
| 4 | agua | 0,59 | 80,77% | 5. Alimentación |
| 5 | tren | 0,58 | 75% | 11. Vacaciones |
| | padre | 0,58 | 76,92% | 4. Relaciones personales |
| 6 | madre | 0,55 | 76,92% | 4. Relaciones personales |
| | autobús | 0,55 | 76,92% | 11. Vacaciones |

| | | | | |
|---|-------|------|--------|-------------------------------------|
| 7 | hotel | 0,52 | 69,23% | 11. Vacaciones |
| 8 | leche | 0,51 | 73,08% | 5. Alimentación |
| 9 | ojo | 0,50 | 65,38% | 1. Descripción física y de carácter |

Así, *nombre, profesor y vino* son las palabras más disponibles en el lexicón de los alumnos cuando se evocan en sus correspondientes centros de interés. Veamos con un ejemplo cómo leer cada una de las líneas de la tabla. Cuando a los informantes se les presenta el tema 2. *Datos personales*, lo primero que al 92,30% de ellos le viene a la cabeza es la palabra *nombre*. Estos vocablos, además de presentar un elevado porcentaje de aparición en las encuestas, suelen aparecer en las primeras posiciones de la lista en su centro de interés correspondiente, de ahí su elevado IDL.

Otro de los fenómenos que hemos podido observar en este análisis es que todas las unidades léxicas más disponibles para cada centro de interés son de tipo simple, están formadas por una sola palabra, lo que nos lleva a pensar que los alumnos taiwaneses tienen más facilidad para recordar unidades léxicas simples en esta primera etapa de aprendizaje de la lengua.

Finalizando ya con las conclusiones sobre los análisis de los resultados obtenidos en las encuestas tomadas a estos 53 alumnos taiwaneses de nivel A1, encontramos que los pares léxicos o clústeres más aparecidos son: *madre-padre, café-leche, agua-vino, ir-cine, frío-calor, perro-gato, estudiante-profesor, guapo-feo y autobús-tren*. En general, vemos que tanto los antónimos como las relaciones asociativas y los cohipónimos son los tipos de vínculos que facilitan la relación entre diferentes unidades léxicas y, por lo tanto, tal y como ya señalan las nuevas corrientes de investigación, el aprendizaje de las mismas se verá facilitado por las mismas. Esto convierte a los pares

léxicos en un ejemplo perfecto que, recordando lo dicho por (Ferreira, Roberto y Echevarría Max S., 2010), demuestra que “la activación de una determinada palabra activará otras palabras similares”, lo que nos da una idea perfecta para poner en práctica en el aula.

Para cerrar este apartado, nos gustaría señalar que este trabajo, tal y como se ha presentado, si bien nos lleva a unas conclusiones incompletas, no por ello dejan de ser ciertas y, además, nos permitirá comenzar futuras investigaciones a partir de unas justificaciones sólidamente construidas. Con ello, nos referimos a trabajos en los que se contraste nuestros resultados con los resultados de otros grupos de estudiantes, con lo aconsejado por el Instituto Cervantes en su Plan Curricular al respecto de las unidades léxicas que han de aprenderse para cada nivel, o con el léxico trabajado en los manuales ELE.

Por último, creemos ver en esta línea de investigación un rico filón que resultará de gran utilidad en la creación de materiales de español para alumnos taiwaneses, en la publicación de diccionarios específicamente dirigidos a sinohablantes o en el análisis de necesidades previo a la creación de los contenidos de la clase y de la evaluación del alumno.

BIBLIOGRAFÍA

董燕生 (2014). *Español Moderno 現代西班牙語學生用書 (1)*. 外語教學與研究出版社

Almela, R. *et al.* (2005). *Frecuencias del español. Diccionario y estudios léxicos y morfológicos*. Madrid: Ed. Universitas.

Alonso Ramos, M. (2012). “Explorando la frecuencia léxica para el Diccionario de colocaciones del español”. *Cum corde et in nova grammatica*. Santiago de Compostela: Universidade de Santiago de Compostela, 19-40 [consulta: 6 de febrero de 2018]. Disponible en la web:

<http://hdl.handle.net/10347/12183>

Alvar Ezquerro, M. (2004). “La frecuencia léxica y su utilidad en la enseñanza de español como lengua extranjera”, en M.A. Castillo Carballo *et al.* (eds.): *Las gramáticas y los diccionarios en la enseñanza de español como segunda lengua: deseo y realidad*, Actas del XV Congreso Internacional de ASELE. Sevilla: Universidad de Sevilla, 19-39.

Áyora Esteban, M.C. (2006). *Disponibilidad léxica en Ceuta. Aspectos sociolingüísticos*. Cádiz: Servicio de Publicaciones de la Universidad de Cádiz.

Baralo Ottonello, M. (dira.) (2012). *Revista Nebrija de Lingüística Aplicada a la Enseñanza de las Lenguas*, 11 [consulta: 5 de febrero de 2018]. Disponible en la web:

<http://www.nebrija.com/revista-linguistica/numero-11-2012>

Baralo Ottonello, M. (2005). “Aspectos de la adquisición del léxico y su aplicación en el aula”. *FIAPE. I Congreso internacional: El español, lengua del futuro*. marzo de 2005. Toledo, 1-32.

Baralo Ottonello, M. (2007). “Adquisición de palabras: redes semánticas y léxicas”. *Foro internacional: aprender y enseñar léxico*, 15 y 16 de junio de 2007. Múnich. [consulta 6 de mayo de 2018]. Disponible en la web:

http://www.cervantes-muenchen.de/es/05_lehrerfortb/Actas06-07/3MBaralo.pdf

Bartol Hernández, J.A. (2014). "Idoneidad de los centros de interés clásicos en los estudios de disponibilidad léxica aplicados al español como lengua extranjera: acerca del artículo de González Fernández". *Revista Nebrija de Lingüística Aplicada*, 8 (5) [consulta 2 de enero de 2018]. Disponible en la web:
<https://www.nebrija.com/revista-linguistica/idoneidad-de-los-centros-de-interes-clasicos-en-los-estudios-de-disponibilidad-lexica-aplicados-al-espa%C3%B1ol-como-lengua-extranjera-2.html>

Bartol Hernández, J.A. (2010). "Disponibilidad léxica y selección del vocabulario", en R. Castañer Martín y V. Lagüéns Gracia (eds.): *De moneda nunca usada. Estudios filológicos dedicados a José M. Enguita Utrilla*, 85-107.

Bartol Hernández, J.A. (2006). "La disponibilidad léxica". *Revista Española de Lingüística*, 36, 380-384.

Bartol Hernández, J.A. (2004). *Manual de usuario de dispolex.com*. Disponible mediante suscripción en la web:
<http://www.dispolex.com>

Benítez Pérez, P. (1992). "Disponibilidad léxica en la zona metropolitana de Madrid". *Boletín de la Academia Puertorriqueña de la Lengua Española*, 1 / 1, 71-102.

Cabrera Domecq, E. *et al.* (2015). "La disponibilidad léxica en estudiantes sinohablantes de ELE en la carrera de Medicina". *Comunicación Social: retos y perspectivas*. Centro de Lingüística Aplicada, Santiago de Cuba.

Callealta Barroso, F.J. & Gallego Gallego, D.J. (2016). "Medidas de disponibilidad léxica: comparabilidad y normalización". *Boletín de Filología*, Tomo LI, 1, 39-92.

Carcedo González, A. (1998). "Sobre las pruebas de disponibilidad léxica para estudiantes de español/ LE". *RILCE*, 14 (2), 205-224.

Carcedo González, A. (2000). *Disponibilidad léxica en español como lengua extranjera: el caso finlandés (estudio del nivel preuniversitario y cotejo con tres fases de adquisición)*. Humaniora, Ser. B, Turku, Turun Yliopisto.

Casas Gómez, M. (2011). "Problemas y criterios lingüísticos subyacentes a una tipología de relaciones en semántica". *Lorenzo Hervás*, 20, 63-108.

Chen, Z. & Lu, L. (2011). *Estudios de disponibilidad léxica en una muestra de estudiantes sinohablantes de español*. Tesis de Licenciatura en Lengua Española para No Hispanohablantes, Facultad de Español para No Hispanohablantes, Universidad de La Habana.

Chia, J. (2016). *Análisis comparativo de la disponibilidad léxica en español como lengua extranjera entre estudiantes chinos y sudafricanos*. Tesis en opción al título de Máster en Estudios Lingüísticos del Español como Lengua Extranjera, Universidad de la Habana.

Consejo de Europa (2002). *Marco común europeo de referencia para las lenguas: aprendizaje, enseñanza, evaluación*. Madrid: Secretaría General Técnica del MECD-Subdirección General de Información y Publicaciones y Grupo Anaya.

Coseriu, E. (1987). *Gramática, semántica, universales. Estudios de lingüística funcional*. Madrid: Gredos.

Del Barrio de la Rosa, F. (2016). "Algunas observaciones sobre la disponibilidad léxica en estudiantes itálofonos de español", en Eugenia Sainz González *et al.* (eds.): *Geométrica explosión. Estudios de lengua y literatura en homenaje a René Lenarduzzi*. Biblioteca di Rassegna iberistica, 1, 127-143.

Echeverría, M.S. *et al.* (2005). *Dispogen II. Programa computacional para el análisis de la disponibilidad léxica*. Concepción, Chile: Universidad de Concepción.

Echeverría, M.S. *et al.* (2008). "DispoGrafo: una nueva herramienta computacional para el análisis de relaciones semánticas en el léxico disponible", *Revista de Lingüística Teórica y Aplicada*, 46, 1, 81-91.

Ferreira, R. y Echeverría M.S. (2010). "Redes semánticas en el léxico disponible de inglés L1 e inglés LE". *Onomázein*, 21: 133-153.

Gallego Gallego, D.J. (2014a). "Léxico disponible de 82 inmigrantes estudiantes de español en la ciudad de Alcalá de Henares". *Lengua y migración*, 6 (2), 95-123 [consulta 14 de diciembre de 2017]. Disponible en la web: <http://hdl.handle.net/10017/22015>

Gallego Gallego, D.J. (2014b). *Léxico disponible de estudiantes de español como lengua extranjera en la Comunidad de Madrid*. Tesis Doctoral (Universidad de Alcalá, Facultad de Filosofía y Letras).

Gómez Molina, J.R. (2004). "La subcompetencia léxico-semántica", en Sánchez Lobato, J. y Santos Gargallo, I. (dir.): *Vademécum para la*

formación de profesores. Enseñar español como segunda lengua (L2) / lengua extranjera (LE). Madrid: SGEL, 491-510.

González Fernández, J. (2014). "Idoneidad de los centros de interés clásicos en los estudios de disponibilidad léxica aplicados al español como lengua extranjera". *Revista Nebrija de Lingüística Aplicada*, 8 (3) [consulta 14 de diciembre de 2017]. Disponible en la web: <https://doi.org/10.26378/rnlael816223>

Gougenheim, G.R. *et al.* (1956). *L'élaboration du Français fondamentale*. Paris: Didier.

Hatch, E. & Fahardy, H. (1982). *Research Design and Statistics for Applied Linguistics*. Rowley, Mass: Newbury House.

Herreros Marcilla, M. (2015a). "Los centros de carácter cultural en los estudios de disponibilidad léxica: análisis y nueva propuesta", en Gordejuela Senosiáin, A., Izquierdo Alegría, D., Jiménez Berrio, F., De Lucas Vicente, A. y Casado Velarde, M. (eds.): *Lenguas, lenguaje y lingüística. Contribuciones desde la Lingüística General*. Pamplona: Servicio de Publicaciones de la Universidad de Navarra, 279-290.

Herreros Marcilla, M. (2015b). "La disponibilidad léxica como propuesta de investigación docente. Análisis en alumnos inmigrantes no hispanoblatantes", en Yuko Morimoto, M. Victoria Pavón Lucero y Rocío Santamaría Martínez (eds.): *La enseñanza de ELE centrada en el alumno. Actas del XXV Congreso de ASELE*, 459-468.

Hidalgo Gallardo, M. (2017a). *La disponibilidad léxica como método de detección del vocabulario y de su selección en manuales: aplicación en una muestra de estudiantes sinohablantes de ELE*. Tesis

doctoral inédita defendida en el Departamento de Filología de la Universidad de Jaén.

Hidalgo Gallardo, M. (2017b). "Evolución diacrónica del léxico disponible de estudiantes sinohablantes de ELE". *SinoELE*, 16 [consulta 14 de mayo de 2018]. Disponible en la web: http://www.sinoele.org/images/Revista/16/Articulos/Hidalgo_SinoELE_16_2017_1-16.pdf

Huo, X. y Yun, S. (2014). *La disponibilidad léxica en estudiantes sinohablantes de ELE en la carrera de Medicina*. Tesis de Licenciatura en Lengua Española para No Hispanohablantes, Facultad de Español para No Hispanohablantes, Universidad de La Habana.

Instituto Cervantes (2006). *Plan Curricular del Instituto Cervantes. Niveles de referencia para el español*. 3 vols. Madrid: Instituto Cervantes, Biblioteca Nueva.

Jiménez Catalán, R.M. (ed.) (2014). *Lexical Availability in English and Spanish as a Second Language*. Springer: Netherlands.

Jiménez Catalán, R.M. (2017). "Estudios de disponibilidad léxica en español y en inglés: revisión de sus fundamentos empíricos y metodológicos". *Revista Nebrija de Lingüística Aplicada a la Enseñanza de las Lenguas*, 22 [consulta 13 de mayo de 2018]. Disponible en la web: <https://revistas.nebrija.com/revista-linguistica/article/view/2/4>

Lewis, M. (1993). *The Lexical Approach: The State of ELT and a Way Forward*. Hove: Language Teaching Publications.

Lin, J. (2006). *El estudio de disponibilidad léxica de los estudiantes chinos de español como lengua extranjera*, Memoria de Máster inédita, Universidad de Alcalá.

Lin, J. (2012). "El estudio de disponibilidad léxica de los estudiantes chinos de español como lengua extranjera". *marcoELE*, 14 [consulta 13 de febrero de 2018]. Disponible en la web: https://marcoele.com/descargas/14/lin-disponibilidad_lexica.pdf

López González, A.M. (2014). "Desarrollo de los estudios de disponibilidad léxica en Español Lengua Extranjera (ELE)". Actas del *XXIV Congreso Internacional de la ASELE La enseñanza del español como LE/L2 en el siglo XXI*: septiembre de 2013, Jaén, 397-408.

López Morales, H. (1973). *Disponibilidad léxica en escolares de San Juan*. Ms. inédito.

Martí Fernández, M. & Penadés Martínez, I. (coords.) (2015). "V Jornadas de Lengua y Comunicación. Léxico: Enseñanza e Investigación". *LinRed*, 13 [consulta 13 de diciembre de 2017]. Disponible en la web: http://www.linred.es/numero13_monografico.html

Medina Arejita, E.T. (2009). *Las nociones específicas del PCIC y la disponibilidad léxica como instrumento de selección del vocabulario. El caso de 43 estudiantes alemanes de español en Berlín*. Memoria de máster inédita. Universidad Internacional Menéndez Pelayo - Instituto Cervantes.

Michéa, R. (1953). "Mots fréquents et mots disponibles. Un aspect nouveau de la statistique du langage". *Les langues modernes*, 47, 338-344.

Moreno Fernández, F. (2012). "Disponibilidad léxica: cuestiones metodológicas: A propósito de disponibilidad léxica de los estudiantes hispanos de Redwood City, CA". *Revista Nebrija de Lingüística Aplicada*, 6 (2) [consulta 10 de diciembre de 2017]. Disponible en la web:
<https://revistas.nebrija.com/revista-linguistica/article/view/178>

Paredes García, F. (2001). "Disponibilidad léxica de los alumnos de Alcalá y su comarca: resultados generales", en Cruz, I., Santamaría, C. y Tejedor, C. (eds.): *La lingüística aplicada a finales del siglo XX. Ensayos y propuestas. Actas del XVII Congreso de AESLA*, 2001. Alcalá de Henares, vol. II, 721-728.

Paredes García, F. (2012). "Desarrollos teóricos y metodológicos recientes de los estudios de disponibilidad léxica". *Revista Nebrija de Lingüística Aplicada*, 11 (6) [consulta 6 de diciembre de 2017]. Disponible en la web:
https://www.nebrija.com/revista-linguistica/files/articulosPDF/articulo_5285dd9dbe4d3.pdf

Paredes García, F. (2014). "A vueltas con la selección de 'centros de interés' en los estudios de disponibilidad léxica: para una propuesta renovadora a propósito de la disponibilidad léxica en ELE". *Revista Nebrija de Lingüística Aplicada*, 8 (4) [consulta 9 de diciembre de 2017]. Disponible en la web:
<https://www.nebrija.com/revista-linguistica/a-vueltas-con-la-seleccion-de-centros-de-interes-en-los-estudios-de-disponibilidad-lexica-para-una-propuesta-renovadora-a-proposito-de-la-disponibilidad-lexica-en-ele.html>

Paredes García, F. (2015). "Disponibilidad Léxica y enseñanza de ELE: el léxico disponible como fuente curricular y como recurso en el aula". *LinRed, Monográfico: V Jornadas de Lengua y Comunicación "Léxico:*

Enseñanza e Investigación" [consulta 23 de diciembre de 2017]. Disponible en la web:
http://www.linred.es/numero13_monografico_Art2.html

Paredes García, F. (2017). "La comparabilidad de los trabajos de disponibilidad léxica". *Revista Nebrija de Lingüística Aplicada*, 22 [consulta 2 de diciembre de 2017]. Disponible en la web:
<https://revistas.nebrija.com/revista-linguistica/article/view/6>

Penadés Martínez, I. (2002). *Diccionario de locuciones verbales para la enseñanza del español*. Madrid: Arco Libros.

Penadés Martínez, I. (2006). *Ni da igual ni da lo mismo*. Madrid: Edinumen.

Penadés Martínez, I. (2008). *Diccionario de locuciones nominales, adjetivas y pronominales para la enseñanza del español*. Madrid: Arco Libros.

Penadés Martínez, I. (2005). *Diccionario de locuciones adverbiales para la enseñanza del español*. Madrid: Arco Libros.

Penadés Martínez, I. (2012). *Gramática y semántica de las locuciones*. Alcalá de Henares: Servicio de publicaciones de la Universidad de Alcalá.

Pérez Ruiz, J. (2011). "Análisis de errores y de la interacción oral como bases para una propuesta didáctica en la clase de conversación. Primeros resultados". *Revista Nebrija de Lingüística Aplicada*, 5 (2) [consulta 2 de diciembre de 2017]. Disponible en la web:
https://www.nebrija.com/revista-linguistica/files/articulosPDF/articulo_52947e01a919a.pdf

Rubio Lastra, M. (2012). *El mecanismo de las interrupciones en la conversación de alumnos taiwaneses de ELE*. Tesis doctoral inédita. Universidad Alcalá de Henares.

Rubio Sánchez, R. (2015a). "Estudio de disponibilidad léxica en aprendices italianos de español: análisis cuantitativo". *e-AESLA* (1) [consulta 2 de febrero de 2018]. Disponible en la web: <https://cvc.cervantes.es/lengua/eaesla/pdf/01/66.pdf>

Rubio Sánchez, R. (2015b). "Estudio de disponibilidad léxica en aprendices italianos de español: análisis cuantitativo", en Gordejuela Senosiáin, A., Izquierdo Alegría, D., Jiménez Berrio, F., De Lucas Vicente, A. y Casado Velarde, M. (eds.): *Lenguas, lenguaje y lingüística. Contribuciones desde la Lingüística General*. Pamplona: Servicio de Publicaciones de la Universidad de Navarra, 429-442.

Samper Hernández, M. (2002). "Disponibilidad léxica en alumnos de español como lengua extranjera". Málaga: ASELE. [consulta 2 de marzo de 2018]. Disponible en la web: <http://www.cervantesvirtual.com/nd/ark:/59851/bmc864g9>

Samper Hernández, M. & Jiménez Catalán, R.M. (2014). "Researching Lexical Availability in L2: Some Methodological Issues". En: Rosa María Jiménez Catalán (ed.). *Lexical Availability in English and Spanish as a Second Language*, Springer: Netherlands, 189-205.

Samper Padilla, J.A. & Samper Hernández, M. (2006). "Aportaciones recientes de los estudios de disponibilidad léxica". *Lynx: Panorámica de estudios lingüísticos*, 5, 5-95.

Sánchez, A. & Cantos, P. (1997). "Predictability of Word Forms (types) and Lemmas in Linguistic Corpora". *International Journal of Corpus Linguistics*, 2 (2), 251-272.

Sánchez Pérez, A. (2001). *Gran Diccionario de Uso del Español Actual*. Madrid: SGEL.

Sánchez-Saus Laserna, M. (2009). "La variable "nivel de español" en el léxico disponible de los estudiantes de español como lengua extranjera". *Pragmalingüística* (17), 140-153.

Sánchez-Saus Laserna, M. (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. Universidad de Cádiz [consulta 2 de marzo de 2018]. Disponible en la web: <http://hdl.handle.net/10498/15862>

Sans Baulenas, N. et al. (2016). *Bitácora. 1 Nueva edición. Libro del alumno*. Barcelona: Difusión.

Ten Have, Paul (1999). *Doing Conversation Analysis: A Practical Guide*. London: Sage Publications.

Tomé Cornejo, C. (2015). *Léxico disponible. Procesamiento y aplicación a la enseñanza de ELE*. Tesis doctoral (Universidad de Salamanca, Departamento de Lengua española). [consulta 14 de marzo de 2018]. Disponible en la web: https://gredos.usal.es/jspui/bitstream/10366/128287/1/DLE_Tom%C3%A9Cornej%C3%A9oC_L%C3%A9xicodisponible.pdf

Verdeses-Mirabal, R.T. (2012). "Disponibilidad léxica de los estudiantes hispanos de Redwood City, California". *Revista Nebrija de Lingüística Aplicada*, 11 [consulta 14 de marzo de 2018].

Disponible en la web:

<https://www.nebrija.com/revista-linguistica/disponibilidad-lexica-de-los-estudiantes-hispanos-de-redwood-city-california.html>

Zuccalá, D. (2014). "Disponibilidad léxica de ELE: el caso brasileño. Estudio piloto realizado con estudiantes de último curso de Traducción en la UNESP". *FIAPE. V Congreso internacional: ¿Qué español enseñar y cómo? Variedades del español y su enseñanza*: 25-28 de junio de 2014, Cuenca [consulta 7 de marzo de 2018].

Disponible en la web:

<https://www.mecd.gob.es/dam/jcr:fe2606fb-edd5-489f-8ac9-272b6bf079e8/34-disponibilidad-lexica-de-ele--zuccaladaniele-pdf.pdf>

FECHA DE ENVÍO: 17 DE OCTUBRE DE 2018

FECHA DE ACEPTACIÓN: 6 DE NOVIEMBRE DE 2018