

ESTUDIO EXPLORATORIO DEL LÉXICO DISPONIBLE DEL ALUMNADO VIETNAMITA DE ELE EN VIETNAM

PHUC TRAN*

Universidad de Ciencias Sociales y Humanidades (USSH)
Universidad Nacional de Vietnam - Ciudad de Ho Chi Minh (VNU-HCMC)

ANTONIO MANUEL ÁVILA MUÑOZ**

Universidad de Málaga

RESUMEN: La lengua española se encuentra en una fase inicial de expansión en Vietnam, impulsada por vínculos históricos con Cuba desde los años sesenta. A pesar de un interés creciente en el aprendizaje del español, en la actualidad este interés se circunscribe a sectores muy restringidos de la población y se ve obstaculizado por la falta de profesionales dedicados a la enseñanza de Español como Lengua Extranjera (ELE) y a la escasez de materiales didácticos. El presente estudio nos ha permitido realizar una primera exploración del léxico disponible del estudiantado vietnamita de ELE en Vietnam y analizar sus características. Hemos comprobado que la cultura vietnamita influye en la naturaleza de las palabras aportadas por el alumnado, así como el contacto que este mantienen con el español y con hispanohablantes fuera del aula. En el análisis exploratorio realizado no se han observado diferencias en función de la identidad de género de los sujetos que componen la muestra, aunque el estudio por grados de nivel de español nos ha ofrecido resultados más complejos que los que se esperaban. A la espera de pruebas estadísticas confirmatorias de nuestros resultados, consideramos que los hallazgos obtenidos ofrecen información valiosa para docentes de ELE, sugiriendo la necesidad de adaptar los materiales de enseñanza a las necesidades léxicas, culturales y contextuales del alumnado y fomentar un enfoque que se adapte a sus intereses reales.

PALABRAS CLAVES: léxico disponible, alumnos vietnamitas, nivel de español, sexo.

* Para correspondencia dirigirse a: Phuc Tran (hongphuc@hcmussh.edu.vn).

** Para correspondencia dirigirse a: Antonio Manuel Ávila Muñoz (amavila@uma.es).

EXPLORATORY STUDY OF THE AVAILABLE VOCABULARY OF VIETNAMESE SPANISH FOREIGN LANGUAGE STUDENTS IN VIETNAM

Abstract: The Spanish language is currently in an early stage of expansion in Vietnam, driven by historical ties with Cuba since the 1960s. Despite a growing interest in learning Spanish, this interest is currently confined to very limited sectors of the population and is hindered by a lack of professionals dedicated to teaching Spanish as a Foreign Language (ELE) and a scarcity of teaching materials. This study has allowed us to conduct an initial exploration of the available vocabulary of Vietnamese ELE students in Vietnam and analyze its characteristics. We have found that Vietnamese culture influences the nature of the words provided by the students, as well as the contact they maintain with Spanish and Spanish speakers outside the classroom. In the exploratory analysis conducted, no differences were observed based on the gender identity of the participants in the sample, although the study by Spanish proficiency levels yielded more complex results than expected. Pending confirmatory statistical tests, we believe that the findings offer valuable information for ELE teachers, suggesting the need to adapt teaching materials to the lexical, cultural, and contextual needs of the students and to promote an approach that aligns with their real interests.

Keywords: available lexicon, Vietnamese students, Spanish level, sex.

1. INTRODUCCIÓN

El español se encuentra en una etapa embrionaria de acceso y expansión en Vietnam como lengua extranjera. Históricamente, los primeros contactos que tuvieron los vietnamitas con el mundo hispánico se realizaron gracias a la relación político-histórica entre Vietnam y Cuba. Desde los años sesenta hasta la actualidad, se ha logrado dar continuidad a programas de movilidad entre estudiantes y especialistas de los dos países, la gran mayoría desde Vietnam hacia Cuba, lo que proporciona a este país asiático un número bastante considerable de personas que se interesan por el aprendizaje del español.

No obstante, esa incipiente comunidad hispanohablante no nativa se encuentra muy localizada en sectores económico-sociales específicos y se relaciona casi exclusivamente en el entorno de redes profesionales, por lo que el aprendizaje de la lengua española no está generalizado en la sociedad vietnamita. Además, es curiosa la falta de profesionales que se dedican a la enseñanza de Español como Lengua Extranjera (ELE) al volver de los mencionados programas internacionales, a pesar de la aparente motivación inicial. No debemos olvidar que Vietnam cuenta con un amplio conjunto de empresas de origen hispano, provenientes tanto de España como de Latinoamérica, lo que ofrece oportunidades laborales a la población vietnamita y, al mismo tiempo, debería ayudar a la expansión de la lengua española en este país, sea por motivo de trabajo o de estudio.

Una importante fuente de formación de especialistas hispanohablantes vietnamitas proviene de las dos universidades públicas más prestigiosas del país vinculadas con la enseñanza de lenguas extranjeras: la Universidad de Hanói (la antigua Universidad

de Estudios Extranjeros) y la Universidad de Ciencias Sociales y Humanidades de Ciudad Ho Chi Minh, donde se imparten grados oficiales de español desde 2002 y 2010, respectivamente. Anualmente se matriculan unos 60 estudiantes de nuevo ingreso y se registran unos 50 egresados en cada uno de los dos centros. La diferencia entre el número de matriculados y egresados se debe a varios factores: primero, a que algunos estudiantes deciden finalizar sus estudios en contextos de inmersión viajando a países hispanohablantes; segundo, porque ciertos alumnos abandonan sus estudios por cambios de universidad o de especialidad; tercero, porque otros solicitan bajas voluntarias y retrasan la fecha de graduación por problemas económicos, familiares o de salud, entre otros. No obstante, anualmente se registra un incremento gradual y estable en las solicitudes de ingreso, aunque los requisitos de admisión son cada vez más estrictos. A pesar del incremento gradual de solicitudes de matrícula, las plazas de admisión no suelen verse ampliadas debido a la falta de personal docente.

En este contexto, el objetivo principal de la presente investigación es observar, evaluar y analizar el nivel de aprendizaje de ELE por parte de alumnos vietnamitas. Para ello, llevaremos a cabo una aproximación exploratoria desde la perspectiva léxica que nos permitirá delimitar el enfoque, ya que abordaremos nuestro estudio usando la metodología propia de una corriente de investigación sólidamente consolidada en el contexto de estudios con aplicación práctica en la enseñanza de ELE: la disponibilidad léxica. Inspirados por los trabajos de Hidalgo Gallardo (2019), Mendoza Puertas (2018) o Rubio Lastra (2018), entre otros, adaptaremos la presente investigación a la comunidad estudiada tanto en el contenido como en la metodología empleada. Se pretende, por tanto, realizar un diagnóstico del nivel de aprendizaje léxico obtenido por estudiantes vietnamitas de ELE a partir de una aproximación exploratoria de naturaleza cualitativa del léxico disponible de estos estudiantes.

A continuación, se enumeran los objetivos secundarios relacionados con nuestra propuesta:

1. Conocer los índices generales y los centros de interés que presentan mayor y menor productividad, con el propósito de determinar sus causas.
2. Observar si existen variables sociales asociadas que condicionen la producción de léxico disponible.
3. Analizar el tipo de léxico obtenido en las listas analizadas de léxico disponible y realizar un estudio cualitativo de su naturaleza, así como de las características propias que influyen en la formación y selección de palabras por parte del alumnado vietnamita.

A partir de los objetivos expuestos hemos construido las siguientes hipótesis de trabajo:

Hipótesis sobre la productividad léxica por centros de interés: La cantidad de palabras disponibles en cada centro de interés estará relacionada con la familiaridad y el uso práctico de ese léxico en la vida diaria de los estudiantes. Por lo tanto, se espera que los centros de interés más cercanos a experiencias cotidianas del estudiantado proporcionen mayor producción léxica que otros menos familiares.

Hipótesis sobre la cohesión léxica: El índice de cohesión léxica en un centro de interés refleja, en última instancia, la homogeneidad en el aprendizaje dentro del campo semántico representado por ese centro de interés. Es probable que aquellos temas más estudiados y discutidos en clase presenten mayores índices de cohesión, mientras que los temas menos trabajados o más amplios semánticamente muestren menor cohesión léxica, lo que podría afectar al aprendizaje del vocabulario.

Hipótesis sobre las diferencias de género: A pesar de pequeñas diferencias observadas en la producción léxica entre hombres y mujeres en trabajos previos, creemos que la identidad de género no será un factor determinante en la cantidad y diversidad del léxico disponible en estudiantes vietnamitas de ELE, ya que la segregación en clase según este factor es inexistente en Vietnam.

Hipótesis sobre el nivel de español: Se espera que los estudiantes de niveles superiores de ELE manifiesten mayor productividad léxica. No obstante, es probable que otros factores complementarios, como la motivación o el interés personal en el aprendizaje, influyan de manera más determinante que el nivel de instrucción formal en la producción léxica.

Hipótesis sobre la influencia de otras lenguas: Los errores léxicos o las composiciones creativas encontradas en las respuestas estarán influidas por la lengua materna de los estudiantes (vietnamita) o por el inglés (lengua vehicular en Vietnam), lo que sugiere que el bilingüismo o el uso de una tercera lengua (inglés) podría ser un factor que afecte tanto a la adquisición y producción léxica en español como a la naturaleza del léxico que encontremos.

La justificación última del presente trabajo se encuentra en la escasez de estudios dedicados específicamente a estudiantes vietnamitas de ELE, un grupo discente con características particulares. Se trata de un colectivo reducido que, además, encuentra un significativo impacto sociocultural y lingüístico cuando se acerca al estudio de la cultura y la lengua española, entre otros motivos por los pocos contactos que los estudiantes tienen con el mundo hispánico. Sin embargo, en nuestra opinión, merece la pena avanzar de diferentes maneras en el conocimiento del estado actual de la situación de la enseñanza de ELE en este país asiático con el objetivo de atraer a más especialistas interesados a construir, en un futuro no demasiado lejano, una comunidad reconocida de ELE en Vietnam.

2. MARCO TEÓRICO

2.1. Estudios de disponibilidad léxica: origen y desarrollo

Los estudios sobre disponibilidad léxica surgieron a mediados del siglo pasado, cuando un grupo de investigadores liderado por G. Gougenheim publicó en 1956 *L'elaboration du Français élémentaire* en el marco de un proyecto financiado por el Ministerio de Educación Nacional de Francia. El objetivo de aquella obra pionera fue la recogida de palabras elementales que sirvieran para la enseñanza de francés a ciudadanos no francófonos, en especial a los procedentes de las colonias. En el desarrollo de dicho trabajo se evidenció que había una serie de palabras muy

conocidas por parte de los hablantes nativos que no estaban registradas en las listas de frecuencia existentes obtenidas a partir del criterio de uso, el único que se barajaba hasta entonces (Michéa, 1953). Esto puso en cuestión la eficacia de la ‘frecuencia de uso’ como argumento para la selección del vocabulario con fines didácticos, puesto que este criterio facilitaba la aparición de vocablos ‘atemáticos’, entendidos como palabras usuales en cualquier contexto comunicativo sin especificación de tema alguno, y no propiciaban la obtención de vocablos ‘temáticos’ muy conocidos por hablantes nativos, pero que solo se activan al tratar determinados temas que pueden llegar a ser muy frecuentes en la vida cotidiana como los medios de transportes, las partes del cuerpo o las partes de una casa.

En este contexto de revisión, surgió el concepto de ‘léxico disponible’, que se refiere a las palabras que suelen actualizarse en situaciones comunicativas frecuentes cuando se abordan determinados temas cotidianos, aunque las palabras en sí no tienen que ser necesariamente frecuentes (Michéa, 1953: 340; Samper Hernández, 2002: 8; Ávila Muñoz y Villena Ponsoda, 2010: 32). La diferencia entre los dos tipos de repertorios, el frecuente y el disponible, obligó a la aplicación de pruebas asociativas en la recogida de datos para la publicación de una versión revisada del *Français Fondamental* (Gougenheim *et al.*, 1964), basadas originalmente en dieciséis estímulos, llamados ‘centros de interés’, que sirvieron de modelo metodológico para muchos trabajos posteriores. Estos 16 centros de interés originales fueron: *El cuerpo humano*, *La ropa*, *Partes de la casa (sin los muebles)*, *Los muebles de la casa*, *Alimentos*, *Objetos colocados en la mesa para la comida*, *La cocina y sus utensilios*, *La escuela: muebles y materiales*, *Calefacción, iluminación y medios de airear un recinto*, *La ciudad*, *El campo*, *Medios de transporte*, *Trabajos del campo y del jardín*, *Los animales*, *Juegos y distracciones* y *Profesiones y oficios*.

2.2. La disponibilidad léxica en el mundo hispánico

Siguiendo los patrones propuestos por los investigadores pioneros franceses, en los años 70 se produjeron los primeros trabajos sobre léxico disponible en el mundo hispánico. Entre ellos, destacaron los producidos por Humberto López Morales (1973, 1978, 1999). Una de las principales aportaciones del profesor López Morales ha sido la puesta en marcha del *Proyecto Panhispánico de Disponibilidad Léxica* (PPHDL) que tiene como objetivo elaborar un diccionario del léxico disponible español capaz de recoger voces que sean reflejo de la diversidad geolectal del mundo hispánico. El proyecto se expandió rápidamente, desde Hispanoamérica (Puerto Rico, México, Chile, República Dominicana, Costa Rica, Uruguay, Colombia, Panamá) hasta la Península Ibérica (Madrid, Gran Canarias, País Vasco, Andalucía, Gibraltar, Ceuta, Castilla y León, Asturias, Valencia, Aragón, Castilla-La Mancha) (Martínez Olmo, 2015: 30-36).

Las investigaciones incorporadas al PPHDL comparten algunos aspectos metodológicos que facilitan la comparación de los resultados obtenidos y el desarrollo del macroproyecto: se trabaja con estudiantes preuniversitarios, las pruebas asociativas incluyen los 16 centros de interés originales, no hay límite en cuanto al número de palabras aportadas, pero sí en el tiempo de reacción de los informantes (dos minutos por estímulo temático), se establecen criterios de edición de datos comunes y se aplica

el mismo índice matemático para el cálculo estadístico de la disponibilidad léxica (Ávila Muñoz y Sánchez Sáez, 2011: 48-49).

2.3. Aplicaciones de la disponibilidad léxica

Los trabajos sobre el léxico disponible proporcionan datos beneficiosos a muchas ramas de la lingüística. Por un lado, permite analizar diferentes variedades diatópicas, lo que ha supuesto el desarrollo de investigaciones relacionadas con la dialectología. Desde el trabajo pionero e inspirador de Mackey (1971) que estudiaba la diferencia entre el léxico disponible de Francia, y el de Canadá, los avances en este campo han sido constantes. En el caso de la dialectología de la lengua española, López Chávez (1992) recogió el léxico disponible de Madrid, Gran Canaria, República Dominicana y Puerto Rico con el objetivo de comprobar la compatibilidad léxica entre las variedades del español empleadas en cada una de esas zonas. Años después, Samper Padilla (1999) recogió exclusivamente los vocablos que compartían Puerto Rico y Gran Canaria y encontró unos índices de compatibilidad muy altos entre el léxico disponible de las dos comunidades.

Por otro lado, los trabajos de disponibilidad léxica ayudan a conocer las diferencias del uso de vocabulario de la misma lengua, hablada en una misma comunidad, pero entre distintos grupos de hablantes cultural y socialmente estratificados, lo que concierne directamente a la sociolingüística. En este ámbito, López Morales (1973) trabajó con 63 informantes escolares de San Juan de Puerto Rico y logró comprobar que la adquisición léxica estaba condicionada por el nivel sociocultural, de modo que las personas de nivel medio disponían de una ventaja notable sobre las de nivel bajo y el obrero. En otros estudios, López Morales también atendió a la variable sociocultural, observando significativas implicaciones en la adquisición léxica (López Morales, 1979). Del mismo modo, Samper Padilla (2020) confirmó recientemente la influencia de los factores generacionales en la disponibilidad léxica.

Además de las aplicaciones mencionadas, los resultados de disponibilidad aportan información valiosa de la que se benefician estudios psicolingüísticos, puesto que ayudan a observar el tipo de relación que existe entre las palabras y la realidad asociada en la mente de los informantes (Ávila Muñoz y Sánchez Sáez, 2011).

Sin embargo, uno de los ámbitos donde más se han desarrollado los estudios de disponibilidad léxica es, sin duda, el de la enseñanza de lenguas, tanto maternas como extranjeras. Los estudios sobre el léxico disponible de una población proporcionan información muy relevante relacionada con los procesos de adquisición léxica y pueden ayudar a realizar una adecuada planificación curricular y a mejorar el diseño de manuales o materiales didácticos (Carcedo, 1998). En el ámbito de enseñanza de ELE, la investigación de Carcedo González (1999) puede considerarse pionera. Carcedo estudió los errores léxicos en el vocabulario disponible de estudiantes finlandeses de español y concluyó que la transferencia léxica de lengua(s) extranjera(s) —factor responsable de un 47% de los errores detectados— desempeñaba un papel importante en la adquisición del vocabulario y, por tanto, en la competencia lingüística de los alumnos de ELE.

No obstante, Samper Hernández (2002) detectó justo lo contrario en su trabajo llevado a cabo con estudiantes de ELE en Salamanca, al llegar a la conclusión de que no existe relación asociativa entre el factor ‘conocimiento de otras lenguas’ y la cantidad de vocablos aportados. Por su parte, Sánchez-Saus Laserna (2009) realizó encuestas a alumnos universitarios de ELE en Cádiz con el objetivo de examinar la influencia de la variable ‘nivel de español’ en las listas obtenidas. A través de su análisis, logró confirmar la hipótesis de que el nivel de español de los alumnos condiciona de manera directamente proporcional el número de palabras que actualizan, de manera que los estudiantes con nivel de español superior muestran más capacidad de producción léxica. Sin embargo, el grado de crecimiento no es homogéneo entre los tres grupos analizados por Sánchez-Saus, sino que se hace más evidente entre el nivel inicial y el medio que entre el medio y el superior.

En cuanto a estudiantes de ELE de origen asiático, García Naranjo (2019) trabajó con informantes universitarios japoneses y comprobó que estos presentaban características propias en la adquisición del léxico condicionadas por su contexto cultural y académico. Además, concluyó que las palabras recogidas antes del proceso de edición de los datos ofrecían información explorable y utilizable para fines didácticos.

Jindong Song (2018) realizó las pruebas asociativas a 15 estudiantes chinos de ELE en las que obtuvo resultados bastante parecidos a los encontrados en otros trabajos anteriores realizados con poblaciones de características similares. Del mismo modo, Lin Jing (2012) llevó a cabo un trabajo con 263 estudiantes chinos de grado de ELE en el que obtuvo resultados inesperados sobre la abundante cantidad de errores ortográficos, tanto en estudiantes de niveles avanzados como iniciales, lo cual originó en la autora reflexiones preocupantes como docente.

Otro estudio realizado con informantes chinos de ELE fue llevado a cabo por Hidalgo Gallardo (2019), con una muestra de 440 sujetos procedentes de seis universidades públicas de China. Por una parte, el autor obtuvo una importante cantidad de errores ortográficos en las listas analizadas que se detectaron incluso en los listados del alumnado que cursaba niveles avanzados. Además, las variables *Sexo* y *Lengua materna* no condicionaron los resultados obtenidos por Hidalgo Gallardo, mientras que el *Nivel de español* y la *Estancia en un país hispanohablante* se revelaron como determinantes en el desarrollo del léxico disponible de estos alumnos. Por otra parte, Hidalgo Gallardo también realizó un análisis contrastivo de dos manuales de ELE —uno de autor y editorial chinos (*Español Moderno 1*) y otro de procedencia española (*Bitácora 1*)—, que se ordena en tres partes: un análisis contrastivo de cada manual por separado con *una lista unificada* (creada mediante la agrupación de los vocablos más disponibles de nativos de las comunidades de Chile, Puerto Rico, Jaén y La Rioja), una comparación entre ambos manuales, y un cotejo entre la lista unificada, el libro *Español Moderno 1* y el léxico más disponible de los informantes. El trabajo y los resultados del estudio de Hidalgo Gallardo resultarán inspiradores para la elaboración de nuestra investigación.

3. METODOLOGÍA DE INVESTIGACIÓN

3.1. Muestra

La muestra usada en nuestro estudio exploratorio está compuesta por 45 estudiantes del grado oficial en Lingüística y Literatura Española de la Universidad de Ciencias Sociales y Humanidades de Ciudad Ho Chi Minh. Se trata de alumnado que cursa segundo, tercero y cuarto año de la carrera y, según los planes de estudio de la citada Universidad, deberían alcanzar los niveles de A1, B1 y B2 de español respectivamente, según el Marco Común Europeo de Referencia para las lenguas (MCER). La Tabla 1 presenta la distribución de la muestra según la identidad de género y el nivel estudiado por el alumnado.

| | Universidad Ciudad Ho Chi Minh | | | Total |
|--------------|--------------------------------|-----------|-----------|--------------|
| | Nivel A2 | Nivel B1 | Nivel B2 | |
| Hombres | 1 | 5 | 4 | 10 |
| Mujeres | 18 | 6 | 11 | 35 |
| Total | 19 | 11 | 15 | 45 |

Tabla 1. *Muestra*

3.2. Cuestionario sociológico

La recogida de datos se planificó a partir de dos cuestionarios, uno sociológico y otro léxico basado en pruebas de asociación de palabras a partir de un estímulo léxico inicial. Ambos cuestionarios se recogieron en una única sesión en presencia del investigador, procedimiento que supuso indudables ventajas metodológicas (Hernández Campoy y Almeida, 2005: 122; Vida Castro, Ávila Muñoz y Carriscondo Esquivel, 2016: 135-136).

A través del cuestionario sociológico se pretendió recoger datos personales (identidad de género, edad, domicilio, idiomas, formación, profesión) e información relacionada con la lengua española (nivel de español, estancia en países hispanohablantes, contacto con nativos, frecuencia de lectura en español, entre otros). Este cuestionario contuvo preguntas de alternativas tanto fijas como abiertas y, al estar destinado a informantes vietnamitas, se redactó en su lengua materna para facilitar la comprensión. Se permitió que el informante se tomara el tiempo necesario para rellenarlo, pero en ningún caso más allá de 20 minutos. Las variables de estudio se determinaron con la intención de evaluar los condicionantes sociales que afectaban a la calidad y cantidad de las palabras proporcionadas por los informantes. Tomando como referencia los resultados obtenidos en trabajos anteriores (Carcedo, 1999; Cummins, 1979; Sánchez-Saus Laserna, 2009), además de preguntas generales destinadas a obtener información sobre datos personales, se añadieron, como hemos señalado, otras específicas con la intención de acercarnos y explorar la realidad social de los informantes. Sin embargo, por problemas de espacio, en el presente estudio solo analizaremos la influencia de la *identidad de género* y el *nivel de español* de los estudiantes en los resultados obtenidos.

3.3. Prueba de léxico disponible

Tras completar el cuestionario sociológico, los informantes procedieron a realizar la prueba de léxico disponible en la que debieron escribir, en un tiempo de dos minutos, todas las palabras que les venían en la mente cuando se les planteaba un estímulo cognitivo en forma de núcleo temático o centro de interés. Conviene señalar que se evitó dar explicaciones adicionales y ejemplos que pudiesen influir en las respuestas de los informantes.

Es oportuno aclarar también que, para la recolección de las palabras, se eligió el modo presencial y escrito, no virtual ni oral por varios factores: el ahorro de tiempo al poder realizarse de forma colectiva, el control en la honestidad de las respuestas debido a la presencia de los entrevistadores y la falta de ayuda de programas informáticos como la autocorrección, lo que facilitó un mejor registro, edición de datos y compatibilidad de los ficheros electrónicos finales con las herramientas estadísticas empleadas en las fases posteriores.

3.4. Centros de interés

Hemos mantenido la mayor parte de los dieciséis estímulos clásicos utilizados en las primeras investigaciones sobre léxico disponible, con algunos ajustes de acuerdo con las características propias del alumnado vietnamita. En este sentido, se tomó como referencia la revisión crítica propuesta por Sánchez-Saus Laserna a los centros de interés (Sánchez-Saus Laserna 2009). De este modo, se estimó que determinados centros resultarían poco productivos por tratarse de un contexto de aprendizaje sin inmersión sociocultural, como ocurre en el caso de los centros *Iluminación, calefacción y medios de airear un recinto* o *Trabajos del campo y del jardín*; otros parecen demasiado específicos para resultar productivos: *Objetos colocados en la mesa para la comida, Partes de la casa (sin muebles)* y *Los muebles de la casa*. En definitiva, en el presente trabajo analizamos los siguientes dieciséis centros de interés: 1) *el cuerpo humano*, 2) *la ropa y complementos*, 3) *muebles de la casa*, 4) *partes de la casa*, 5) *comidas y bebidas*, 6) *la cocina y sus utensilios*, 7) *la escuela: muebles y materiales*, 8) *la ciudad*, 9) *medios de transporte*, 10) *los animales*, 11) *entretenimiento y tiempo libre*, 12) *profesiones y trabajos*, 13) *tiempo meteorológico y clima*, 14) *rutina diaria*, 15) *tecnología e internet*, 16) *colores*.

3.5. Edición y tratamiento estadístico de los datos

En el proceso de registro de datos, se establecieron dos corpus: uno integrado por datos ‘en bruto’, donde se mantuvieron exactamente las palabras tal y como las habían escrito los informantes, y otro editado según los criterios que se expresan más adelante. La primera base de datos nos sirvió para llevar a cabo el análisis cualitativo de algunos factores como la detección de errores ortográficos o la observación de las preferencias de los informantes hacia el género de las palabras aportadas, las formas gramaticales o las variaciones diatópicas, entre otros aspectos. La segunda contenía una lista ‘limpia’ y preparada para la obtención de los listados definitivos.

Para la edición de datos, se tuvieron en cuenta los criterios propuestos por Samper Padilla (2003: 205-206) para el PPHDL con adaptaciones propias que se consideran más adecuadas para nuestra muestra. De modo general, los criterios fueron estos:

1. Eliminación de términos repetidos por el mismo informante en el mismo centro de interés.
2. Corrección ortográfica. Decidimos guiarnos principalmente por el Diccionario de la lengua española (DLE) y, además, acudimos al Corpus de Referencia del Español Actual (CREA) en caso de que la palabra no estuviese contenida en el DLE.
3. Unificación ortográfica de palabras con doble ortografía. En principio, seguimos la recomendación del PPHDL de mantener la forma que aparece como lema, no como una variante o aceptación que remite a la primigenia en los diccionarios, especialmente el DLE (*video* por *video/video*). Sin embargo, en algunas ocasiones nos encontramos con que las dos variantes tienen su propia entrada independiente en el DLE; entonces, decidimos lematizar según la forma que más se repite en las encuestas (*beige* por *beige/beis*).
4. Neutralización de variantes puramente flexivas optando por las formas preferidas en los diccionarios (DLE, CREA): infinitivo para los verbos, forma singular, masculina para sustantivos y adjetivos a no ser que sean heterónimos; de darse el caso, se mantienen las dos formas.
5. Tratamiento de formas derivadas. Quedan excluidos de la homogeneización los derivativos que aportan significado semántico a la palabra porque se considerarían vocablos diferentes (*pan/panadería*). Se trata, sobre todo, de palabras formadas por *sufijos emotivos*, que son *diminutivos*, *augmentativos* o *despectivos*. Aunque en el PPHDL se recomienda usar la forma preferida por los diccionarios si se trata de una derivación puramente morfológica, —es decir, que no altere el significado semántico de la palabra o que no le marque matiz adicional—, decidimos finalmente mantener las dos formas, independientemente de la presencia o ausencia de diferencias semánticas.

Pensamos que para un estudio de disponibilidad léxica en nativos hispanohablantes, quizás las palabras formadas por derivación regular no presenten una significativa riqueza léxica puesto que se adquieren según patrones mecánicos productivos; sin embargo, en el caso de estudiantes de ELE, el hecho de que sepan que existen estas herramientas morfológicas para la formación de palabras y que las usen correctamente es un señal de su capacidad lingüística tanto en desde el punto de vista gramatical como léxico, lo que aporta mucho valor a su conocimiento. Por ejemplo, no es indiferente si un alumno escribe únicamente *abrazo* que si actualiza *abrazo*, *abracito* o *abrazote*. Téngase en cuenta, además, que el vietnamita es una lengua muy sintética y muy poco flexiva, con lo que estas variaciones morfológicas pueden ser muy significativas en nuestro caso particular.

6. Agrupación de las formas completas y sus acortamientos bajo una sola entrada donde el empleo del paréntesis indica alternancia de aparición: *tele(visión)*, *bici(cleta)*, *auto(bus)*.
7. Agrupación mediante paréntesis de unidades léxicas que se refieren al mismo concepto, pero que se pueden escribir con una o más palabras. Hemos encontrado abundantes ejemplos de estos casos en los centros de interés ‘Entretenimiento y tiempo libre’ y ‘Rutina diaria’ donde predominan las acciones como (*ir al*) *cine*, (*leer*) *libros*, (*escuchar*) *música*, etc.
8. En casos particulares como los nombres de marcas comerciales, redes sociales o tendencias juveniles, el PPHDL recomienda incluir solo las palabras que ya estén totalmente introducidas en los diccionarios. En nuestro caso, incluimos en la lista definitiva también palabras que no están lexicalizadas oficialmente, aunque su uso está muy extendido (*Facebook*, *Google*).

Una vez realizada la lematización de todas las formas, se codificaron los datos según un plan de codificación anteriormente elaborado y se procesaron mediante el programa *DispogenW* para obtener tanto las listas de disponibilidad generales por centros de interés como las específicas según las variables sociológicas analizadas.

4. ESTUDIO EXPLORATORIO GENERAL

En este apartado se analizan los índices generales del léxico disponible de los 45 estudiantes en relación con los dieciséis centros de interés usados y, posteriormente, se observa el posible efecto de las variables sociales consideradas en el presente estudio en la producción léxica de estos alumnos que, recordemos, son *Identidad de género* y *Nivel de español*. No debemos olvidar que el análisis que efectuamos es meramente exploratorio y descriptivo general, sin pretensiones estadísticas de representatividad, con lo que nuestro propósito se centra exclusivamente en buscar tendencias que puedan ser corroboradas en posteriores fases de esta investigación.

4.1. Índices generales

Los índices generales, presentados en la Tabla 2, incluyen el total de palabras (es decir, cada una de las respuestas editadas, también denominada ‘unidad léxica’ por otros autores (Paredes, 2015; Rubio, 2018) y el número de vocablos (cada entrada diferente, es decir, sin repeticiones en el mismo informante ni entre todos ellos), el promedio de palabras por informante y el índice de cohesión obtenidos en cada centro de interés explorado. Además, también se aporta la suma total palabras y vocablos de todos los centros.

| | Total de palabras | Número de palabras | Promedio de palabras | Índice de cohesión |
|---|-------------------|--------------------|----------------------|--------------------|
| <i>El cuerpo humano</i> | 480 | 86 | 10,67 | 0,12 |
| <i>La ropa y complementos</i> | 441 | 90 | 9,80 | 0,11 |
| <i>Muebles de la casa</i> | 399 | 107 | 8,87 | 0,08 |
| <i>Partes de la casa</i> | 317 | 61 | 7,04 | 0,12 |
| <i>Comidas y bebidas</i> | 694 | 170 | 15,42 | 0,09 |
| <i>La cocina y sus utensilios</i> | 248 | 66 | 5,51 | 0,08 |
| <i>La escuela: muebles y materiales</i> | 379 | 112 | 8,42 | 0,08 |
| <i>La ciudad</i> | 527 | 184 | 11,71 | 0,06 |
| <i>Medios de transporte</i> | 343 | 70 | 7,62 | 0,11 |
| <i>Los animales</i> | 457 | 93 | 10,16 | 0,11 |
| <i>Entretenimiento y tiempo libre</i> | 535 | 196 | 11,89 | 0,06 |
| <i>Profesiones y trabajos</i> | 447 | 139 | 9,93 | 0,07 |
| <i>Tiempo meteorológico y clima</i> | 378 | 84 | 8,40 | 0,10 |
| <i>Rutina diaria</i> | 594 | 166 | 13,20 | 0,08 |
| <i>Tecnología e internet</i> | 374 | 185 | 8,31 | 0,04 |
| <i>Los colores</i> | 447 | 44 | 9,93 | 0,23 |
| PROMEDIO | 441,25 | 115,81 | 9,81 | 0,10 |
| TOTAL | 7060 | 1853 | 156,89 | 0,10 |

Tabla 2. *Índices generales*

Se actualizaron un total de 7060 palabras en los dieciséis centros de interés estudiados. Los centros más y menos productivos fueron ‘Comidas y bebidas’ y ‘La cocina y sus utensilios’, con 694 y 248 palabras, respectivamente. Nótese que las 248 palabras encontradas en el centro de interés ‘La cocina y sus utensilios’ quedan bastante distantes en relación con el promedio de las palabras producidas por centro de interés (441,25). Estos dos centros de interés también ocupan las posiciones extremas respecto al promedio de palabras proporcionadas por informante, con 15,42 y 5,51 palabras respectivamente; esta última cifra se encuentra casi en la mitad por debajo de la media, que corresponde a 9,81 palabras.

Asimismo, se producen más palabras diferentes (*vocablos*) en ‘Entretenimiento y tiempo libre’ y menos en ‘Los colores’, con 196 y 44 entradas, respectivamente. Esto demuestra, una vez más, un intervalo considerable en comparación con el promedio de vocablos actualizados por centro de interés, que es 115,81 palabras.

El índice de cohesión permite medir el grado de coincidencia entre las respuestas de los informantes en un mismo centro de interés, es decir, nos ayuda a saber si las palabras son conocidas por muchos alumnos o solo por unos pocos. Se calcula dividiendo el promedio de palabras por informante entre el número de vocablos en cada centro de interés. El resultado varía entre 0 y 1. Cuanto más cerca de 1 esté el valor, mayor cohesión dispone, esto significa que más informantes activan las mismas palabras y el centro de interés es más compacto. Este índice nos podría dar pistas de si se ha trabajado lo suficiente en el léxico de un determinado tema en clase o se debería consolidar más. La Tabla 2 nos muestra que el centro de interés ‘Los colores’

obtiene el mayor índice de cohesión (0,23) y ‘Tecnología e internet’ el menor índice de cohesión (0,04).

4.2. Factores condicionantes del léxico disponible del alumnado vietnamita

Como hemos anunciado anteriormente, hemos intentado determinar la influencia de dos factores que pudieran condicionar al léxico disponible de los alumnos vietnamitas de ELE mediante el análisis exploratorio por variables sociales: *Identidad de género* y *Nivel de español*.

4.2.1. Identidad de género

La Tabla 3 y el Gráfico 1 muestran el promedio de palabras por informante en función de la variable *Identidad de género*.

| | M | H |
|---|--------------|--------------|
| <i>El cuerpo humano</i> | 9,69 | 14,1 |
| <i>La ropa y complementos</i> | 9,77 | 9,9 |
| <i>Muebles de la casa</i> | 8,34 | 10,7 |
| <i>Partes de la casa</i> | 6,71 | 8,2 |
| <i>Comidas y bebidas</i> | 16,03 | 13,3 |
| <i>La cocina y sus utensilios</i> | 5,40 | 5,9 |
| <i>La escuela: muebles y materiales</i> | 8,46 | 8,3 |
| <i>La ciudad</i> | 11,43 | 12,7 |
| <i>Medios de transporte</i> | 7,29 | 8,8 |
| <i>Los animales</i> | 10,06 | 10,5 |
| <i>Entretenimiento y tiempo libre</i> | 11,57 | 13 |
| <i>Profesiones y trabajos</i> | 10,11 | 9,3 |
| <i>Tiempo meteorológico y clima</i> | 8,26 | 8,9 |
| <i>Rutina diaria</i> | 12,83 | 14,5 |
| <i>Tecnología e internet</i> | 7,97 | 9,5 |
| <i>Los colores</i> | 9,54 | 11,3 |
| MEDIA | 9,59 | 10,56 |

Tabla 3. Promedio de palabras por Identidad de género

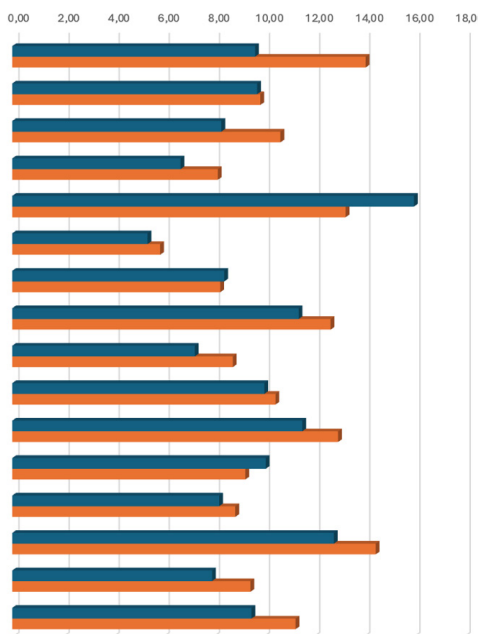


Gráfico 1. Promedio de palabras por Identidad de género

De los dieciséis centros de interés, los hombres superan a las mujeres en número de palabras actualizadas en casi todos los campos analizados, menos en ‘Comidas y bebidas’, ‘La escuela: muebles y materiales’ y ‘Profesiones y trabajos’ aunque con diferencias muy pequeñas: 2,73, 0,16 y 0,81 palabras respectivamente. Se observa que hay bastante similitud entre la media de palabras producidas por mujeres y por hombres: 9,59 y 10,56 palabras respectivamente.

Por centros de interés, las mujeres activan más palabras en ‘Comidas y bebidas’ (16,03 palabras) mientras que los hombres son más productivos en ‘Rutina diaria’ (14,5 palabras). Los dos grupos coinciden al aportar menos respuestas en ‘La cocina y sus utensilios’ con 5,4 palabras por partes de las mujeres y 5,9 palabras por partes de los hombres.

Respecto a las divergencias en las respuestas entre los dos grupos por centro de interés, se registra la mayor distancia en ‘Cuerpo humano’ con 4,41 palabras (9,69 palabras por las mujeres y 14,1 palabras por los hombres). Sin embargo, en ‘La escuela: muebles y materiales’, los dos grupos presentan una cifra muy similar, con solo 0,16 palabra de diferencia (8,46 palabras para las mujeres y 8,3 palabras para los hombres).

No debemos olvidar que el número de mujeres y hombres que compone la muestra es bastante desequilibrado (35 mujeres frente a 10 hombres, Tabla 1), por lo que la variable *Identidad de género* no parece ser un factor que afecte a la producción léxica de los alumnos en función de los resultados obtenidos en este avance exploratorio.

4.2.2. Nivel de español

Debido al momento en que se realizaron las encuestas, a finales del primer semestre del curso académico, a los grupos de segundo, tercero y cuarto año les deberían corresponder los niveles de A2, B1 y B2 según el MCER. En cada uno de los niveles se obtuvieron respuestas de 15 alumnos (Tabla 1). La concordancia en el número de encuestados en los tres grupos permite analizar con intención contrastiva los índices de palabras por centro de interés, vocablos y promedio de palabras por informante.

| | Total palabras por CI | | | Vocablos | | | Promedio de palabras | | |
|---|-----------------------|------------|---------------|--------------|--------------|--------------|----------------------|--------------|--------------|
| | A2 | B1 | B2 | A2 | B1 | B2 | A2 | B1 | B2 |
| <i>El cuerpo humano</i> | 162 | 147 | 171 | 47 | 51 | 46 | 8,53 | 13,36 | 11,40 |
| <i>La ropa y complementos</i> | 203 | 112 | 126 | 51 | 42 | 38 | 10,68 | 10,18 | 8,40 |
| <i>Muebles de la casa</i> | 157 | 116 | 126 | 57 | 51 | 56 | 8,26 | 10,55 | 8,40 |
| <i>Partes de la casa</i> | 139 | 81 | 97 | 39 | 32 | 31 | 7,32 | 7,36 | 6,47 |
| <i>Comidas y bebidas</i> | 343 | 158 | 193 | 126 | 66 | 81 | 18,05 | 14,36 | 12,87 |
| <i>La cocina y sus utensilios</i> | 102 | 66 | 80 | 34 | 38 | 30 | 5,37 | 6,00 | 5,33 |
| <i>La escuela: muebles y materiales</i> | 152 | 88 | 139 | 62 | 45 | 63 | 8,00 | 8,00 | 9,27 |
| <i>La ciudad</i> | 224 | 146 | 157 | 89 | 77 | 92 | 11,79 | 13,27 | 10,47 |
| <i>Medios de transporte</i> | 111 | 102 | 130 | 18 | 31 | 49 | 5,84 | 9,27 | 8,67 |
| <i>Los animales</i> | 227 | 115 | 115 | 71 | 44 | 41 | 11,95 | 10,45 | 7,67 |
| <i>Entretenimiento y tiempo libre</i> | 240 | 141 | 154 | 120 | 74 | 86 | 12,63 | 12,82 | 10,27 |
| <i>Profesiones y trabajos</i> | 204 | 106 | 137 | 77 | 54 | 59 | 10,74 | 9,64 | 9,13 |
| <i>Tiempo meteorológico y clima</i> | 183 | 97 | 98 | 61 | 33 | 37 | 9,63 | 8,82 | 6,53 |
| <i>Rutina diaria</i> | 242 | 158 | 194 | 101 | 69 | 78 | 12,74 | 14,36 | 12,93 |
| <i>Tecnología e internet</i> | 129 | 103 | 142 | 83 | 70 | 85 | 6,79 | 9,36 | 9,47 |
| <i>Los colores</i> | 192 | 120 | 135 | 31 | 24 | 20 | 10,11 | 10,91 | 9,00 |
| MEDIA | 188,13 | 116 | 137,13 | 66,69 | 50,06 | 55,75 | 9,90 | 10,55 | 9,14 |

Tabla 4. Índices generales por nivel de español

Según los datos contenidos en la Tabla 4, referidos al total de palabras producidas por centro de interés en función del nivel de español de los encuestados, los alumnos del nivel A2 muestran el mejor resultado con 188,13 términos, seguidos de los del nivel B2 con 137,13 palabras y, por último, los de nivel B1 con 116. La escala comparativa decreciente sería A2-B2-B1.

El centro de interés más productivo para los estudiantes de nivel A2 es el ‘Comidas y bebidas’ con 343 palabras, lo que crea una diferencia de casi el doble en comparación con las 158 palabras de los estudiantes que cursan el nivel B1 y las 193 de los que se sitúan en el nivel B2 en el mismo centro. A pesar de que el centro de interés más productivo para los estudiantes de nivel B1 y B2 es ‘Rutina diaria’ con 158 y 194 palabras respectivamente, estos no logran superar a los de nivel A2 en cuanto al número de respuestas, que se sitúa en 242 palabras.

El centro de interés que presenta menos respuestas para los tres grupos es ‘La cocina y sus utensilios’, con 102, 66 y 80 palabras que corresponden a los niveles de A2, B1 y B2, respectivamente. La coincidencia que se observa referida al hecho de que los alumnos de los tres niveles revelan capacidad léxica limitada en este estímulo nos sugiere señales de que, quizás, los docentes deberíamos reforzar más esta temática relacionada con experiencias bastante cotidianas.

En lo relativo a los vocablos, se observa un patrón de mayor a menor A2-B2-B1 con una media de 66,69-55,75-50,06 respectivamente. Los centros de interés más y menos productivos para los estudiantes de nivel A2 son ‘Comidas y bebidas’ y ‘Medios de transporte’ con 126 y 18 términos respectivamente. Otra vez, los de nivel B1 y B2 comparten los centros de interés con mayor y menor respuestas: ‘La ciudad’ y ‘Los colores’, con 77 y 92 términos en el primero y 24 y 20 en el segundo estímulo.

Respecto al promedio de palabras actualizadas por informante, se reciben más respuestas en el nivel B1, seguido del nivel A2 y, por último, nivel B2, con una media de 10,55; 9,90 y 9,14 palabras. ‘Comidas y bebidas’ y ‘La cocina y sus utensilios’ constituyen los centros de interés más y menos productivos para los informantes de nivel A2, con una media de 18,05 y 5,37 palabras por sujeto. ‘Comidas y bebidas’ junto a ‘Rutina diaria’ conforman los centros de interés con mayor amplitud léxica, con 14,36 palabras en cada centro; mientras que se muestra una menor productividad en ‘La cocina y sus utensilios’ con una media de 6 palabras. Por último, en el nivel B2 se registran 12,93 palabras por informante en el centro de interés más prolífico (‘Rutina diaria’) y 5,33 palabras en el menos productivo (‘La cocina y sus utensilios’).

Se observa una curiosa tendencia en el sentido de que los estudiantes de nivel A2 presentan mejores resultados en la mayoría de los índices en comparación con los de nivel B1 y B2. Sin embargo, esta mejora no se mantiene en dirección creciente o decreciente gradual por nivel de español, por lo que no podríamos concluir que el nivel de español sea uno de los factores influyentes a la productividad léxica del alumnado.

5. ANÁLISIS CUALITATIVO

En este apartado comentaremos algunos de los rasgos particulares que se observan en las palabras que componen las listas aportadas por los alumnos vietnamitas de ELE. En concreto, prestaremos atención a los siguientes aspectos: características generales de las unidades léxicas, variantes diatópicas, composiciones creativas e influencias de otras lenguas.

5.1. Características generales

Durante la edición y análisis de datos, hemos encontrado abundantes elementos léxicos poco frecuentes por tratarse de términos no presentes en los manuales de ELE empleados en las clases o que se actualizan en ámbitos especializados. En ‘El cuerpo humano’ se hallan palabras que no son básicas ni previsibles de encontrar en situaciones cotidianas como *caño*, *colon*, *torso*, *viscera*, *riego sanguíneo*, *pene*, *vagina*. Lo curioso es que son activadas por alumnos de diferentes niveles, incluso de los de nivel A2. Por un lado, nos sorprende desde el punto de vista docente, puesto que no suelen aparecer en las clases. Por otro lado, la activación de estos términos puede considerarse positiva porque nos permite pensar que parte del alumnado tiene contacto con el español o con hispanohablantes fuera de clase. También se observan términos especializados como *palatal*, *faringe*; es probable que la fuente de este tipo de vocabulario provenga de la asignatura *Fonética y Fonología Española*, presente en su plan de estudios.

En ‘La ropa y complementos’, aparte del vocabulario básico de las prendas y accesorios como *camisa*, *camiseta*, *pantalón*, etc., se localizan unidades léxicas como *sujetador* y *traje de gala*. En primer lugar, respecto a *sujetador*, podría no llamar la atención el uso de esta palabra en hablantes nativos, pero en nuestro caso, por influencia de la cultura asiática, para el alumnado no es cómodo referirse a partes íntimas ni a prendas relacionadas con ellas, de modo que los alumnos suelen recurrir a *ropa interior* como solución colectiva para todo tipo de prendas íntimas. Las especificaciones de *sujetadores*, *bragas* o *calzoncillos* se trabajan en clase, pero su empleo se deja a libre selección del alumno y no forman parte del grupo de palabras indispensables, por lo que los alumnos suelen evitar su uso. Lo mismo ocurre con *pene* y *vagina* anteriormente mencionadas, que son partes íntimas del cuerpo que no suelen aparecer en clase por cuestión cultural.

En segundo lugar, *traje de gala*, es una unidad léxica formada por combinación abierta, es decir, se pueden cambiar sus componentes para formar unidades nuevas como *traje de vestir*, *traje de novio/a*. En clase se le proporciona al alumnado la palabra *traje* en su forma aislada, así que su uso en la forma compuesta nos sugiere un nivel de español más elevado por parte del alumnado que la ha actualizado en sus listas. Se ha revisado el perfil sociológico de quienes lo han hecho y, en efecto, se asocia a estudiantado que cursa nivel B2 de español.

Entre el listado de palabras ‘sorprendentes’, conviene mencionar *horchata* en ‘Comidas y bebidas’. Se refiere a una especie de bebida elaborada con chufa, cuyo zumo se mezcla con agua y azúcar. Nos llama la atención que algunos estudiantes conozcan el término y que lo escriban correctamente, con *h* y con *r*, puesto que es un cultrema que no se suele trabajar en clase. Otra palabra que conviene destacar en este inventario es *caqui*, árbol y fruto de origen japonés y chino. Su aparición en nuestras listas constituye otra muestra más de que algunos estudiantes están enriqueciendo su capacidad léxica de modo personal fuera de clase.

En ‘Profesiones y trabajos’ se registra la palabra *nini* que hace referencia a una persona joven que ni estudia ni trabaja. Esta voz no está lexicalizada en el DLE, lo que nos ha llevado a estudiar el perfil sociológico del encuestado que la

actualiza. El resultado es que posee nivel A2 de español y tiene contacto con nativos hispanohablantes, según se desprende de los datos aportados en el cuestionario sociológico.

5.2. Variantes diatópicas

La lengua española, como cualquier otra, destaca por su riqueza y diversidad léxica. Un mismo concepto puede ser representado por varias palabras dependiendo de diferentes elementos como el contexto donde se realiza la comunicación, el registro que emplean las personas que participan en ella y la zona geográfica en que se documenta. Este último aspecto resulta de enorme interés ya que cada región presenta sus propias características léxicas. En nuestro estudio, se perciben múltiples usos de variantes diatópicas; muchas de ellas son formas asociadas a la norma en países latinoamericanos y, a pesar de que en las clases de ELE en Vietnam se utilizan manuales procedentes de España, hemos registrado algunas variantes diatópicas hispanoamericanas. Para corroborar la procedencia de estos elementos léxicos, hemos acudido al Diccionario de americanismo (DA) y al DEL, lo que nos ha servido para afianzar el origen geográfico de los términos.

En primer lugar, hemos constatado una evidente preferencia por el uso de *computador/a* frente a *ordenador*. La primera forma es más usada en países latinoamericanos, mientras que la segunda es preferida mayormente por hablantes del español en Europa. Además, se constatan otros ejemplos de pares de palabras que denotan el mismo concepto y que aparecen indistintamente en las listas. Hemos documentado casos como *refrigerador/ frigorífico/ nevera* en ‘La cocina y sus utensilios’; *banana/ plátano, papa/ patata, jugo/ zumo, res/ carne de vaca, ron/ licor* en ‘Comidas y bebidas’; *tomar fotos/ hacer fotos* en ‘Entretenimiento y tiempo libre’; *bluyín/ pantalón vaquero, sobretodo/ abrigo/ impermeable* en ‘La ropa y complementos’. Es de reseñar que las variantes de América suelen tener formas semejantes en inglés, *refrigerador/ refrigerator, banana/ banana, tomar fotos/ take photos, bluyín/ blue jean*.

En segundo lugar, nos ha sorprendido el uso exclusivo de variantes procedentes del español de América. Destaca la aparición de *parqueo*, término usado en Bolivia, Colombia, Cuba, El Salvador, Honduras, Nicaragua y Perú para denotar *aparcamiento* en el centro de interés ‘La ciudad’. En ‘Comidas y bebidas’ aparece *guasaya*, vocablo empleado en Honduras para denotar ‘dos frutos que comparten la misma cáscara’ (DLE); *guaba*, propio de Guatemala, Honduras, Ecuador, Nicaragua, Costa Rica, Panamá, Cuba, República Dominicana, Puerto Rico, Perú para referirse al fruto o árbol de *guamo* que, aunque tiene otras variantes ortográficas (*huaba* y *guabo*) en nuestras listas solo aparece bajo la forma de *guaba*; *frijol*, procedente de México, Guatemala, Honduras, Nicaragua, Costa Rica, Panamá, Cuba, República Dominicana, Puerto Rico, Colombia, Ecuador, Perú, Bolivia para referirse a *judía*. En ‘Muebles de la casa’ se presenta *lampa*, usado en Ecuador, Bolivia, Chile, Argentina, para denominar a la *azada*.

Asimismo, documentamos *perico* en ‘Profesiones y trabajos’ con acepciones diferentes en función de la zona geográfica donde se usa. En España es una voz coloquial que denota a una ‘persona, especialmente mujer, que gusta de callejear, y es

a veces de vida desventurada' (DLE) mientras que en Nicaragua se refiere a un aprendiz. En 'Los animales' se registra *mosco*, variante chilena de *caballo* o *yegua*; *cuy*, variante de Argentina, Bolivia, Colombia, Ecuador y Perú que significa *conejillo de Indias*.

Es importante resaltar que, en todos los casos de variantes diatópicas anteriores, los encuestados señalaron en sus cuestionarios sociológicos complementarios que tenían frecuentes contactos con nativos hispanohablantes.

5.3. Composiciones creativas

En las encuestas se ha observado la formación de palabras mediante diferentes procedimientos, unos más ortodoxos que otros, pero todos presentan rasgos significativos para el estudio del léxico disponible de nuestro grupo de informantes.

Se hallan casos en los que, debido a la falta de vocabulario, el informante aplica reglas morfológicas adquiridas para formar nuevas voces que pueden no existir en español normativo. Por ejemplo, algunos informantes combinan una raíz española con un sufijo derivativo, como *escritoria** en vez de *escritora*, *compañador** en lugar de *compañero* en 'Profesiones y trabajos'.

De igual modo, consideramos como una buena señal que, en niveles de español superiores, los alumnos creen unidades léxicas para expresar conceptos sin limitarse a su capital léxico. Se encuentran casos como '*amplificador de voz*' para indicar *altavoz*, '*plano de proyector*' para referirse a '*pantalla de proyector*' en 'La escuela: muebles y materiales', '*edificio de gobierno*' para describir a *ayuntamiento* o *comité popular* en 'La ciudad'.

5.4. Influencia de otras lenguas

También se registran unidades léxicas curiosas por la influencia de otras lenguas. Resulta interesante la combinación de una raíz inglesa con un sufijo derivativo español. Por ejemplo, *gardería**, tomada de *garden* en inglés y el sufijo español *-ería* para referirse a *jardín* en el centro de interés 'Partes de la casa'; *translatoria**, tomada de *translator* en inglés y el sufijo *-ria* en español para referirse a *traductora* en 'Profesiones y trabajos'.

Además, se ha documentado la transferencia morfosemántica del inglés en *arrocero* en 'La cocina y sus utensilios', puesto que en inglés se añade el sufijo *-er* al verbo *cook* para formar *cooker* que significa *el objeto usado para cocinar arroz*. Este mismo patrón ha sido trasladado al español agregando el sufijo *-ero* a la raíz *arroz-* para formar *arrocero* con la intención de designar al objeto, cuando *strictu sensu* se refiere a la persona que cocina. En otros casos los estudiantes recurren a la formación de unidades léxicas por traducción literal del vietnamita. Por ejemplo, en *tren de agua** para denotar *tren de vapor* en 'Medios de transporte'. Resulta que, en la lengua materna, al *tren de vapor* se le llama *tren de agua*; es posible que ante el desconocimiento del término *vapor*, el alumno decidiera traducirlo literalmente.

6. DISCUSIÓN

Presentamos a continuación una recopilación de los hallazgos más relevantes que hemos presentado hasta ahora y su relación con las hipótesis planteadas al principio de nuestro trabajo.

En primer lugar, nuestros resultados confirman que los centros de interés más cercanos a la vida cotidiana de los estudiantes, como “Comidas y bebidas” o “Rutina diaria”, tienden a generar una mayor producción léxica. Esto coincide con estudios previos que señalan la relevancia de los temas familiares y frecuentemente abordados en clase en la productividad léxica del alumnado de ELE (Paredes, 2015; Rubio, 2018). La familiaridad temática facilita la activación del léxico, ya que los estudiantes poseen un mayor bagaje léxico relacionado con estos ámbitos. Por el contrario, los centros de interés menos productivos, como “La cocina y sus utensilios”, reflejan áreas temáticas que, probablemente, no solo son menos tratadas en el aula, sino también menos relevantes para la vida diaria de los estudiantes vietnamitas. Estos resultados subrayan la necesidad de una mayor atención pedagógica en temas que, aunque puedan resultar menos interesantes para el alumnado, son esenciales para un dominio más amplio del idioma.

En segundo lugar, hemos encontrado que los resultados obtenidos relacionados con temas más concretos y con vocabulario más limitado, como “Los colores”, tienden a presentar un mayor índice de cohesión, lo que sugiere que muchos estudiantes comparten el mismo léxico en estas áreas menos productivas. Este hallazgo tiene implicaciones didácticas, ya que demuestra que la enseñanza de léxico muy específico y homogéneo facilita el aprendizaje de vocabulario por parte de los estudiantes. En contraste, temas más amplios y con mayor variabilidad léxica, como “Tecnología e internet”, muestran una menor cohesión, lo que podría motivar el hecho de que los estudiantes activen un repertorio léxico más diversificado y disperso. Esta falta de cohesión podría deberse a que estos temas no han sido suficientemente tratados en clase o no cuentan con el mismo grado de interés por parte de los estudiantes fuera del aula.

En tercer lugar, en relación con la influencia de las variables sociales estudiadas, los resultados exploratorios obtenidos respecto a la identidad de género de los participantes revelan que, en general, no existen diferencias significativas en la producción léxica entre hombres y mujeres. Aunque en algunos centros de interés, como “Cuerpo humano” o “Comidas y bebidas”, se observan pequeñas variaciones, estas no son lo suficientemente marcadas como para considerar que esta variable influya en la cantidad o diversidad del léxico disponible. Estos hallazgos coinciden con la literatura previa (Rubio, 2018), que también concluye que el género de estudiantes de ELE no suele ser un factor determinante en estudios de léxico disponible.

En cuarto lugar, el análisis de la influencia del nivel de español en los resultados obtenidos ha generado resultados más complejos a los esperados. Si bien se observó una mayor productividad léxica en el nivel A2 en comparación con los niveles B1 y B2, esta tendencia no sigue en absoluto la progresión esperada. Los estudiantes de nivel B2, que deberían mostrar una competencia lingüística más avanzada que el resto, no superan de manera significativa a los del nivel A2 en muchos casos, lo que obliga a reflexionar sobre posibles factores externos en la capacidad léxica de los informantes, como el contacto con hablantes nativos fuera del aula, la motivación

personal, o incluso la propia metodología de enseñanza. Este hallazgo invita a una revisión crítica de cómo se estructuran los niveles de enseñanza en ELE y qué otros elementos, además del nivel de instrucción formal, pueden estar influyendo en la productividad léxica de los estudiantes.

Creemos que los resultados de este estudio tienen claras implicaciones para la enseñanza del español como lengua extranjera en contextos similares a los aquí considerados. En primer lugar, subrayan la importancia de reforzar temas menos productivos como “La cocina y sus utensilios” para equilibrar la competencia léxica de los estudiantes en diferentes áreas temáticas. Además, la baja cohesión en algunos centros de interés sugiere que ciertos temas podrían beneficiarse de un enfoque más sistemático en el aula, para garantizar que todos los estudiantes manejen un repertorio léxico común y suficiente. Finalmente, la variabilidad en la productividad léxica entre niveles de español refuerza la necesidad de atender a factores adicionales, como la exposición al idioma fuera del aula, que pueden ser tan influyentes como el nivel formal en el aprendizaje del léxico.

7. CONCLUSIONES

En este estudio hemos presentado una primera exploración del léxico disponible de estudiantes vietnamitas de Español como Lengua Extranjera (ELE) que nos ha permitido analizar las características de las palabras contenidas en listados obtenidos. El análisis efectuado nos ha permitido comprobar la consistencia de varias de las hipótesis planteadas al inicio de la investigación.

Productividad léxica por centros de interés: Los resultados confirman que la familiaridad y la relevancia práctica del léxico en la vida cotidiana de los estudiantes influyen significativamente en su producción léxica, tal y como se había anticipado. El centro de interés “Comidas y bebidas” resultó ser el más productivo en términos tanto de la cantidad total de palabras obtenidas como del promedio de palabras por informante, lo cual respalda la hipótesis de que los temas más cercanos a la vida diaria de los estudiantes generan mayor producción léxica. Por otro lado, el centro “La cocina y sus utensilios” fue el menos productivo, apoyando la idea de que los temas menos familiares provocan menor productividad léxica.

Cohesión léxica: Los resultados también apoyan la hipótesis de partida referida a la cohesión léxica. El tema “Los colores” presentó el mayor índice de cohesión, con una alta coincidencia entre las respuestas de los estudiantes, lo que indica una mayor homogeneidad en el aprendizaje de este campo semántico. En contraste, “Tecnología e internet” mostró el menor índice de cohesión, con respuestas más variadas entre los informantes, lo que confirma que los temas más amplios y menos trabajados en clase tienden a mostrar menor cohesión léxica.

Diferencias de género: La hipótesis de que el género no sería un factor determinante en la producción léxica también fue confirmada. El análisis de las palabras aportadas por hombres y mujeres no reveló diferencias exploratorias significativas, lo que refuerza la idea de que la identidad de género no tiene un impacto relevante en la cantidad ni en la diversidad del léxico disponible en estudiantes vietnamitas de ELE.

Nivel de español: Aunque se esperaba que los estudiantes de niveles superiores mostraran una mayor productividad léxica, los resultados obtenidos obligan a matizar esta hipótesis. La mayor producción léxica encontrada entre los estudiantes de nivel A2 en comparación con los de nivel B1, refuta este supuesto. Además, esta tendencia no se mantuvo entre los niveles B1 y B2, ya que estos últimos sí que manifestaban mayores índices de producción léxica. Esto sugiere que, aunque el nivel de español influye en la productividad, factores como la motivación o el contacto con el español o con hispanohablantes fuera de clase podrían jugar un papel más determinante, tal como se había planteado en la hipótesis inicial.

Influencia de otras lenguas: El análisis cualitativo corroboró la influencia de otras lenguas, como el vietnamita y el inglés, en la producción léxica de los estudiantes. Se encontraron desviaciones léxicas y composiciones creativas que indican la presencia de transferencias lingüísticas desde las lenguas señaladas. Este resultado coincide con la hipótesis de que el bilingüismo o el uso de una tercera lengua vehicular (inglés) puede afectar a la adquisición y producción léxica en español. Además, se documentó la presencia de variantes diatópicas no contenidas en los manuales ni trabajadas en clase, lo que sugiere que los estudiantes están en contacto con hablantes nativos de español fuera del aula, lo cual es un dato significativo para la planificación pedagógica.

Hemos de tener en cuenta, no obstante que, aunque los hallazgos encontrados en nuestro estudio proporcionan una primera aproximación útil al léxico disponible de los estudiantes vietnamitas de ELE, la investigación presenta ciertas limitaciones. En primer lugar, la muestra de estudiantes es limitada y puede no ser representativa de la totalidad de los aprendices de español en Vietnam, lo que afecta a la posibilidad de generalización de los resultados que hemos encontrado. Además, un aspecto relacionado con la imposibilidad de generalizar nuestros resultados se refiere a la falta de un análisis estadístico exhaustivo que permita corroborarlos y validarlos de manera más rigurosa. Si bien se identificaron tendencias generales relacionadas con la productividad y cohesión léxica, la ausencia de pruebas estadísticas dificulta la confirmación de la significancia de tales diferencias entre niveles, géneros y centros de interés. Esto limita, como es lógico, la capacidad de inferir con precisión la magnitud de estas diferencias en una población más amplia. Además, el nivel de exposición de los estudiantes a la lengua española fuera del aula no se midió con precisión, lo que podría afectar a la comprensión del impacto del contacto con hablantes nativos en su producción léxica. Aunque se identificaron influencias de otras lenguas, un análisis más detallado del grado y tipo de interferencia lingüística podría haber proporcionado una visión más completa de este fenómeno, pero las limitaciones de espacio nos han impedido obtenerla. En definitiva, en futuros estudios sería recomendable ampliar la muestra, realizar un análisis estadístico robusto y llevar a cabo un análisis más profundo de la interacción entre el léxico y las lenguas en contacto para obtener resultados más concluyentes.

A pesar de las limitaciones señaladas, consideramos que los hallazgos de este estudio proporcionan información valiosa para los docentes de ELE en Vietnam que podrían extrapolarse a otras sintopías. Los resultados resaltan la importancia de adaptar los materiales de enseñanza a las necesidades léxicas y contextuales de los estudiantes, teniendo en cuenta no solo su nivel formal de español, sino también otros factores relacionados con la cultura del alumnado, su exposición a variedades diatópicas, el

contacto con hablantes nativos, la influencia de otras lenguas y el interés personal en determinados centros de interés más relacionados con actividades cotidianas. Además, las composiciones léxicas creativas observadas sugieren que la enseñanza del español puede beneficiarse de un enfoque que fomente la experimentación y el uso flexible del lenguaje, permitiendo a los estudiantes desarrollar una competencia comunicativa más productiva y variada.

8. REFERENCIAS

- ÁVILA MUÑOZ, A. M. Y VILLENNA PONSODA, J. A. 2010. *Variación social del léxico disponible en la ciudad de Málaga*. Málaga: Editorial Sarriá.
- ÁVILA MUÑOZ, A. M. Y SÁNCHEZ SÁEZ, J. M. 2011. La posición de los vocablos en el cálculo del índice de disponibilidad léxica: procesos de reentrada en las listas del léxico disponible de la ciudad de Málaga. *Estudios de Lingüística de la Universidad de Alicante (ELUA)*, 25: 45-74.
- CARCEDO GONZÁLEZ, A. 1998. Sobre las pruebas de disponibilidad léxica para estudiantes de español/LE. *RILCE*, 14(2): 205-224.
- CARCEDO GONZÁLEZ, A. 1999. Análisis de errores léxicos del español en la interlingua de los finlandeses. En M. C. Losada Aldrey *et al.*, (coords.), *Español como lengua extranjera: enfoque comunicativo y gramática*, pp. 465-472. Actas del IX Congreso Internacional de ASELE, Santiago de Compostela, 23-26 de sep- tiembre de 1998. Santiago de Compostela: Universidad de Santiago de Compostela.
- CUMMINS, J. 1979. Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research*, 49(2): 222-251.
- GARCÍA NARANJO, J. 2019. *Estudio de disponibilidad léxica en España y Japón mediante encuestas online*. Tesis doctoral. Universidad Nacional de Educación a Distancia.
- GOUGENHEIM, G.; MICHÉA, R.; RIVENC, P. Y SAUVAGEOT, A. 1956. *L'élaboration du Français Élémentaire (I Degré). Étude sur l'établissement d'un vocabulaire et d'une grammaire de base*. Paris: Didier.
- GOUGENHEIM, G.; MICHÉA, R.; RIVENC, P. Y SAUVAGEOT, A. 1964. *L'élaboration du Français Fondamental (I Degré). Étude sur l'établissement d'un vocabulaire et d'une grammaire de base*. Paris: Didier.
- HERNÁNDEZ CAMPOY, J. M. Y ALMEIDA, M. 2005. *Metodología de la investigación sociolingüística*. Albolote: Comares.
- HIDALGO GALLARDO, M. 2019. *Factores y fuentes que inciden en el léxico disponible de estudiantes sinohablantes de ELE*. ASELE. Colección: Monografías (Tesis Doctorales), 23.
- JINDONG, SONG. 2018. El léxico disponible de alumnos chinos de español: estudio basado en seis centros de interés. IX Congreso Internacional de la Asociación Asiática de Hispanistas (Bangkok, 2016). *Monográficos SinoELE* 17: 373-379.
- LIN, JING. 2012. El estudio de disponibilidad léxica de los estudiantes chinos de español como lengua extranjera. *MarcoELE Revista de didáctica ELE*, 14: 1-14.
- LÓPEZ CHÁVEZ, J. 1992. Alcances panhispánicos del léxico disponible. *Lingüística*, 4: 26-124.
- LÓPEZ MORALES, H. 1973. *Disponibilidad léxica en escolares de San Juan*. MS.

- LÓPEZ MORALES, H. 1979. Disponibilidad léxica y estratificación socioeconómica. López Morales, H. (ed.), *Dialectología y sociolingüística. Temas puertorriqueños*, pp. 173-181. Madrid, Miami, New York, San Juan: Hispanova de Ediciones.
- LÓPEZ MORALES, H. 1999. *Léxico disponible de Puerto Rico*. Madrid: Arco/Libros.
- MACKEY, W. F. 1971. *Le sondage dans les enquêtes de disponibilité*. Quebec: CIRB, Université eLaval.
- MARTÍNEZ OLMO, E. 2015. *El léxico disponible de las comunidades de habla alicantinas. Estudio sociolingüístico y diccionarios*. Alicante: Publicaciones de la Universitat d'Alacant.
- MENDOZA PUERTAS, J. D. 2018. El léxico disponible de 82 estudiantes coreanos de español como lengua extranjera. MarcoELE. *Revista de didáctica ELE*, 26: 1-26. Disponible en <https://www.redalyc.org/journal/921/92155180004/92155180004.pdf> [Consulta 15/05/20224]
- MICHÉA, R. 1953. Mots fréquents et mots disponibles. Un aspect nouveau de la statistique du langage. *Les Langues Modernes*, 47: 338-344.
- PAREDES GARCÍA, F. 2015. Disponibilidad léxica y enseñanza de ELE: el léxico disponible como fuente curricular y como recurso en el aula. V Jornadas de Lengua y Comunicación. Léxico: enseñanza e investigación. Linred: *Lingüística en la Red*, 13. Disponible en <http://hdl.handle.net/10017/25257> [Consulta 18/06/2024].
- REAL ACADEMIA ESPAÑOLA. Banco de datos (CREA) [en línea]. Corpus de referencia del español actual. Disponible en <http://www.rae.es> [Consulta 25/05/2024]
- REAL ACADEMIA ESPAÑOLA. DICCIONARIO DE LA LENGUA ESPAÑOLA, 23.ª ed., [versión 23.7 en línea]. Disponible en <https://dle.rae.es> [Consulta 25/05/2024]
- RUBIO LASTRA, M. 2018. Disponibilidad léxica de 52 estudiantes taiwaneses universitarios de ELE A1. MarcoELE. *Revista de Didáctica Español Lengua Extranjera*, 27. España. Disponible en: <https://www.redalyc.org/articulo.oa?id=92155498009> [Consulta 20/05/2024]
- SAMPER HERNÁNDEZ, M. 2002. *Disponibilidad léxica en alumnos de español como lengua extranjera*. ASELE. Colección: Monografías (Tesis Doctorales), 4.
- SAMPER PADILLA, J. A. 1999. Léxico disponible y variación dialectal: datos de Puerto Rico y Gran Canaria. En Amparo Morales et al. (eds.), *Estudios de lingüística hispánica. Homenaje a María Vaquero*, pp. 550-573. San Juan de Puerto Rico: Universidad de Puerto Rico.
- SAMPER PADILLA, J. A. 2003. *El proyecto panhispánico de disponibilidad léxica: logros y estado actual. VII Jornadas de lingüística. Universidad de Cádiz. Servicio de publicaciones*. pp. 193-226.
- SAMPER PADILLA, J. A. 2020. El factor edad en los estudios de disponibilidad léxica. *Linred: Lingüística en la Red*, 17: 1-20.
- 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.
- VIDA CASTRO, MATILDE; ÁVILA MUÑOZ, A. M. Y CARRISCONDO ESQUIVEL, F. M. 2016. *Manual práctico de sociolingüística*. Red de bibliotecas Universitarias REBIUN. ISBN: 978-84-9077-380-2.