diacrítica
Revista do Centro de Estudos Humanísticos

Vol. 40, n.º 1, 2026, pp. 75–98. doi: 10.21814/diacritica.6871

Variación diatópica y léxico disponible: un análisis desde la centralidad léxica

Variação diatópica e léxico disponível: uma análise com base na centralidade lexical

Diatopic variation and available vocabulary: an analysis based on lexical centrality

Vicente Martínez Aránguiz[1]
0009-0009-6246-7961

Maribel Serrano Zapata[2]
0000-0003-4378-4389

[1] Facultad de Filosofía y Letras, Universidad de Málaga, España. Facultad de Letras, Pontificia Universidad Católica de Chile, Chile.
[2] Departamento de Filología y Comunicación, Facultad de Letras, Universitat de Lleida, España.

Autor correspondiente: vcmartinez@uma.es

Resumen. En el presente trabajo comparamos el léxico disponible de dos grupos de estudiantes preuniversitarios en torno a dos centros de interés (*Alimentos y bebidas* y *Juegos y distracciones*). Por un lado, el análisis cuantitativo de las respuestas de 360 estudiantes, 180 de Lleida (España) y 180 de Santiago de Chile (Chile), indica que el alumnado catalán tiene unos promedios superiores de cantidad de respuestas actualizadas, mientras que el chileno obtiene un mayor número de vocablos y, en consecuencia, un vocabulario más heterogéneo. Los índices de compatibilidad entre los dos inventarios son del 41 % y del 43 % considerando las 100 primeras palabras. Por otro lado, el análisis cualitativo se realiza a partir de la teoría de la centralidad léxica, lo que supone una novedad en este tipo de investigaciones, evidenciando algunas disimilitudes entre los hábitos alimenticios y de ocio de cada grupo: en Chile se detecta una alimentación menos saludable, y en Lleida, un mayor uso de actividades de ocio vinculadas al uso de dispositivos electrónicos.

Palabras clave: Disponibilidad léxica. Dialectología. Léxico disponible.

Resumo. No presente trabalho, comparamos o léxico disponível de dois grupos de estudantes pré-universitários em torno de dois centros de interesse (Alimentos e bebidas e Jogos e distrações). Por um lado, a análise quantitativa das respostas de 360 estudantes, 180 de Lleida (Espanha) e 180 de Santiago do Chile (Chile), indica que os estudantes catalães têm médias superiores em quantidade de palavras atualizadas, enquanto os chilenos obtêm um maior número de palavras e, consequentemente, um vocabulário mais heterogéneo. Os índices de compatibilidade entre os dois inventários são de 41 % e 43 %, considerando as 100 primeiras palavras. Por outro lado, a análise qualitativa é realizada a partir da teoria da centralidade lexical, o que constitui uma novidade neste

tipo de investigação, evidenciando algumas diferenças entre os hábitos alimentares e de lazer de cada grupo: no Chile, detecta-se uma alimentação menos saudável e, em Lleida, um maior recurso a atividades de lazer relacionadas com o uso de dispositivos eletrónicos.

Palavras-chave: Disponibilidade lexical. Dialetologia. Léxico disponível.

Abstract. In this study, an investigation is made into the available vocabulary of two groups of pre-university students around two areas of interest (*Food and Beverages* and *Games and Entertainment*). A quantitative analysis of the responses of 360 students, 180 from Lleida (Spain) and 180 from Santiago of Chile (Chile), was conducted. The results indicated that Catalan students exhibited higher productivity averages of words, while Chilean students obtained a greater number of unique words and, consequently, a more diverse vocabulary. The compatibility indices between the two inventories are 41% and 43% when considering the initial 100 words. Conversely, the qualitative analysis is predicated on lexical centrality theory, a novel concept in this field of research, which reveals discrepancies between the dietary and leisure habits of the two groups. In Chile, a less healthy diet is observed, while in Lleida, a greater prevalence of leisure activities involving electronic devices is evident.

Keywords: Lexical availability. Dialectology. Available lexicon.

1. Introducción

Una de las posibles aplicaciones de la disponibilidad léxica son los estudios de variación dialectal (López Morales, 1995). Desde el trabajo clásico de López Chávez (1992), se han realizado investigaciones que buscan comparar el léxico disponible de distintas comunidades hispánicas con el objetivo de conocer cuál es el vocabulario compartido (Bartol Hernández, 2022; Samper Hernández, 2008; Samper Padilla, 1999). Sin embargo, esta línea de estudio se ha encontrado con la dificultad de decidir qué cantidad de léxico disponible entra en consideración a la hora de realizar este tipo de cotejos y qué criterio es el más idóneo en este sentido, como veremos a continuación.

En la presente contribución, se propone una solución a este inconveniente abordando el estudio del léxico común desde una perspectiva novedosa: la de la centralidad léxica (Villena Ponsoda et al., 2010). Para ello, de acuerdo con las pautas del Proyecto Panhispánico de Disponibilidad Léxica, se analiza un corpus correspondiente a 360 estudiantes preuniversitarios de dos sintopías distintas, Santiago de Chile (Chile) y Lleida (España), en relación con dos centros de interés —*Alimentos y bebidas* y *Juegos y distracciones*—. Dado que el propósito de este trabajo es comparar dos repertorios léxicos, hemos elegido dos áreas donde pudieran evidenciarse de forma clara diferencias culturales. Por tanto, a lo largo de estas páginas, pretendemos responder a las siguientes preguntas:

a) ¿Qué diferencias cuantitativas y cualitativas existen entre la producción de léxico disponible en Lleida (España) y Santiago de Chile (Chile)?

b) ¿Cómo se puede abordar el estudio dialectológico del léxico disponible a partir de la teoría de la centralidad léxica?

Cabe notar que las dos comunidades de hablantes comparadas corresponden a dos variedades distintas del español. La zona de Lleida se adscribe a la variedad castellana, caracterizada por su conservadurismo fonético, al mismo tiempo que es posible reconocer rasgos procedentes del contacto con el catalán en todos los niveles lingüísticos, por ser una zona de bilingüismo social extenso que cuenta con una lengua vernácula propia. La variedad chilena, una de las variedades americanas del español, es, por el contrario, una de las más innovadoras, con rasgos fonéticos de relajación y asibilación de varias consonantes. En lo que respecta al plano léxico —eje central de este trabajo—, en Lleida es posible encontrar un vocabulario de base común de la zona septentrional española, además de préstamos e interferencias de la lengua catalana, muchos de ellos compartidos con otras zonas colindantes, como Aragón o Valencia. En Santiago, se utilizan palabras y usos dialectales propios de la variedad chilena, así como de la andina; en el vocabulario también se observa la influencia de las lenguas indígenas, sobre todo del mapuche y del quechua. Las dos variedades que estudiamos no solo están alejadas geográficamente, sino que tampoco tienen una cercanía cultural, demográfica o histórica. No cabe esperar, por tanto, una gran compatibilidad entre inventarios.

2. Marco teórico

En el presente apartado abordamos los dos ejes teóricos sobre los que se sustenta este trabajo: en primer lugar, las comparaciones intermuestrales dialectológicas llevadas a cabo desde la disponibilidad léxica y, en segundo lugar, la teoría de la centralidad léxica y sus posibles aplicaciones en distintos campos.

2.1. Los cotejos entre repertorios de léxico disponible

Los estudios de disponibilidad léxica aplicados a la dialectología han tenido dos grandes líneas: el análisis de los dialectalismos que aparecen en las encuestas de cada territorio y el cotejo de distintos repertorios con la finalidad de establecer el léxico disponible común. Esta segunda perspectiva es la que adoptamos en el presente trabajo. Desde el principio, esta línea de estudio ha contado con bastantes contribuciones. Tras comprobar que el análisis del léxico disponible arrojaba porcentajes bajos de léxico compartido (López Chávez, 1992), se constató que, estableciendo medidas de corte en el cotejo, los resultados eran muy superiores (Alba, 1998, 2000; Bartol Hernández, 2008; Carcedo González, 2003, 2004; González Martínez, 2000–2001; González Martínez & Orellana Ramírez, 2000; Samper Hernández, 2008; Samper Padilla, 1999). El debate sobre cuál es la forma idónea de acotar el léxico que debe someterse a la comparación interdialectal sigue abierto. Hasta el momento se han adoptado varias soluciones consistentes en limitar los repertorios léxicos, utilizando para ello el número de palabras (ya sean las 20, 50 o 100 primeras), la frecuencia acumulada por estas (el 75 u 80 %), el índice de disponibilidad que alcanzan (el 0,01; 0,02 o 0,05) o el número de informantes (descartando los términos que aporta una sola persona) (Bartol Hernández, 2020; López González, 2014; Samper Padilla & Samper Hernández, 2007). Tampoco es infrecuente

tener en cuenta más de un criterio (Bartol Hernández & Borrego Nieto, 2019) y observar las diferencias que se aprecian en este sentido (López González, 2014, pp. 186–245).

Así pues, se ha constatado que el porcentaje de vocabulario compartido, considerando algún corte en los listados, es siempre alto, de más del 50 %. Este hecho se ha comprobado tanto en los cotejos entre distintas regiones de España (Prado & Galloso, 2015; Samper Hernández, 2008; Samper Padilla et al., 2003) como en las comparaciones que incluyen territorios de ambos lados del Atlántico (Alba, 1998, 2000; Carcedo González, 2003, 2004; González Martínez, 2000–2001; Samper Padilla, 1999), aunque las cifras son superiores en el primer caso (Carcedo González, 2004; Samper Padilla & Samper Hernández, 2007). Los porcentajes altos de convergencia se siguen encontrando en trabajos recientes: Samper Padilla et al. (2019), por ejemplo, comparan la producción léxica de escolares de 6 años de Gran Canaria, México y Costa Rica en torno a tres centros de interés y concluyen que predominan los vocablos comunes en el léxico más disponible. Bartol Hernández (2022) llega a una conclusión similar con los datos de 20 países hispanohablantes alrededor de *Animales*, un centro especialmente homogéneo. Algunas de las investigaciones actuales donde se comparan bancos léxicos de disponibilidad se realizan en torno al Proyecto del léxico disponible de España (Bartol Hernández & Borrego Nieto, 2019) o al Proyecto de enseñanza del español en el mundo hispánico (Bartol Hernández, 2022). Suelen centrarse en el análisis de un solo centro de interés, como *La ropa* (Bartol Hernández, 2020; Bartol Hernández & Borrego Nieto, 2019), *Animales* (Bartol Hernández, 2022) o *El mar* (Medina Peñate, 2022).

Tienen cierta relevancia, para el trabajo que nos ocupa, los cotejos dialectales que incluyen alguna de las comunidades estudiadas. Con respecto a Lleida, no hay ningún estudio —que conozcamos— donde se compare el vocabulario recogido en esta provincia con alguna otra[1], mientras que sí ha sido más frecuente la incorporación del léxico disponible de Chile, basándose en los datos de Valencia Espinoza y Echeverría (1999). Alba (1998, 2000) lo considera, junto con el de República Dominicana, Puerto Rico, Madrid y México. Carcedo González (2004) también lo incluye en su contraste del léxico de Asturias con el de Cádiz, Zamora, Puerto Rico y República Dominicana. Valencia Espinoza (2011), en una investigación con objetivos muy similares a los nuestros, compara el vocabulario de Chile con el de Aragón. Los datos chilenos corresponden a 2052 informantes de todo el país; en cambio, los de Aragón retratan el léxico disponible de esta comunidad autónoma y se basan en 417 informantes. La investigadora chilena analiza la compatibilidad en cada centro de interés considerando las 100 palabras más disponibles y analiza los vocablos compartidos y exclusivos de cada listado. En la comparación de los promedios de producción, se observa que los resultados son siempre superiores en la muestra aragonesa, salvo en dos centros de interés (*Partes de la casa* y *Trabajos del jardín*). En relación con los centros analizados en el presente trabajo, en *Alimentos y bebidas*, las dos sintopías comparten el 45 % de vocabulario; mientras que, en *Juegos y distracciones*, esta cifra desciende al 41 %. La autora comenta también las variantes léxicas halladas en cada campo, con las que se alude a una misma realidad, así

[1] El estudio de Samper Padilla y Samper Hernández (2007) sí la incluye, pero se basa en los datos de 48 informantes de la ciudad de Lleida, puesto que en ese momento aún no se habían publicado los resultados provinciales.

como la procedencia del vocabulario, y observa una mayor presencia de extranjerismos en los listados chilenos. En las conclusiones comenta que:

> …[e]ventualmente, se presentarían problemas comunicativos al hablar de la ropa, alimentación y cocina …. En cuanto a los *Juegos y diversiones*, si desean interactuar aragoneses y chilenos, seguramente necesitarían describirlos, pues en la mayoría de los casos se trata de diferencias de denominación. (Valencia Espinoza, 2011, p. 197)

En último lugar, cabe mencionar el estudio de Gómez Devís et al. (2023), donde se comparan datos de Valencia (España) y Talca (Chile), en un trabajo dedicado al centro de interés *Partes del cuerpo*, cuyo objetivo es el análisis de la organización del lexicón mental.

2.2. La teoría de la centralidad léxica y sus aplicaciones

De acuerdo con lo mencionado anteriormente, la disponibilidad léxica ha supuesto un método eficaz para aproximarse a diversos fenómenos sociales y comunidades de habla, siendo prueba de ello la prolífica producción académica asociada con este campo de estudio. Sin embargo, durante los últimos años, algunos investigadores han desarrollado un enfoque novedoso para estudiar el léxico disponible, con el propósito de superar algunos escollos teórico-metodológicos que impiden a la disponibilidad léxica pasar de un nivel descriptivo a una explicación profunda sobre la relación entre el caudal léxico actualizado y la comunidad de habla particular.

Así, considerando los aportes de la teoría de los prototipos (Rosch, 1978) y la teoría matemática de los conjuntos difusos (Zadeh, 1965), Ávila Muñoz y Villena Ponsoda (2010) proponen la adaptación del concepto de disponibilidad léxica al de centralidad, con el propósito de transitar desde un indicador de las palabras (disponibilidad) a un atributo propio de las personas (compatibilidad). Este concepto sigue una premisa fundamental que establece que, por razones socioculturales, en los diferentes grupos o comunidades de habla existiría un vocabulario compartido o meta, estructura compartida a la que todos sus miembros pueden acceder (Ávila Muñoz, 2024). De este modo, a partir de la centralidad léxica se busca reconstruir el prototipo comunitario que tienen las distintas comunidades acerca de determinado tópico de la realidad, a la vez que se identifican los términos que son más o menos representativos de dicha estructura.

Desde un punto de vista cognitivo, la centralidad léxica implica una reflexión en torno al concepto de accesibilidad, pues el ejercicio asociativo aplicado para captar el léxico disponible es un ejercicio similar a las pruebas de fluencia semántica que provienen de la psicología (Hernández Muñoz, 2006), a través de las cuales los informantes acceden a una determinada categoría por medio de una unidad léxica principal. Por lo tanto, identificar aquellos vocablos que sirven de puntos de acceso a la red léxica permite reconstruir la estructura nuclear de accesibilidad del grupo y comparar los grados en que las estructuras individuales se asocian con la estructura grupal. Ahora bien, las unidades se organizan de acuerdo con su nivel de centralidad bajo una perspectiva de núcleo-

periferia, por lo que según sus valores alcanzados podemos identificar aquellos conceptos que son más centrales —más cercanos al núcleo— y menos centrales —más alejados— hasta aquel parámetro que Ávila Muñoz (2023) denomina *línea de incertidumbre*, y que sucede en el momento en que los informantes sospechan que se han alejado significativamente del núcleo del centro de interés y ponen en marcha mecanismos de reentrada a la red léxica original. Así, en contraste con la noción tradicional de disponibilidad, la centralidad léxica permite visibilizar aquellos vocablos que se alejan de las primeras posiciones, reivindicando el grado de asociación que tienen dichos términos con el prototipo comunitario independientemente de la posición que hayan obtenidos en los listados.

Por su parte, el índice de centralidad se fundamenta en la inclusión del concepto matemático de conjunto difuso, a partir del cual se evalúa la pertenencia de un elemento a un conjunto sin las restricciones propias de un grupo con límites bien definidos mediante una distinción de compatibilidad gradual. La incorporación de un modelo matemático en el estudio del léxico disponible tiene como finalidad superar las dificultades metodológicas propias de la selección de los vocablos que se consideran importantes para ser analizados en un determinado centro de interés. De esta forma, gracias a la herramienta *Fuzzy Expected Value* (*FEV*), es posible calcular eficazmente el índice de compatibilidad de un vocablo con respecto al espectro grupal en un centro de interés y determinar niveles objetivos de pertenencia de las unidades a dicho prototipo a partir de una constante k, cuya finalidad es definir el nivel de corte sobre el que deben ubicarse las palabras para ser consideradas más o menos representativas del léxico de la población (Ávila Muñoz, 2014; Ávila Muñoz & Sánchez Sáez, 2011, 2014).

En aras de sistematizar el cálculo de los índices de centralidad o compatibilidad léxica, Ávila Muñoz et al. (2021) desarrollan el programa computacional gratuito DispoCen, cuyo diseño se llevó a cabo mediante lenguaje de programación R. Gracias a esta herramienta, resulta posible calcular tanto el índice tradicional de disponibilidad —según la fórmula de López Chávez y Strassburger Frías (2000)— como también el índice de centralidad —según la fórmula Ávila Muñoz y Sánchez Sáez (2014)—.[2]

La inclusión de este nuevo paradigma en los estudios de disponibilidad léxica permite no solo la formulación de propuestas novedosas, sino también volver a analizar resultados de investigaciones previas a la luz de este nuevo enfoque complementario. En esta línea, durante los últimos años se han desarrollado diversos trabajos que adoptan el enfoque de la centralidad léxica para aproximarse, por un lado, a nuevos fenómenos como las percepciones sobre eventos o procesos socialmente relevantes como la pandemia (Ávila Muñoz, 2022; Ávila Muñoz et al., 2020; Delgado-Olmos et al., 2023), la inmigración (Extremera Pérez, 2024; Extremera Pérez & Galloso Camacho, 2023), el movimiento feminista (Ávila Muñoz & Segura Lores, 2024; Martínez Aránguiz & Ávila Muñoz, 2025) o el cambio climático (Ávila Muñoz & Segura Lores, 2024; Extremera Pérez, 2024), entre otros. Por otro lado, destaca la aplicación de este enfoque en

[2] Además, a diferencia de las demás herramientas que sirven para estos fines, DispoCen destaca por su autonomía y capacidad de integración de la información, pues al estar alojado en R, permite beneficiarse de los diferentes paquetes y librerías existentes tanto para análisis estadístico como para la representación gráfica de los resultados.

propuestas relacionadas con la selección de léxico disponible y enseñanza de lenguas (Ávila Muñoz, 2016, 2017; Santos Díaz, 2017; Santos Díaz et al., 2020), y el análisis del léxico disponible de futuros docentes en formación (Cruz Ventura, 2024).

De acuerdo con el estado actual de los estudios de centralidad léxica, no se ha realizado aún ninguna contribución desde una perspectiva interdialectal, por lo que este trabajo supone una oportunidad para evaluar el aporte de este enfoque novedoso en el cotejo de las diferencias en el léxico disponible de dos comunidades de habla distintas.

3. Metodología

El presente estudio se basa en una muestra de 360 estudiantes, la mitad de los cuales se encuentra en la provincia de Lleida, mientras que la otra mitad se ubica en la ciudad de Santiago de Chile. Todos cursan el último año de estudios antes de entrar en la educación superior, lo que corresponde con el 2.º de bachillerato en España y el 4.º de educación media en Chile y tienen, mayoritariamente, 17 o 18 años. La muestra se halla estratificada por cuotas uniformes según el sexo, tal como se ve en la Tabla 1.

Tabla 1. Distribución de los informantes en Lleida y Santiago según el sexo

Lleida (España)		Santiago (Chile)	
Mujeres	Varones	Mujeres	Varones
90	90	90	90
180		180	

El alumnado de Lleida, provincia española ubicada en el nordeste de la Península Ibérica, procede de un total de doce centros educativos, seis ubicados en la ciudad de Lleida —tres públicos y tres concertados— y los otros seis, en distintos puntos. Es relevante indicar que las personas encuestadas en Lleida son bilingües, dada la situación sociolingüística de Cataluña, donde el castellano y el catalán conviven de forma histórica y sin que haya, en la actualidad, una distribución rígida de los ámbitos de uso de cada lengua. La lengua vehicular de escolarización es el catalán, lengua inicial de la mayoría de la población en Lleida. Con respecto a la muestra chilena, estos estudiantes proceden de seis escuelas, correspondientes a establecimientos educativos públicos, particulares-subvencionados[3] y privados ubicados en cinco comunas de la ciudad de Santiago (Santiago centro, La Granja, Puente Alto, Ñuñoa y Las Condes), capital y núcleo urbano más poblado del país.

A todas las personas encuestadas, se les administró durante el año 2024 una prueba de fluidez categorial, el test de disponibilidad, donde debían escribir todas las palabras que se les ocurrieran relacionadas con un estímulo en un tiempo máximo de dos minutos.

[3] Este tipo de escuelas equivalen a los establecimientos concertados del sistema educativo español. La distribución de los informantes de cada tipo de escuela fue homogénea.

Los estudiantes de Lleida completaron un total de 17 categorías[4] y los chilenos, 11[5]. En la provincia catalana, las pruebas se administraron de forma escrita, a razón de una página por centro; en Santiago, la encuesta se completó en línea, gracias a la aplicación DLEx (De la Maya & López Pérez, 2022).

En el presente trabajo, analizamos dos centros de interés: *Alimentos y bebidas* y *Juegos y distracciones*, que recibieron los nombres de *Comidas y bebidas*, y *Juegos y actividades recreativas* en la prueba administrada en Santiago de Chile. Aunque los enunciados difieren ligeramente, como también sucede en otros estudios comparativos, hay una gran homogeneidad en otros aspectos metodológicos que aseguran la validez del contraste, como el número de informantes, la edad y el año de aplicación de las encuestas, aspectos señalados a veces como inconvenientes en los cotejos dialectales[6]. La uniformidad alcanza a los criterios de edición, en los que se han adoptado las mismas decisiones para evitar distorsiones en los resultados. En general, se siguen los criterios clásicos de Samper Padilla (1998), con la única salvedad de las marcas comerciales, que se mantienen en todos los casos.

Para el procesamiento de los datos empleamos, fundamentalmente, la herramienta DispoCen, cuya interfaz permite acceder a la categorización conceptual colectiva de cada estímulo. Por un lado, la herramienta realiza el cálculo del índice de compatibilidad léxica y, por otro, establece niveles de corte objetivos según el grado de compatibilidad de las unidades con el prototipo comunitario construido para un centro de interés particular. Para completar el análisis nos ha sido de gran utilidad el programa LexPro (Hernández Muñoz et al., 2023), con el que se obtienen los índices cuantitativos y los porcentajes de vocabulario compartido. Como resulta lógico, las posibilidades de las herramientas informáticas utilizadas en el análisis de la disponibilidad influyen en el desarrollo de los cotejos dialectales. Por este motivo, nuestra aportación pretende ser una aplicación de las nuevas posibilidades que confieren los programas actuales de análisis de la disponibilidad a este campo de estudio. Finalmente, hemos utilizado los paquetes estadísticos R para llevar a cabo los análisis de las diferencias entre los promedios de respuestas.

4. Análisis y discusión

Dividimos el presente apartado en dos subapartados, en función de la perspectiva adoptada en el análisis, ya sea la cuantitativa (4.1.) o la cualitativa (4.2). Con la finalidad de interpretar adecuadamente los datos, vamos a caracterizar brevemente los centros de interés considerados. Según su caracterización semántica, se trata de centros inclusivos, cuyos elementos se vinculan por una relación estricta de pertenencia (Hernández Muñoz,

[4] *El cuerpo humano; Ropa y complementos; Partes de la casa (sin los muebles); Muebles de la casa; Comidas y bebidas; Objetos colocados en la mesa para la comida; La cocina y sus utensilios; La escuela: muebles y materiales; Iluminación*, *Calefacción y medios para airear un recinto; La ciudad; El campo; Medios de transporte; Trabajos del campo y del jardín; Animales; Juegos y distracciones; Profesiones y oficios*, y *Colores*.

[5] *Deportes*, *Ansiedad*, *Juegos y actividades recreativas*, *Depresión*, *La escuela*, *Psicólogo*, *Alimentos y bebidas*, *Psiquiatra*, *Redes sociales*, *Trastorno Obsesivo compulsivo*, y *Salud mental*.

[6] Bartol Hernández (2008, p. 225) indica que las diferencias en el número de informantes pueden influir en las disimilitudes halladas. Samper Hernández (2008, pp. 270–271) llama la atención acertadamente sobre el problema de comparar muestras alejadas en el tiempo.

2006, p. 150). En relación con su productividad, medida a través de la cantidad de respuestas, *Alimentos y bebidas* es uno de los centros más productivos de los 16 clásicos; mientras que *Juegos y distracciones* no se caracteriza por un promedio alto de respuestas[7]. Sin embargo, es uno de los campos con mayor número de vocablos, por encima de *Alimentos*, con una posición media en este parámetro. Respecto al índice de cohesión, las coincidencias en *Alimentos* suelen ser varias, con lo que se configura como un centro bastante cohesionado (con una media de 0,03–0,05); en cambio, *Juegos* es claramente más disperso (con índices de 0,01–0,03) (López González, 2014, pp. 58–82). El hecho de que la cohesión no sea muy alta en ninguno de los dos campos puede influir en la cantidad de diferencias que encontremos[8].

4.1. Análisis cuantitativo

En primer lugar, vamos a comparar los índices cuantitativos obtenidos en las dos sintopías. Tal como se observa en la Tabla 2, el promedio de respuestas (o palabras), parámetro idóneo en las comparaciones inter e intramuestrales, es superior en *Alimentos* en los dos territorios frente a *Juegos*, donde se recoge un mayor número de vocablos (o palabras diferentes) en consonancia con la mayor dispersión de este centro, con un índice de cohesión menor.

Tabla 2. Índices cuantitativos en Lleida y Santiago: *Alimentos y bebidas* y *Juegos y distracciones*

	Alimentos y bebidas		***Juegos y distracciones***	
	Lleida	**Santiago**	**Lleida**	**Santiago**
Total de palabras	4432	2611	2787	1883
Vocablos	474	484	494	587
Promedio palabras	24,62	14,51	15,48	10,46
Índice de cohesión	0,05	0,03	0,03	0,02

Con respecto a la comparación entre las dos comunidades, en el total y el promedio de respuestas, los resultados son mejores en la provincia catalana donde se obtienen 24,62 palabras de promedio en *Alimentos*, mientras que en Santiago solo se obtienen 14,51. En *Juegos* se repite este patrón y también el resultado es más favorable en Lleida, con 15,48 frente a 10,46, aunque la diferencia en este caso es 5,02 respuestas y en la anterior de 10,1. Resulta llamativa esta disimilitud que, sin embargo, no se observa en la cantidad de vocablos, con cifras muy similares alrededor de *Alimentos* (474 vs. 484) y superiores en Chile en *Juegos* (494 vs. 587). Este hecho provoca que los índices de cohesión sean

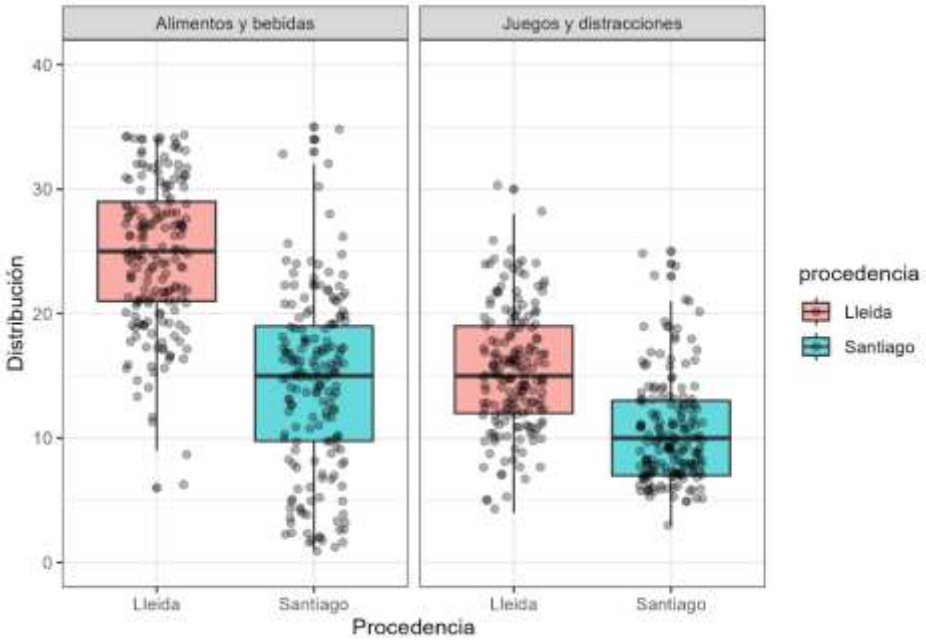
[7] López González (2014, p. 60), al revisar los datos de 14 investigaciones realizadas sobre disponibilidad léxica, sitúa *Alimentos y bebidas* en la segunda posición y *Juegos y distracciones* en la décima según el promedio de palabras. Abreviamos los nombres de los centros de interés a partir de ahora para facilitar la lectura.

[8] Samper Padilla y Samper Hernández (2007, p. 107) afirman: "Los resultados de los cotejos realizados hasta ahora permiten concluir que hay una relación inversamente proporcional entre el grado de cohesión de los centros de interés y el número de diferencias cualitativas". También Bartol Hernández (2008, p. 226): "cuanto mayor cohesión tiene un centro de interés, menores son las diferencias entre las diferentes regiones en el léxico disponible de informantes pertenecientes a un mismo grupo generacional".

inferiores en Chile, puesto que se reportan más vocablos, siendo más baja la cantidad de respuestas. Los promedios recabados en Lleida son cercanos a la media de estos centros de interés, situada en 25,7 en el primer campo y en 17,8 en el segundo (López González, 2014, p. 58); los datos de Santiago son bajos en relación con los que obtienen informantes de la misma edad en otros lugares hispánicos. Esta preponderancia de los resultados peninsulares, también la observa Valencia Espinoza (2011) al comparar el promedio de Chile con el de Aragón en estudiantes de la misma edad y Gómez Devís et al. (2023) al cotejar el léxico disponible de niños de Talca y Valencia.

Con el fin de verificar la significancia de las marcadas diferencias registradas en la producción léxica en ambos estímulos, llevamos a cabo procedimientos estadísticos que permiten ahondar en la distribución de las respuestas de cada grupo estudiado (Figura 1). En términos generales, la distribución de los datos de los estudiantes de Lleida tiende a estar por encima de la de los alumnos de Santiago en los dos campos léxicos analizados. Además, según el rango de las respuestas de los grupos, se aprecia que los estudiantes catalanes poseen una distribución más homogénea que los santiaguinos en el estímulo *Alimentos* (Lleida: 28; Santiago: 43); no obstante, en *Juegos* se produce el efecto contrario, puesto que el rango de los alumnos españoles es levemente superior que el de los chilenos (26 vs. 22). Finalmente, al considerar los valores máximos de cada grupo, el estudiantado americano registra el mayor rendimiento en el estímulo *Alimentos* y el leridano, en *Juegos*.

Figura 1. Distribución de las respuestas en cada centro de interés en Lleida y Santiago



En cuanto al análisis específico de contraste estadístico, por un lado, tras analizar la naturaleza de los datos de cada grupo mediante la prueba de *Shapiro*, se corrobora que la producción de Lleida posee una distribución normal en *Juegos* (W = 0,98, p-value = 0,16), pero no en *Alimentos* (W = 0,97, p-value = 0,00); mientras que, en Santiago,

ninguno de los dos centros de interés posee una distribución normal (*Juegos*: W = 0,9653, p-value = 0,00; *Alimentos*: W = 0,91, p-value = 8,63e-09). Por otro lado, a través de la prueba de *Levene*, se comprueba que las varianzas de ambos grupos no son homogéneas ni en el centro *Alimentos* (Df= 1, F = 4,64, *p-value* = 0,03) ni en *Juegos* (Df= 1, F = 5,13, *p-value* = 0,02).

Por lo tanto, luego de comprobar que los datos no muestran una distribución normal ni cumplen con el requisito de homocedasticidad, se emplea la prueba no paramétrica *U de Mann Whitney*, cuyas funciones permiten contrastar la distribución de los datos de dos grupos diferentes bajo tres escenarios distintos: (1) la distribución de los grupos es diferente; (2) la distribución del Grupo 1 tiende a presentar valores superiores a la del Grupo 2; (3) la distribución del Grupo 1 tiende a presentar valores inferiores a la del Grupo 2. Así, según se constata en la Tabla 3, los resultados de la prueba realizada indican que para ambos centros de interés existe suficiente evidencia estadística para concluir que la distribución de las respuestas de los grupos es diferente y que los estudiantes de Lleida tienden a presentar un rendimiento superior al de los alumnos santiaguinos.

Tabla 3. Índices cuantitativos en Lleida y Santiago en cada centro de interés.

CI	Grupo 1	Grupo 2	Escenario	*w-value*	*p-value*
Alimentos	Lleida	Santiago	(1)	28092,5	1,8e-27
			(2)	28092,5	9,2e-28
			(3)	28092,5	1
Juegos			(1)	25757,5	3,0e-16
			(2)	25757,5	1,5e-16
			(3)	25757,5	1

Como ya advertimos hubo una diferencia metodológica importante en la aplicación de la encuesta: las de Lleida se completaron en papel y las santiaguinas, en línea. Esta podría ser una explicación a la distancia observada entre las comunidades en los promedios; no obstante, los estudios realizados hasta la fecha que han comparado estas dos opciones metodológicas no han hallado diferencias considerables (De la Maya Retamar & López Pérez, 2025; Martínez Lara, 2023; Martínez Lara & Santos Díaz (e.p.)).

Pasamos a comparar ahora los repertorios obtenidos (Tabla 4), en relación con *Alimentos*. El léxico compartido alcanza un porcentaje bajo, del 35,02 % cuando se compara Lleida con Santiago y del 34,29 % en dirección inversa. Las formas coincidentes en los dos periodos son 166 (de los 474 y 484 que se obtienen en cada sintopía, v. Tabla 2), dato que implica una compatibilidad general del 20,95 % sobre la suma de los conjuntos (792 vocablos)[9]. En *Juegos*, los datos son similares. El vocabulario compartido es aún menor, con índices también semejantes en los dos territorios y siempre inferiores en Santiago, del 34,41 % y del 29,01 %; los vocablos compartidos son 170, cifra que representa una coincidencia del 18,68 % sobre el total de las unidades (911).

[9] Conviene ser prudentes respecto a esta cifra, puesto que cada corpus tiene variación intrínseca y no podemos saber qué variación se debe al cotejo diatópico o a la variación inherente en cada uno de los corpus.

Tabla 4. Datos cuantitativos de los vocablos compartidos y exclusivos en el corpus completo en las dos sintopías: *Alimentos y bebidas* y *Juegos y distracciones*

Alimentos y bebidas					
Sintopía	**Vocablos exclusivos**	**% exclusivos**	**% compartidos**	**Vocablos compartidos**	**% compatibilidad**
Lleida	308	64,98	35,02	166	20,95
Santiago	318	65,70	34,29		
Juegos y distracciones					
Lleida	324	65,58	34,41	170	18,68
Santiago	417	70,98	29,01		

Estos resultados recuerdan los índices de compatibilidad bajísimos que encontró López Chávez (1992) al considerar la nómina completa de palabras. Sin embargo, si tenemos en cuenta las 100 primeras palabras según el índice de disponibilidad, hallamos resultados distintos, en la línea de los estudios que nos anteceden sobre cotejos dialectales. En *Alimentos*, el porcentaje de compatibilidad aumenta hasta el 43 %, mientras que en *Juegos* es del 41 %. Comparando estos datos con los de Valencia Espinoza (2011), quien contempló en su contraste entre Chile y Aragón esta cantidad de palabras, observamos una gran similitud en las cifras: la investigadora chilena halló una compatibilidad del 45 % en el primer centro de interés y del 41 %, precisamente, en el segundo. Al extender el cotejo más allá de las 100 primeras posiciones, observamos que el índice de convergencia aumenta, puesto que, de los 57 vocablos exclusivos en Lleida en *Alimentos* entre los 100 primeros, 35 se encuentran en Chile, solo que más allá de esa posición. De forma inversa, 19 de los 57 vocablos exclusivos de Santiago se encuentran en las listas leridanas. Estos datos hacen que la compatibilidad entre los inventarios aumente al 78 % (Lleida→ Santiago) y al 62 % (Santiago→ Lleida). En *Juegos*, de las 59 palabras exclusivas de Lleida, 27 las mencionan los estudiantes de Santiago en otras posiciones; a la inversa sucede con 19 vocablos. En este centro de interés, por tanto, la compatibilidad considerando estas palabras, es del 68 % (Lleida→ Santiago) y del 60 % (Santiago→ Lleida).

Por otro lado, si ceñimos el análisis comparativo a las 50 primeras palabras, el 50 % son idénticas en los dos territorios en *Alimentos*. Dentro de este margen solo hay 6 unidades que pertenecen a la sintopía catalana ausentes por completo en la nómina santiaguina (*patatas*, *macarrón*, *zumo*, *espagueti*, *aquarius* y *Nestea*) y 11 en sentido inverso (*fideos*, *papas fritas*, *completo*, *mc donald's*, *papas*, *asado*, *cazuela*, *sopaipilla*, *KFC*, *comida chatarra y comida china*), lo que indica un porcentaje de compatibilidad global dentro de este margen del 88 % (Lleida→ Santiago) y de 78 % en sentido contrario (Santiago→ Lleida). En *Juegos*, hay 24 palabras compartidas, lo que supone una compatibilidad algo menor, del 48 %. Ahora bien, las palabras que realmente no están en

el corpus completo del otro territorio son solo 14[10], lo que indica un porcentaje del léxico compartido del 72 % en las dos direcciones.

4.2. Análisis cualitativo

A través de DispoCen, se ha diseñado un modelo específico que consta de seis niveles de centralidad definidos para cada centro de interés, entre los cuales se concentra el léxico nuclear aportado por cada subgrupo de hablantes de la muestra en las categorías semánticas propuestas. De este modo, se registran los vocablos correspondientes a cada nivel de centralidad, en donde el Nivel 6 es el máximo nivel y el Nivel 1 el mínimo o periférico. La pertenencia de los vocablos a cada nivel se determina mediante valores de corte, definidos por la herramienta informática, que las unidades deben obtener para ubicarse en un determinado nivel. Además, mediante este enfoque es posible obtener la representación de la red mental de los informantes, a la vez que se accede a la estructura léxica de cada grupo en los dos estímulos propuestos, permitiendo identificar aquellos vocablos que funcionan como puntos de acceso a las redes léxicas comunitarias (Nivel 6).

4.2.1. Alimentos y bebidas

De acuerdo con las respuestas registradas en el estímulo *Alimentos* (Ver Tabla 5), tanto en los niveles de máxima centralidad (Niveles 6 y 5) como en los intermedios (Nivel 4 y 3) se aprecian diferencias significativas en la calidad de los alimentos descritos por los hablantes de cada grupo. En los vocablos de los informantes de Lleida se observa una mayor proporción de alimentos con un elevado índice nutricional, entre los que destacan la presencia de bebestibles como el *agua*; algunas frutas como la *manzana*, el *plátano* y el *tomate*; carbohidratos como la *pasta* y el *macarrón*; junto con proteínas como la *carne*. En cambio, para los hablantes de Santiago se observa un elevado número de vocablos relacionados con comida rápida y alimentos de escaso valor nutricional, a saber, bebidas azucaradas como la *(coca) cola*; alimentos ultraprocesados o poco saludables como la *hamburguesa*, la *pizza*, el *completo,* las *papas fritas* o el *sushi*[11]; y, de forma ocasional, se presentan conceptos como el *agua* y la *(carne de) pollo*, los cuales podrían ser considerados como alimentos que forman parte de una alimentación balanceada.

[10] Exclusivas en Lleida: *teléfono móvil*, *oca*, *pilla pilla*, *escondite*, *tres en raya*, *ordenador*, *videoconsola*, *cluedo*, *baloncesto*, *pica pared*, *Catán*, *billar*, *comba*, *Clash of Clans*. Exclusivas en Santiago: *pinta*, *salir*, *diversión*, *ludo*, *escondida*, *saltar la cuerda*, *carioca*, *felicidad*, *amistad*, *risa*, *entretenimiento*, *sillita musical*, *quemadas*, *Counter Strike.*

[11] Los hablantes de Santiago consideran el *sushi* como un alimento poco saludable, pues su preparación más habitual y consumida implica un procedimiento de rebozado y fritura en aceite, lo cual afecta significativamente su contenido nutricional.

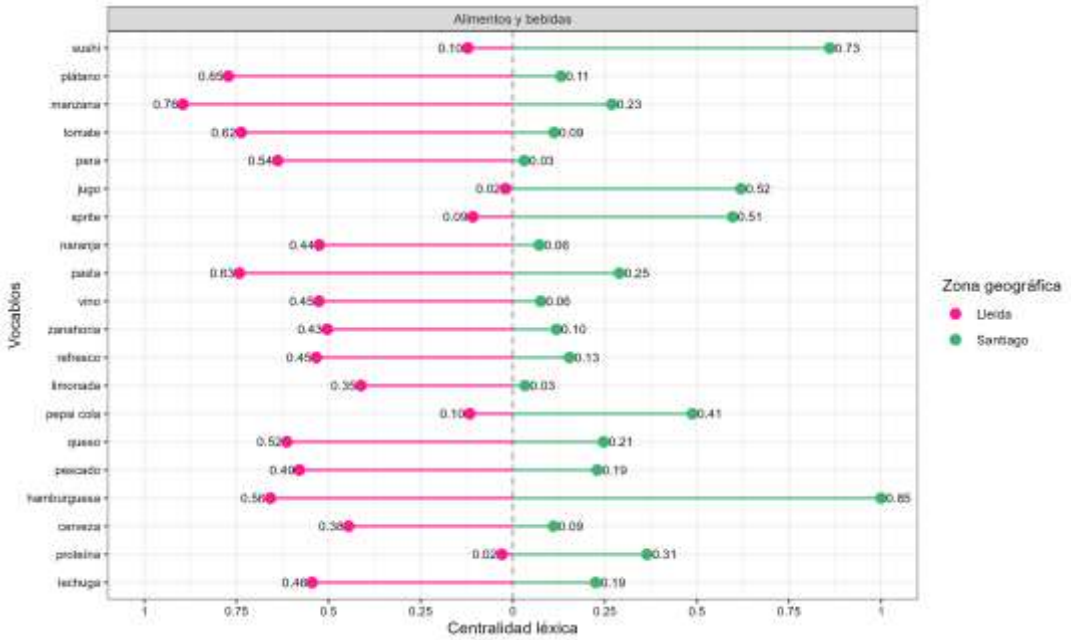
Tabla 5. Elementos más compatibles en *Alimentos y bebidas*

Nivel	Lleida	ICL	Santiago	ICL
6	agua	0,90	(coca) cola	0,90
5	(coca) cola	0,84	hamburguesa	0,84
4	manzana	0,72	pizza, sushi	0,72
3	macarrón, plátano, carne, pasta, tomate	0,60	agua, completo, (carne de) pollo, papas fritas	0,59
2	patata, (carne de) pollo, pan, hamburguesa, arroz, pera, pizza, verdura, queso, zumo, pescado, aquarius, fruta, lechuga	0,46	fideos, arroz, mcdonald's, carne, jugo, fanta, sprite, papas	0,42
1	espagueti, refresco, vino, naranja, fanta, zanahoria, sopa, café, leche, nestea, cerveza, jamón, chocolate, limonada, sandía, ensalada, galleta, huevo, yogur	0,29	pepsi cola, ensalada, fruta, pan, puré, proteína, verdura, alcohol, galleta, sopaipilla, (comida) chatarra, lasaña, bebida, empanada	0,24

Luego, al observar los niveles periféricos de centralidad (Niveles 2 y 1), se corroboran las diferencias registradas en los rangos anteriores, puesto que los estudiantes de Lleida mantienen la tendencia hacia la inclusión de alimentos con elevado valor nutricional (por ejemplo: *pera*, *fruta*, *lechuga*, *leche*, *café*, *sandía*, *huevo*, *yogur*, *ensalada*, *zanahoria*), incluyendo ocasionalmente algunas bebidas alcohólicas (*vino*, *cerveza*), bebidas azucaradas (*aquarius*, *fanta*) y comidas ultraprocesadas (*galleta*, *hamburguesa*, *pizza*, *chocolate*). Por su parte, en los estudiantes santiaguinos se observa lo contrario, pues en este grupo predominan las comidas de bajo nivel nutricional como los alimentos fritos y las bebidas azucaradas (por ejemplo: *mcdonald's*, *sopaipilla*, *sprite*, *fanta*, *pepsi cola*, *(comida) chatarra*, *bebida*, *empanada*) y, en algunos casos, se registran alimentos nutritivos, los cuales aparecen hacia el final de las redes léxicas de este grupo (*arroz*, *papas*, *ensalada*, *fruta*, *pan*, *verdura*).

Con el propósito de corroborar las tendencias registradas a través de los seis niveles de centralidad, a continuación, en la Figura 2, se analizan los veinte vocablos que obtuvieron una mayor diferencia según el índice de centralidad léxica asignado por cada grupo. En este sentido, a través de la diferencia de sus valores en ambos grupos, se constata que para los informantes de Lleida son mucho más centrales elementos como el *plátano* (ICL = 0,65), la *manzana* (ICL = 0,76), la *naranja* (ICL = 0,44), la *pasta* (ICL = 0,62), el *tomate* (ICL = 0,62), la *pera* (ICL = 0,54) y la *lechuga* (ICL = 0,46); mientras que los estudiantes de Santiago tienen cierta preferencia por alimentos ultraprocesados, como el *sushi* (ICL = 0,73), el *sprite* (ICL = 0,51), la *pepsi cola* (ICL = 0,41) y la *hamburguesa* (ICL = 0,85).

Figura 2. Contraste según índice de centralidad léxica en *Alimentos y bebidas*



Sin embargo, en conformidad con los resultados presentados, existen algunas unidades que permiten matizar la tendencia señalada más arriba. Por ejemplo, en los estudiantes chilenos aparecen conceptos que estarían asociados con un mayor valor nutricional, a saber, el *jugo* (ICL = 0,52), con escasa presencia en Lleida por el predominio del geosinónimo *zumo*, y la *proteína* (ICL = 0,31) —de hecho, este último término, en particular, podría reflejar un gran nivel de conciencia nutricional, en el caso de que esté refiriéndose a suplementos alimenticios como la proteína en polvo, la cual está fuertemente asociada con una cultura de actividad física—. Del mismo modo, en los estudiantes de Lleida, a diferencia de los santiaguinos, se observan valores de centralidad significativos en unidades relacionadas con bebidas alcohólicas como el *vino* (ICL = 0,45) y la *cerveza* (ICL = 0,38), cuyo consumo excesivo suele estar asociado con el riesgo de padecer obesidad, desnutrición severa y desequilibrios de glucosa, entre otros efectos en la salud de las personas.

4.2.2. Juegos y distracciones

Con respecto al centro de interés *Juegos* (Tabla 6), en los niveles de máxima centralidad (Niveles 6, 5) y en los intermedios (Niveles 4 y 3) se observan algunas diferencias entre los grupos dadas por las características de las actividades mencionadas por los estudiantes. Por un lado, en los estudiantes de Lleida destacan las actividades relacionadas con dispositivos digitales (*teléfono móvil*, *playstation)*, juegos de mesa (*parchís*, *cartas*, *uno*, *juego de mesa*) y con videojuegos (*videojuego*, *Fortnite*, FIFA). Por otro lado, en los alumnos santiaguinos predominan las ocupaciones al aire libre (*fútbol*, *pinta*, *parque*, *dibujar*, *salir*), los juegos de mesa (*ludo*, *uno*), a la vez que se

presentan unidades relacionadas con un ámbito emocional (*diversión*) y social de la recreación (*amigo*).

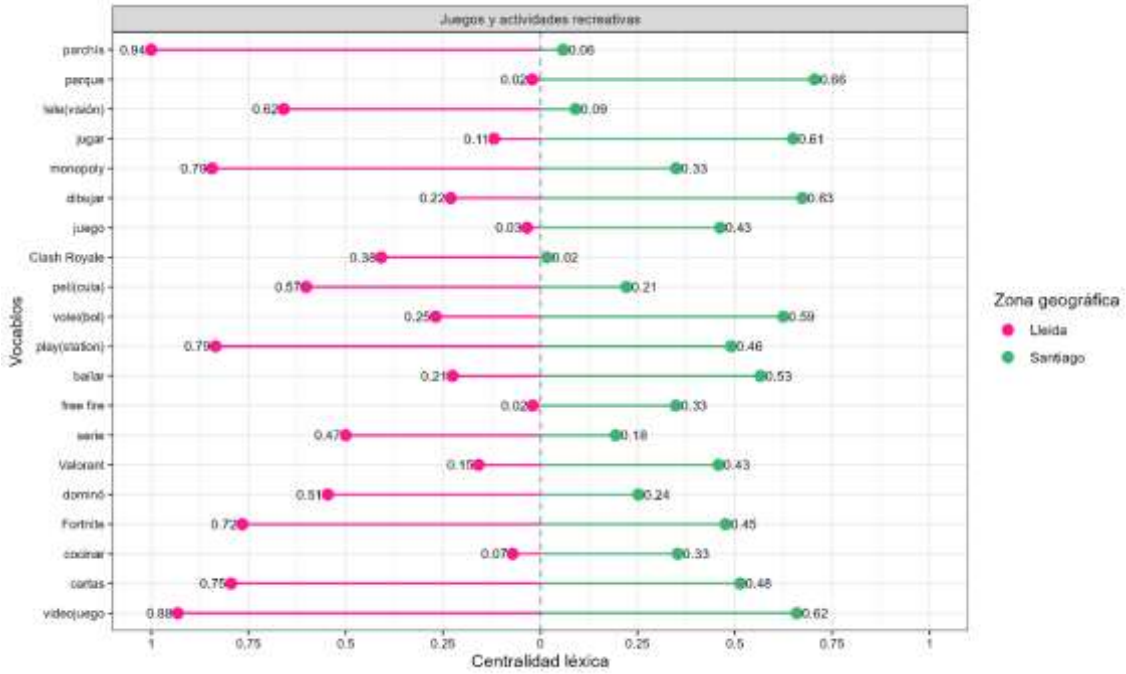
Tabla 6. Elementos más compatibles en *Juegos y distracciones*

Nivel	Lleida	Nivel de corte	Santiago	Nivel de corte
6	parchís	0,90	fútbol	0,90
5	videojuego, fútbol	0,85	diversión	0,82
4	(teléfono) (móvil), monopoly, play(station)	0,75	pinta, amigos	0,73
3	cartas, uno, Fortnite, FIFA, juego de mesa	0,62	salir, ludo, parque, uno, dibujar, videojuego, jugar	0,60
2	tele(visión), oca, pilla pilla, ajedrez, escondite, pelí(cula), básquet(bol), (vídeo)consola, dominó, tres en raya, leer, serie	0,46	básquet(bol), volei(bol), escondida, saltar la cuerda, deporte, bailar, ajedrez, correr, cartas, juego de mesa, leer, Minecraft, FIFA, play(station)	0,45
1	ordenador, Brawl Stars, Clash Royale, póquer, nintendo (switch), Minecraft, deporte, GTA, cluedo, música, Catán, correr, TikTok, ruleta	0,27	Fortnite, música, entretenimiento, juego, Valorant, felicidad, League of legends, Brawl Stars, carioca, pintar, risa, pimpón, Call of Duty, cocinar, cantar, familia, monopoly, Free fire, divertido, sillita musical, cine, entretención	0,28

Luego, hacia los niveles periféricos de centralidad (Nivel 2 y 1), la tendencia descrita se mantiene, aunque con algunos matices, pues en los estudiantes de Lleida aparecen nuevamente numerosos conceptos relacionados con videojuegos (*Brawl Stars*, *Clash Royale*, *GTA*, *Minecraft*), a los que se suman nuevos dispositivos digitales (*tele(visión)*, *ordenador*, *nintendo (switch)*, *(video)consola*); no obstante, en los listados de este grupo comienzan a vislumbrarse también otras actividades sin pantallas como los juegos de mesa (*ajedrez*, *tres en raya*, *dominó*, *cluedo*, *póquer*, *Catán*) y formas de ocio al aire libre (*escondite*, *pilla pilla*, *básquet(bol)*, *correr*), junto con otras actividades recreativas individuales (*música*, *leer*). Por su parte, pese a que los estudiantes de Santiago muestran preferencia por pasatiempos analógicos como los juegos en el exterior (*básquet(bol)*, *volei(bol)*, *escondida*, *saltar la cuerda*, *correr*, *deporte*, *bailar*, *sillita musical*) y otras ocupaciones como los juegos de mesa (*ajedrez*, *cartas*, *juego de mesa*, *carioca*) y entretenciones de carácter individual (*leer*, *cine*, *pintar, música, cantar*); al igual que en los alumnos de Lleida, resulta numerosa la inclusión de videojuegos (*Minecraft*, *FIFA*, *League of legends*, *Brawl Stars*, *Call of Duty*, *Free fire*), lo cual constituye una importante similitud en los hábitos de ambos grupos de estudiantes.

Tal como se llevó a cabo en el estímulo anterior, nuevamente se opta por el análisis de los veinte vocablos en los que se registró una mayor diferencia entre los índices de centralidad léxica estipulados a través de las respuestas de los jóvenes de ambos grupos. Así, según se aprecia en la Figura 3, por una parte, en los estudiantes de Lleida destacan elementos como *parchís* (ICL = 0,94), *tele(visión)* (ICL = 0,62), *monopoly* (ICL = 0,79), *play(station)* (ICL = 0,79), *Fortnite* (ICL = 0,72), *cartas* (ICL = 0,75) y *videojuego* (ICL = 0,88). Por otra parte, en el grupo de Santiago alcanzan un elevado índice de centralidad vocablos como *parque* (ICL = 0,66), *jugar* (ICL = 0,61), *dibujar* (ICL = 0,63), *bailar* (ICL = 0,53) y *cocinar* (ICL = 0,33).

Figura 3. Contraste según índice de centralidad léxica en *Juegos y distracciones*



A partir de estos resultados, es posible colegir que, pese a las diferencias contextuales entre Lleida y Santiago, ambos grupos de estudiantes comparten un interés significativo por los videojuegos, los cuales se mantienen como un núcleo común dentro de sus repertorios léxicos. Sin embargo, las divergencias se hacen evidentes en los niveles periféricos de centralidad: mientras en Lleida se privilegia la mención de dispositivos digitales y juegos de mesa específicos, en Santiago se observa una mayor inclinación hacia actividades recreativas y creativas variadas, como el dibujo, el baile o la cocina. Estas variaciones reflejan no solo la diversidad de referentes culturales y de ocio en cada contexto, sino también la coexistencia de prácticas digitales y tradicionales en las formas de entretenimiento juvenil.

5. Conclusiones

Con respecto al análisis cuantitativo, hemos constatado la superioridad en los promedios de los estudiantes de Lleida con diferencias considerables (de 10,1 respuestas en *Alimentos*; y de 5,02 en *Juegos*) que tienen relevancia estadística. Parece difícil encontrar una explicación a este resultado, a no ser la posible relación con la aplicación de la encuesta, con diferencias en las dos comunidades, aunque esto sería contradictorio con los estudios realizados hasta el momento en los que no se observa diferencia en función del parámetro papel/en línea. A la espera de más investigaciones que analicen esta disimilitud metodológica, sería recomendable ampliar el estudio con más informantes para confirmar estos resultados. Sin embargo, cabe notar que en el número de vocablos se aprecia la tendencia contraria y los informantes chilenos tienen una mayor variedad léxica, con diferencias distintas en función del centro de interés (de 10 vocablos más en *Alimentos* y 93 en *Juegos*).

En cuanto al cotejo entre los dos repertorios léxicos, al considerar únicamente las las 100 palabras más disponibles, se detectan índices de compatibilidad bajos (el 41 y el 43 %), pero en consonancia con los estudios que comparan territorios españoles y americanos (Samper Padilla & Samper Hernández, 2007) y con los datos que ya obtuvo Valencia Espinoza (2011), al comparar Aragón y Chile hace veinte años. Cabe comentar algunas tendencias en este sentido. Por un lado, la convergencia es siempre mayor en *Alimentos* que en *Juegos*, independientemente del criterio que se contemple en el cotejo, lo que parece indicar que hay más diferencias diatópicas ligadas a los juegos y las formas de ocio que a la gastronomía. Por otro lado, la similitud del inventario léxico de Lleida con el de Santiago es mayor en esta dirección que en la inversa. Es decir, hay más palabras que los estudiantes de Lleida desconocen del español americano que al revés. Esta circunstancia puede explicarse por la mayor cercanía del español de esta comunidad con el estándar, basado tradicionalmente en la variedad septentrional castellana, a la que se adscribe el español hablado en Lleida.

El análisis cualitativo, basado en los niveles de centralidad, también ha arrojado algunos resultados interesantes. Con respecto a los hábitos alimenticios, se observa una mayor preferencia en la comunidad española por alimentos y bebidas saludables, propias de la dieta mediterránea (*agua*, *manzana*, *tomate*, *carne*, *fruta*…) frente al estudiantado santiaguino, que opta por una alimentación menos equilibrada, donde se reconocen comidas típicas chilenas (*completo*, *sopaipilla*, *empanada*…), pero también de otros lugares (*pizza*, *sushi*) que alcanzan niveles altos de centralidad. En relación con los juegos y las formas de ocio, se detecta una mayor presencia de entretenimientos relacionados con las pantallas en la sintopía catalana, donde *videojuego*, *teléfono móvil* y *playstation* ostentan posiciones privilegiadas en cuanto a centralidad. A pesar de que estas distracciones también se encuentran en las listas chilenas, en este alumnado predominan las formas de ocio al aire libre y en compañía (*amigos*, *deporte*, *parque*, *diversión*…). Tras haber realizado este análisis, consideramos que los avances de la centralidad léxica pueden contribuir eficazmente a los estudios interdialectales de léxico disponible, tanto para abrir nuevas vías de análisis a partir de las pruebas

asociativas, como también para proponer un marco objetivo para la selección de léxico en este tipo de investigaciones.

Hay que indicar que, tanto en el análisis cualitativo como en las palabras exclusivas de cada sintopía, se encuentran diferencias que responden a la presencia de geosinónimos o sinónimos diatópicos, en los que la misma realidad recibe distinto nombre (*patatas-papas*, *zumo-jugo*, *billar-pool*, *comba-cuerda*, *escondite-escondida*, *ordenador-computadora*, *pilla pilla-pinta*, *teléfono móvil-celular*, *tres en raya-gato*). Otras veces, se trata de elementos culturales distintos que no cuentan con un equivalente; se constatan tanto en las comidas (*macarrón*, *espagueti*[12] vs. *asado*, *cazuela*[13]) como en los juegos (*oca*, *tres en raya*, *pica pared* vs. *ludo*, *carioca*).

De cualquier modo, tiene un enorme valor conocer qué palabras comparten o no los hispanohablantes: ahondar en este conocimiento interdialectal de las variedades del español puede ser de gran utilidad para la adaptación de la enseñanza a contextos específicos, así como para lograr una planificación adecuada del vocabulario.

Contribución de los autores: Vicente Martínez Aránguiz y Maribel Serrano Zapata confirman haber contribuido de manera equitativa y relevante en el diseño, desarrollo y ejecución de la presente contribución académica. Ambos autores participaron activamente en las distintas etapas de la elaboración de la investigación, esto es, en el marco teórico, análisis y discusión de los resultados. Asimismo, los autores sostienen haber ejercido labores de redacción y revisión de acuerdo con los lineamientos dispuestos por la revista con el propósito de garantizar el cumplimiento de los estándares editoriales y el rigor científico acorde a esta instancia académica. Además, el orden de autoria fue definido de común acuerdo entre los autores.

Financiación: Esta investigación es parte del proyecto ECOS-C/N, Estudio de los condicionantes sociales del español actual en el centro y norte de España: nuevas identidades, nuevos retos, nuevas soluciones (ref. PID2023-148371NB-C42).

Disponibilidad de datos: Los datos analizados en el presente estudio consisten en los resultados de encuestas de disponibilidad léxica aplicadas y registradas por los propios autores. En función de garantizar el cumplimiento de los estándares éticos necesarios, los datos se encuentran bajo la custodia de los autores, resguardando la confidencialidad de la información personal los informantes.

Referencias

Alba, O. (1998). Variable léxica y dialectología hispánica. *La Torre. Revista de la Universidad de Puerto Rico*, *3*(7-8), 299–316. https://dialnet.unirioja.es/servlet/articulo?codigo=2416487

Alba, O. (2000). Variable léxica y comparación dialectal. In O. Alba, *Nuevos aspectos*

[12] Son llamativos los casos de *macarrón* y *espagueti*, en cuanto no contarían con un equivalente específico en Chile, pues en este país son más comunes las denominaciones “tallarines” y “pasta” para referirse a las diferentes variedades de este alimento con independencia de su forma o su receta.

[13] Tanto *Cazuela* y *asado* constituyen dos comidas con matices específicos en la cocina chilena, en tanto la primera corresponde a un guiso que contiene caldo, papas, carne (cerdo, vacuno o pollo) y verduras; mientras que la segunda corresponde a una preparación de carne asada en parrilla similar a la barbacoa.

del español en Santo Domingo (pp. 99–132). Librería La Trinitaria y Brigham Young University.

Ávila Muñoz, A. (2014). Patrones sociolingüísticos de la riqueza léxica. Estudio basado en una propuesta original para el cálculo del índice de la densidad léxica virtual de los hablantes. *Lingüística Española Actual*, *36*(2), 249–272. https://dialnet.unirioja.es/servlet/articulo?codigo=5222822

Ávila Muñoz, A. M. (2016). El léxico disponible y la enseñanza del español. Propuesta de selección léxica basada en la teoría de los conjuntos difusos. *Journal of Spanish Language Teaching*, *3*(1), 31–43. https://doi.org/10.1080/23247797.2016.1163038

Ávila Muñoz, A. (2017). La utilidad del vocabulario dialectal en el aula de lenguas extranjeras. Propuesta de selección léxica basada en coronas concéntricas. *marcoELE. Revista de Didáctica Español Lengua Extranjera*, (25), 1–21. https://marcoele.com/descargas/25/avila-vocabulario_dialectal.pdf

Ávila Muñoz, A. M. (2022). Algunas percepciones categoriales compartidas por preuniversitarios andaluces sobre la crisis del coronavirus y sus consecuencias. ¿Deberíamos preocuparnos? Un acercamiento desde la disponibilidad y la centralidad léxica. *Tejuelo*, *35*(3), 17–42. https://doi.org/10.17398/1988-8430.35.3.17

Ávila Muñoz, A. M. (2023). Initial proposal for the development of the hispanic lexico-statistical tradition: the lexical centrality index. *Nueva revista de filología hispánica*, *71*(1), 3–30. https://doi.org/10.24201/nrfh.v71i1.3838

Ávila-Muñoz, A. M. (2024). Spanish students' categorical perceptions of feminist movements. *Journal of Psycholinguistic Research*, *53*(4), 52. https://doi.org/10.1007/s10936-024-10091-8

Ávila Muñoz, A. M., & Sánchez Sáez, J. M.ª. (2010). La disponibilidad léxica. Antecedentes y fundamentos. En A. M. Ávila Muñoz & J. M. Sánchez Sáez (Eds.) *Variación social del léxico disponible en la ciudad de Málaga: Diccionario y análisis* (pp. 37–81). Editorial Sarriá.

Ávila Muñoz, A. M., & Sánchez Sáez, J. M.ª. (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. *ELUA: Estudios de Lingüística. Universidad de Alicante*, *25*, 45–74. https://doi.org/10.14198/ELUA2011.25.02

Ávila-Muñoz, A. M., & Sánchez Sáez, J. M.ª. (2014). Fuzzy sets and prototype theory: representational model of cognitive community structures based on lexical availability trials. *Review of Cognitive Linguistics*, *12*(1), 133–159.

Ávila Muñoz, A. M., & Segura Lores, A. (2024). *Los estudiantes andaluces ante los nuevos retos de una sociedad en transformación: percepciones, reflexiones y propuestas de actuación administrativa y educativa*. Tirant Humanidades.

Ávila Muñoz, A. M., Santos Díaz, I. C., & Trigo Ibáñez, E. (2020). A lexical-cognitive approach to the influence of the media on the perceptions of Spanish university students regarding COVID-19. *Círculo de Lingüística Aplicada a la Comunicación*, *84*, 85–95. https://doi.org/10.5209/clac.70560

Ávila Muñoz, A. M., Sánchez Sáez, J. M.ª, & Odishelidze, N. (2021). DispoCen. Mucho más que un programa para el cálculo de la disponibilidad léxica. *ELUA: Estudios*

de Lingüística. Universidad de Alicante, *35*, 9–36. https://doi.org/10.14198/ELUA2021.35.1

Ávila Muñoz, A. M., & Villena Ponsoda, J. A. (Eds.). (2010). *Variación social del léxico disponible en la ciudad de Málaga*. Sarriá.

Bartol Hernández, J. A. (2008). Variación léxica del español: los léxicos disponibles de Aragón y Soria. En M. L. Arnal Purroy (Ed.), *Estudios sobre disponibilidad léxica en los jóvenes aragoneses* (pp. 207–226). Institución "Fernando el Católico".

Bartol Hernández, J. A. (2020). Léxico disponible de España: la ropa. En F. Gimeno (Ed.), *Lengua*, *Sociedad*, *Cultura* (pp. 41–59). Publicaciones de la Universidad de Alicante.

Bartol Hernández, J. A. (2022). Léxico panhispánico: animales. En C. Díaz Alayón (Coord.), *Studia philologica in honorem José Antonio Samper* (pp. 97–134). Arco/Libros y Academia Canaria de la Lengua.

Bartol Hernández, J. A., & Borrego Nieto, J. (2019). El léxico disponible de la ropa en Cádiz, Valladolid y el conjunto de España: norma y variación. *Ogigia. Revista Electrónica De Estudios Hispánicos*, *25*, 57–82. https://doi.org/10.24197/ogigia.25.2019.57-82

Carcedo González, A. (2003). Unidad y variedad diatópica de la disponibilidad léxica del español: comparación de los inventarios de Puerto Rico, Cádiz y Asturias. En Moreno Fernández et al. (Coords.). *Lengua*, *variación y contexto. Estudios dedicados a Humberto López Morales* (Vol. 1, pp. 199–226). Arco Libros.

Carcedo González, A. (2004). La variable léxica disponible en la comparación interdialectal: compatibilidad de la norma asturiana con otras sintopías hispanohablantes. En V. M. Sánchez Corrales (Ed.), *Actas del XIII Congreso Internacional de la Asociación de Lingüística y Filología de la América Latina* (pp. 661–679). Universidad de Costa Rica.

Cruz Ventura, S. (2024). Las palabras de la escuela: incidencia de la formación docente en el léxico disponible del (futuro) profesorado. En *Procesos de variación y cambio en el español de Sevilla y Las Palmas de Gran Canaria* (pp. 215–235). Iberoamericana Vervuert. https://doi.org/10.31819/9783968695822_009

De la Maya Retamar, G. & López-Pérez, M. (2022). DLEx [en línea]. http://dlex.capptura.com

De la Maya Retamar, G. & López Pérez, M. (2025). DLEx, herramienta para la recuperación del léxico disponible. En A. M. Ávila Muñoz & I. C. Santos Díaz, *Avances y desarrollo de los estudios sobre el léxico disponible. Una aproximación desde la variación léxica* (pp. 299–319). Peter Lang.

Delgado-Olmos, A., Tellez-Pérez, C., & Ugueto Colina, M. (2023). El léxico de la pandemia: un estudio de disponibilidad léxica con estudiantes de la Universidad de Salamanca. *Rassegna iberistica*, *119*, 7–26. https://doi.org/10.30687/RI/2037-6588/2023/01/001

Extremera Pérez, B. (2024). *Compatibilidad léxica y percepciones categoriales compartidas por estudiantes malagueños hacia acontecimientos socialmente relevantes. Resultados parciales del proyecto PULSO Andaluz* [Tesis doctoral, Universidad de Málaga]. https://hdl.handle.net/10630/31470

Extremera Pérez, B., & Galloso Camacho, M. V. (2023). Centralidad léxica: percepción social de la inmigración en alumnos de 4. º de ESO en Huelva y Málaga. *Pragmalingüística*, *31*, 101–126. https://doi.org/10.25267/Pragmalinguistica.2023.i31.05

Gómez Devís, M.ª B., Cepeda Guerra, M. & Herranz-Llácer, C. V. (2023). Léxico disponible y redes asociativas de escolares de seis años. La configuración del lexicón en las etapas iniciales de adquisición. *Círculo de Lingüística Aplicada a la Comunicación*, *96*, 215–227. https://dialnet.unirioja.es/servlet/articulo?codigo=9237010

González Martínez, A. (2000-2001). Compatibilidad y discrepancia entre los léxicos disponibles de Puerto Rico y Cádiz. *Pragmalingüística*, 8-9, 179–196. https://dialnet.unirioja.es/servlet/articulo?codigo=266468

González Martínez, A. & Orellana Ramírez, P. (2000). Cotejo de los léxicos disponibles de Cádiz y Zamora. *Documentos de Español Actual*, *2*, 139–160.

Hernández Muñoz, N. (2006). *Hacia una teoría integrada de la disponibilidad léxica: el léxico disponible de los estudiantes castellanos-manchegos.* Ediciones de la Universidad de Salamanca.

Hernández Muñoz, N., Tomé Cornejo, C., López García, M. & Bartol Hernández, J. A. (2023). *Manual de LexPro*. https://dispogram.usal.es/

López Chávez, J. (1992). Alcances panhispánicos del léxico disponible. *Lingüística*, *4*, 26–124. https://dialnet.unirioja.es/servlet/articulo?codigo=1401403

López Chávez, J., & Strassburger Frías, C. (2000). El diseño de una fórmula matemática para obtener un índice de disponibilidad léxica confiable. *Anuario de Letras. Lingüística y Filología*, *38*, 227–251. https://revistas-filologicas.unam.mx/anuario-letras/index.php/al/article/download/985/983

López González, A. M. (2014). *Disponibilidad léxica. Teoría*, *método y análisis*. Universidad de Lodz.

López Morales, H. (1995). Los estudios de disponibilidad léxica: pasado y presente. *Boletín De Filología*, *35*(1), 245–259. https://boletinfilologia.uchile.cl/index.php/BDF/article/view/19231

Martínez Aránguiz, V., & Ávila Muñoz, A. (2025). Brechas de género en la percepción universitaria de la realidad social. *Clepsydra. Revista Internacional De Estudios De Género Y Teoría Feminista*, (29), 119–142. https://doi.org/10.25145/j.clepsydra.2025.29.06

Martínez Lara, J. A. (2023). *Léxico disponible de estudiantes universitarios de Educación Básica y Letras Hispánicas: un estudio sociolingüístico comparativo y análisis de técnicas de recolección de datos* [Tesis Doctoral, Pontificia Universidad Católica de Chile]. https://doi.org/10.7764/tesisUC/LET/75302

Martínez Lara, J. A., & Santos Díaz, I. C. [en prensa]. Análisis comparativo del léxico disponible según el formato de recogida de datos: en papel frente a digital. *Literatura y Lingüística*.

Medina Peñate, I. (2022). El léxico disponible del mar en Las Palmas de Gran Canaria y Madrid. En C. Díaz Alayón (Coord.), *Studia philologica in honorem José Antonio Samper* (pp. 691–711). Arco/Libros y Academia Canaria de la Lengua.

Prado Aragonés, J., & Galloso Camacho, M. V. (2015). *El léxico disponible de Extremadura y comparación con el de Andalucía*. Universidad de Huelva.

Rosch, E. (1978). Principles of categorization. En E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 27–48). Lawrence Erlbaum.

Samper Hernández, M. (2008). Datos comparativos entre dos léxicos disponibles: Aragón y Gran Canaria. En M. L. Arnal Purroy (Ed.), *Estudios sobre disponibilidad léxica en los jóvenes aragoneses* (pp. 245–282). Institución "Fernando el Católico".

Samper Padilla, J. A. (1998). Criterios de edición del léxico disponible. *Lingüística*, *10*, 311–333. https://dialnet.unirioja.es/servlet/articulo?codigo=1400825

Samper Padilla, J. A. (1999). Léxico disponible y variación dialectal: datos de Puerto Rico y Gran Canaria. En A. Morales et al. (Eds.), *Estudios de Lingüística Hispánica. Homenaje a María Vaquero* (pp. 551–573). Editorial de la Universidad de Puerto Rico.

Samper Padilla, J. A. & Samper Hernández, M. (2007). El proyecto Panhispánico de disponibilidad léxica y los cotejos dialectales. *Lingüística*, *19*, 88–116. https://dialnet.unirioja.es/servlet/articulo?codigo=2779602

Samper Padilla, J. A., Hernández Cabrera, C. E., & Bellón Fernández, J. J. (2003). Comparación de los léxicos disponibles de Gran Canaria y Córdoba. En C. Díaz Alayón et al. (Eds.), *Estudios sobre el español de Canarias. Actas del I Congreso Internacional sobre el español de Canarias* (Vol. 2, pp. 1077–1099). Academia Canaria de la Lengua.

Samper Padilla, J. A., Hernández Cabrera, C. E., & Samper Hernández, M. (2019). Disponibilidad léxica en niños grancanarios de 6 años. Comparación con México y Costa Rica. *Ogigia. Revista Electrónica De Estudios Hispánicos*, *25*, 139-163. https://doi.org/10.24197/ogigia.25.2019.139-163

Santos Díaz, I. C. (2017). Selección del léxico disponible: propuesta metodológica con fines didácticos. *Porta Linguarum: revista internacional de didáctica de las lenguas extranjeras*, *27*, 122–139. https://digibug.ugr.es/handle/10481/53965

Santos Díaz, I. C., Trigo Ibáñez, E., & Romero Oliva, M. (2020). Propuesta de una taxonomía de los centros de interés en los estudios de disponibilidad léxica. *DELTA: Documentação de Estudos em Lingüística Teórica e Aplicada*, *36*(4). https://revistas.pucsp.br/delta/article/view/52559

Valencia Espinoza, A. (2011). Disponibilidad léxica en Aragón y Chile. Revisión contrastiva. *Archivo de Filología Aragonesa*, *67*, 173–200. https://dialnet.unirioja.es/servlet/articulo?codigo=4056763

Valencia Espinoza, A., & Echeverría, M. S. (1999). *Disponibilidad léxica en estudiantes chilenos*. Universidad de Chile y Universidad de Concepción.

Villena Ponsoda, J. A., Ávila Muñoz, A. M., & Sánchez Sáez, J. M. (2010). Patrones sociolingüísticos del vocabulario disponible. Condicionamiento estratificacional de la capacidad léxica en la ciudad de Málaga. Proyecto CONSOLEX. En E. Ridruejo Alonso, T. Solías Arís, N. Mendizábal de la Cruz & S. Alonso Calvo (Coords.), *Tradición y progreso en la Lingüística General* (pp. 409–432). Universidad de Valladolid.

Zadeh, L. (1965). Fuzzy sets. *Information and Control*, *8*, 338–353.

https://doi.org/10.1016/0165-0114(78)90029-5

[recibido el 29 de septiembre de 2025 y aceptado para su publicación el 5 de diciembre de 2025]