

Variación diacrónica del léxico disponible de una comunidad bilingüe

Diachronic variation in the available lexicon of a bilingual community

Maribel Serrano Zapata

Universitat de Lleida, España

maribel.serrano@udl.cat

<https://orcid.org/0000-0003-4378-4389>

Abstract

In this paper we analyze lexical variation in the diachronic axis based on the data offered by lexical availability. To do this, we compared two samples of 120 pre-university students from the city of Lleida taken at different times (2003-2023). The data with respect to the two languages of the community, Spanish and Catalan, and in relation to two centers of interest (*The Human Body* and *Games and Distractions*) reveal a loss in the production of available lexicon that is accentuated in the case of the Catalan language. The two lexical fields show a contrary behavior in the period analyzed: while the vocabulary of *The Human Body* shows great stability, changes in the forms of youth leisure cause a great variability in *Games and Distractions*. The two languages follow similar patterns of diachronic evolution.

Keywords: language change, lexical availability, Spanish, Catalan

1. INTRODUCCIÓN¹

La disponibilidad léxica nace en los años 50 del siglo XX, cuando un grupo de profesores franceses intenta establecer cuál es el vocabulario fundamental que debe aprender una persona no nativa (Gougenheim *et al.*, 1964). Este tipo de estudios de

¹ Esta investigación es parte del proyecto EOS-C/N, *Estudio de los condicionantes sociales del español actual en el centro y norte de España: nuevas identidades, nuevos retos, nuevas soluciones* (ref.

léxico-estadística han tenido hasta el momento un gran desarrollo en la lingüística hispánica, gracias a Humberto López Morales, quien empezó a coordinar a finales de los 90 un ambicioso Proyecto Panhispánico con el que se pretendía describir el léxico disponible del español y al que se fueron sumando distintos estudios regionales (Samper Padilla & Samper Hernández, 2006).

La metodología del léxico disponible permite recabar palabras que están a disposición del hablante y que se activan ante un estímulo verbal determinado. De este modo, se obtienen recuentos de vocabulario donde las unidades se ordenan en función de su índice de disponibilidad, un parámetro que contempla la frecuencia y el orden de aparición de las respuestas. Entre otras posibles aplicaciones, estos datos sirven para estudiar la variación léxica, dado que son una muestra fidedigna y real del léxico disponible de las personas participantes en el estudio. De hecho, desde sus inicios, la disponibilidad ha dado cuenta de las diferencias sociales y ha sido un instrumento eficaz para la estratificación de comunidades de habla. También la variación diatópica ha encontrado en el léxico disponible un campo fecundo para estudiar y cuantificar las diferencias dialectales (López González, 2014, pp. 103-142, 183-245).

La variación diacrónica, por el contrario, ha sido pocas veces estudiada desde esta perspectiva. Esto se debe al propio desarrollo de la disciplina que, como ya hemos comentado, es relativamente reciente, aunque ha resultado fundamental en la consolidación de la lingüística aplicada en España (Mateo García & García Marcos, 2019). Hoy en día existen estudios fuera del Proyecto Panhispánico basados en muestras de población adulta (por ejemplo, Medina, 2018 o Delgado, 2024) o infantil (Samper Hernández, Hernández Cabrera & Samper Padilla, 2020; Gómez-Devís & Herranz-Llácer, 2022; Gómez-Devís, 2024; Medina & Samper, 2025), pero la mayor parte de los trabajos se enmarcan en el Proyecto Panhispánico y, por tanto, utilizan informantes preuniversitarios, de 17 o 18 años². Al abordar una única franja de edad, se hace imposible estudiar el cambio en tiempo aparente, que implica trabajar con hablantes de distintas generaciones. Sin embargo, sí es posible observar la variación léxica en tiempo real mediante un estudio de tendencia, es decir, aplicando la misma metodología de una investigación anterior al universo equivalente y utilizando un instrumento de análisis idéntico (Alba, 2014, pp. 26-28), tal como pretendemos realizar en el presente trabajo.

Por otro lado, aunque la mayoría de los estudios realizados en España son sobre el español, también se han realizado investigaciones sobre otras lenguas, como el gallego (López Meirama & Álvarez de la Granja, 2014) o el catalán (Llopis & Gómez-Devís, 2010), con lo que se ha demostrado también la validez de estos repertorios léxicos

PID2023-148371NB-C42). Agradezco, desde aquí, los comentarios de las dos personas que revisaron de forma anónima la primera versión del trabajo.

² Este es uno de los acuerdos metodológicos de las investigaciones que participan en el Proyecto, junto con el tiempo de reacción de dos minutos y los 16 centros de interés (www.dispolex.com).

para el estudio del contacto lingüístico. Aunando estos dos enfoques, el objetivo de la presente contribución es realizar un análisis de la variación léxica en el eje diacrónico comparando dos muestras de léxico disponible de estudiantes preuniversitarios de la ciudad de Lleida de dos momentos distintos (2003-2023). La situación sociolinguística de esta ciudad, caracterizada por un bilingüismo social estable y extenso, nos brinda una oportunidad interesante de evaluar la evolución diacrónica del léxico disponible del castellano y del catalán.

Por tanto, a lo largo de estas páginas, pretendemos contestar las siguientes preguntas:

1. ¿Hay diferencias entre la producción de léxico disponible en el año 2003 y en el 2023?
2. ¿Qué porcentaje del léxico disponible se muestra estable en el tiempo? ¿Cuál, por el contrario, cambia?
3. ¿Las dos lenguas siguen patrones similares con respecto a la evolución diacrónica?

Para responder a estos interrogantes, nos basaremos en los resultados obtenidos en torno a dos centros de interés, *El cuerpo humano* y *Juegos y distracciones*, y en los datos de las dos lenguas que conviven en Lleida, esto es, el castellano y el catalán.

2. MARCO TEÓRICO

Como ya hemos indicado, en el ámbito de la disponibilidad no abundan los estudios realizados con una perspectiva histórica. La obra de Alba (2013, 2014) sobre el cambio del léxico disponible en el español de la República Dominicana es pionera en este sentido. En este trabajo longitudinal, se comparan las respuestas de 347 estudiantes en 1990; con las de 420, en 2008. La aplicación de la misma metodología en la selección de los informantes y en el mecanismo de obtención de los datos permite comparar los resultados y extraer algunas conclusiones. Por una parte, se comprueba que los centros de interés con mayor o menor productividad léxica coinciden; además, hay un aumento general en la producción de léxico disponible pasados casi 20 años en todos los campos, salvo en *La ciudad*. Por otra parte, al analizar los datos en función de tres factores sociales (sexo, nivel sociocultural y zona del país), se concluye que solo el nivel sociocultural crea diferencias relevantes: la disponibilidad aumenta a medida que asciende el nivel socioeconómico de los sujetos en los dos períodos. El contraste entre los distintos estratos es mayor en 2008, hecho que Alba (2014, p. 132) relaciona con el incremento de las desigualdades sociales en el país.

Desde un punto de vista cualitativo, se comprueba cómo existen grandes diferencias en la variación léxica en función del centro de interés. De este modo, los vocablos de ciertas áreas semánticas (como *El cuerpo humano* o *Animales*) tienden a mantenerse estáticos y a perdurar en el tiempo; mientras que otras áreas (como *Comidas y be-*

bidas, *La ropa o Juegos y distracciones*) son dinámicas y susceptibles al cambio, dada su estrecha relación con las circunstancias sociales, históricas o culturales. Para ilustrar esta diferencia, se analizan los vocablos que desaparecen del repertorio de léxico disponible, o cuyo índice de disponibilidad disminuye, y también aquellos otros que se incorporan o cuya disponibilidad aumenta (Alba, 2014, pp. 85-93).

Existen otros estudios sobre variación diacrónica basados en el léxico disponible, aunque de menor alcance. Hernández & Vivar (2009) comparan los datos de 1994 sobre el español de Chile (Valencia & Echeverría, 1999) con los obtenidos en la zona urbana de Concepción en el 2005. Con dos muestras integradas por 65 sujetos, constatan que el promedio de respuestas es similar en los dos períodos, mientras que el número de vocablos asciende. Mateo García & García Mateo (2018) cotejan el léxico disponible del deporte en el español de Almería basándose en los resultados de 1992 y de 2015 referidos a *Juegos y distracciones*, *Profesiones y oficios*, *La ciudad* y *La escuela* en un análisis, sobre todo, cualitativo. Reyes Valdés *et al.* (2024), por su parte, comparan dos corpus de un corto periodo (2019-2022) con el propósito de analizar la adquisición del vocabulario durante la pandemia de la COVID-19. Dentro de la óptica del español como lengua extranjera, merece mencionarse el trabajo de Hidalgo (2017) sobre el léxico disponible de aprendices sinohablantes de ELE, también de corte diacrónico.

En relación con otras lenguas peninsulares, el único estudio de variación léxica en tiempo real es el de Lago Caamaño (2024). La investigadora analiza los cambios en el vocabulario gallego de los peces y los productos del mar, relacionándolos con los hábitos alimentarios. Para estudiar esta variación, utiliza datos de 2008 de López Meirama & Álvarez de la Granja (2014) y sus propios datos de 2020.

3. METODOLOGÍA

La presente investigación sigue fielmente la metodología utilizada por Serrano Zapata (2014) con la intención de estudiar el cambio lingüístico en tiempo real. En el citado trabajo, se analizan las respuestas de 240 informantes de la provincia de Lleida que completan en dos ocasiones la encuesta de disponibilidad: la primera, en castellano; y la segunda, en catalán. De los 240 estudiantes de 2.^º de bachillerato encuestados en aquella ocasión (curso 2002-2003), se han tomado los 120 que procedían de centros educativos ubicados en la ciudad de Lleida (seis en total: tres concertados y tres de titularidad pública). En el 2023, se ha replicado este estudio y se ha aplicado la misma encuesta de disponibilidad, tanto en castellano como en catalán, en cinco de los seis centros educativos que colaboraron en la primera investigación³. De este modo, con-

³ Agradecemos desde aquí su colaboración en nuestro estudio. Los centros que han participado en ambas investigaciones han sido: *INS Joan Oró*, *INS Ronda*, *Col·legi Episcopal*, *Col·legi Maristes Montserrat* y *Col·legi Lestonnac*. El *INS Escola del Treball*, donde pasamos encuestas en el año 2003, dejó de

tamos con una muestra final de 240 informantes que han completado la prueba asociativa de léxico disponible aportando en un tiempo de reacción de dos minutos todas las palabras que se les ocurrían relacionadas con un estímulo. Dado que las mismas personas han realizado la prueba en dos ocasiones (una en cada lengua), la cantidad total de encuestados que manejan asciende a 480.

Los estudiantes han respondido a 17 centros de interés⁴, igual que en el 2003. Por razones de espacio, en este trabajo analizamos solo los datos de *El cuerpo humano* y *Juegos y distracciones*. Hemos elegido estos dos campos asociativos tras constatar en el estudio de Alba (2014) que presentan un comportamiento claramente distinto en la variación diacrónica. Por otra parte, las personas encuestadas en 2003 conformaron el 8,6 % del total de estudiantes matriculados en 2.º de bachillerato en la ciudad de Lleida; esta cifra asciende a 9,8 % en el 2023; en ambos casos los porcentajes son representativos del universo estudiado.

Para garantizar la fiabilidad del análisis diacrónico ha sido necesario adoptar las mismas decisiones que tomamos con respecto al corpus de 2003 relativas al proceso de edición de los datos⁵ (Serrano Zapata, 2014, pp. 258-311). Además de las directrices ya clásicas de Samper Padilla (1998), conservamos algunos plurales, derivados y muchas marcas comerciales, cuya presencia es significativa en *Juegos y distracciones*. Tras este proceso de edición, las respuestas se han volcado en el software LexPro (Hernández Muñoz *et al.*, 2023). Este programa informático permite obtener los diccionarios de léxico disponible donde las palabras se ordenan en función de su disponibilidad, según la fórmula matemática de López Chávez & Strassburger Frías (2000), además del cálculo de distintos índices cuantitativos y comparativos, como veremos a continuación. Finalmente, hemos comprobado la significación estadística de los datos con el programa SPSS (versión 30.0.0).

impartir 2.º de bachillerato hace algún tiempo, con lo que fue sustituido por el INS *Manuel de Montsuar*. Se seleccionaron 20 alumnos por cada centro educativo.

⁴ Estos son *El cuerpo humano*, *Ropa y complementos*, *Partes de la casa*, *Muebles de la casa*, *Alimentos y bebidas*, *Objetos colocados en la mesa para la comida*, *La cocina y sus utensilios*, *La escuela: muebles y materiales*, *Iluminación, calefacción y formas de 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 Lago Caamaño (2023) adopta la misma postura en su estudio diacrónico de la variación del léxico disponible gallego. En nuestro caso esto ha implicado mantener, por ejemplo, “volei(bol)” y *disco(teca)*, a pesar de que en el corpus de 2023 no aparecen los acortamientos. Sin embargo, “nintendo” es ahora “nintendo (switch)”; y “teléfono” y “móvil”, separados en 2003, se unen en “(teléfono) (móvil)”, con la posible vacilación de cualquiera de los dos elementos. En 2003 los distinguimos, creyendo que el primer miembro podía hacer referencia a un juego de grupo o al teléfono fijo, opción que parece muy improbable en la actualidad.

4. ANÁLISIS

Planteamos el análisis en dos apartados: en el primero, abordamos un análisis cuantitativo general basado en los índices de palabras, vocablos y la cohesión de cada centro de interés; en el segundo, llevamos a cabo una comparación cuantitativa y cualitativa del léxico disponible de los dos campos léxicos estudiados.

4.1. ANÁLISIS CUANTITATIVO GENERAL

Antes de empezar con el cotejo de los resultados, es necesario distinguir entre los conceptos de *palabra* y *vocablo*. Ante un estímulo verbal, las personas encuestadas escriben una retahíla de términos que conforman la cantidad de *palabras* que dichas personas actualizan. Con el término *vocablo*, en cambio, se alude a las palabras *diferentes* que los encuestados producen en conjunto bajo ese mismo estímulo y con las que se conforman los diccionarios de disponibilidad. La cantidad de palabras recabada es una forma de evaluar la productividad léxica de los hablantes; en concreto, el promedio de palabras se ha reconocido muchas veces como el índice más adecuado para las comparaciones inter e intramuestrales. El número de vocablos, por su parte, informa de la riqueza léxica de un determinado centro de interés; por su propia naturaleza, no sirve para medir aspectos individuales, puesto que ofrece información sobre el comportamiento global. Con el índice de cohesión se evalúa el grado de coincidencia entre las respuestas; este se obtiene dividiendo el promedio de palabras entre los vocablos que integran una determinada área (*cf.* Serrano Zapata, 2014, p. 317)⁶.

Con respecto a los resultados obtenidos en lengua española, tal como queda patente en la tabla 1 y la figura 1, se aprecia un descenso en el total de palabras obtenidas en el periodo estudiado en los dos centros de interés. En 2003 se recogen 3114 en *El cuerpo humano*, y 1959, en *Juegos y distracciones*. En 2023, las cifras son menores, de 2942 y 1916, respectivamente. El promedio por informante, resultado de dividir la cifra anterior entre el número de participantes en la prueba, también disminuye, como es lógico: en 2003 un estudiante aportaba casi 26 respuestas en *El cuerpo humano* (25,95 exactamente) mientras que en 2023 actualiza 24,52 en el mismo tiempo de reacción, lo que supone una diferencia de casi una unidad y media (1,43). Los datos referentes a *Juegos y distracciones* corroboran la tendencia a la reducción en la productividad léxica: el promedio del año 2003 (16,32) es mayor que el del 2023 (15,97),

⁶ Los centros de interés cuyo valor de cohesión se acerca a 1 son aquellos en los que la coincidencia entre los informantes es mayor –los denominados centros compactos o cerrados–, mientras que los que propician respuestas más variadas corresponden a áreas donde la convergencia entre las personas encuestadas es menor –los centros abiertos o difusos– (Echeverría, 1991, p. 62).

aunque la diferencia es en este caso de 0,35. Tras realizar la prueba T de Student para muestras independientes, se comprueba que la diferencia es significativa en *El cuerpo humano* ($p = 0,024$; $F = 0,001$; $t = 1,993$; $gl = 238$), pero no en *Juegos y distracciones* ($p = 0,279$; $F = 0,035$; $t = 0,587$, $gl = 238$).

En relación con el número de vocablos, en cambio, se produce un incremento en el primero de los centros de interés y las personas encuestadas en 2023 aportan 47 más que en el periodo anterior; sin embargo, en *Juegos y distracciones* se recogen 14 vocablos menos. Cabe recordar que este parámetro puede alterarse fácilmente por el comportamiento de un solo sujeto atípico. Por otro lado, el índice de cohesión se mantiene constante en este centro de interés, con un 0,04 en los dos años estudiados. Esta área semántica se caracteriza por una cohesión baja, con respuestas dispares. En cambio, *El cuerpo humano* tiene un índice de cohesión próximo a 1: 0,09, en 2003; y 0,07; en 2023, lo que indica que hay una gran coincidencia entre las palabras actualizadas por las personas que realizan la prueba. El hecho de que el centro alcance una cohesión mayor en 2003 se relaciona con el descenso en el número de palabras y el aumento en el de vocablos en 2023.

Tabla 1. Índices cuantitativos (2003-2023): *El cuerpo humano* y *Juegos y distracciones*

Índices cuantitativos	<i>El cuerpo humano</i>		<i>Juegos y distracciones</i>	
	2003	2023	2003	2023
Total de palabras	3114	2942	1959	1916
Total de vocablos	287	334	417	403
Promedio palabras	25,95	24,52	16,32	15,97
Índice de cohesión	0,09	0,07	0,04	0,04

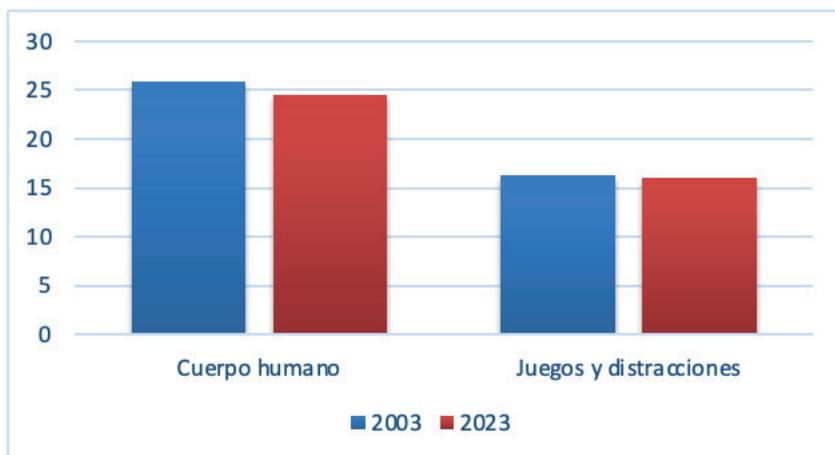


Figura 1. Promedio de palabras años 2003-2023: *El cuerpo humano* y *Juegos y distracciones*

En lengua catalana (tabla 2 y figura 2) se confirma la tendencia al descenso en el número de respuestas. Se recaban en los dos campos léxicos menos palabras y vocablos en el periodo actual, lo que evidencia una disminución tanto de productividad como de riqueza léxica. Concretamente, en el año 2023, el promedio de palabras en *El cos humà* es de 26,65; 2,09 respuestas menos que en 2003, cuando la media se situaba en 28,74. Como en lengua española, la disminución es menor en *Jocs i distraccions*, de 0,92 en este caso, la diferencia que separa las 15,72 palabras por informante de 2003 y las 14,8 actuales. Las disimilitudes en los totales son de 251 palabras en el primer centro y de 111, en el segundo. Tras realizar la prueba T de Student para muestras independientes, se comprueba que la diferencia entre los promedios es significativa en *El cos humà* ($p = 0,002$; $F = 0,698$; $t = 2,874$; $gl = 238$), pero no en *Jocs i distraccions* ($p = 0,085$; $F = 0,119$; $t = 1,377$, $gl = 238$), tal y como sucedía en lengua castellana.

Como ya hemos indicado, la pérdida léxica también se constata en el número de vocablos, aunque las cifras son realmente cercanas (288 y 285; y 390 y 374). El índice de cohesión se muestra invariable en el análisis diacrónico: se mantiene en 0,04 en

Tabla 2. Índices cuantitativos (2003-2023): *El cos humà* y *Jocs i distraccions*

Índices cuantitativos	<i>El cos humà</i>		<i>Jocs i distraccions</i>	
	2003	2023	2003	2023
Total de palabras	3449	3198	1887	1776
Total de vocablos	288	285	390	374
Promedio palabras	28,74	26,65	15,72	14,8
Índice de cohesión	0,1	0,09	0,04	0,04

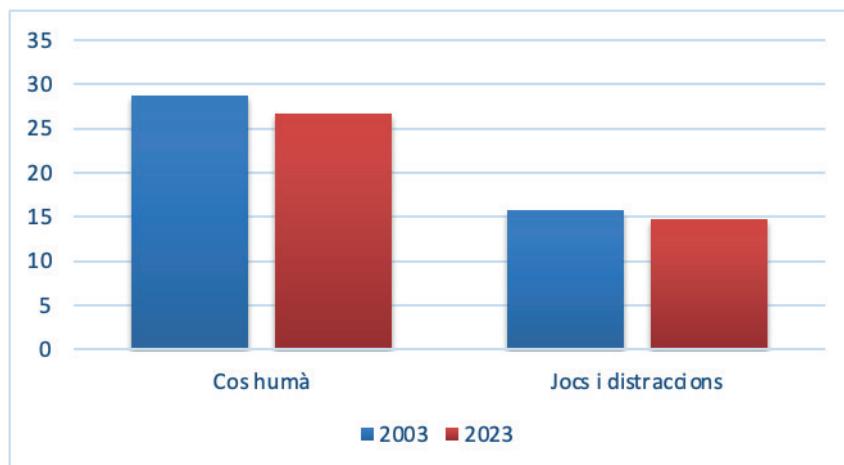


Figura 2. Promedio de palabras años 2003-2023: *El cos humà* y *Jocs i distraccions*

Jocs i distraccions, igual que en castellano, y alcanza un resultado muy similar en *El cos humà* (0,1 en 2003 y 0,09 en 2023). Sumando el descenso en los promedios de los dos campos léxicos analizados, se constata claramente que este es mayor en lengua catalana, donde se pierden 3,01 palabras por informante; en castellano, en cambio, la disminución supone una pérdida de 1,78.

Por otro lado, cabe destacar la enorme semejanza de las dos lenguas con respecto a los centros de interés: en ambos casos *El cuerpo humano* tiene una mayor productividad y una menor riqueza léxica, y se vislumbra como un área cerrada o compacta desde el punto de vista de la cohesión. En cambio, *Juegos y distracciones* propicia un mayor número de asociaciones al tratarse de un centro relacional o radial (Hernández Muñoz, 2006, p. 155), lo que se confirma con un índice de cohesión menor y un mayor número de vocablos.

4.2. COMPATIBILIDAD ENTRE INVENTARIOS LÉXICOS

A continuación, analizamos la evolución del vocabulario durante el periodo estudiado en los dos campos léxicos seleccionados (3.2.1 y 3.2.2). Esta parte del análisis se complementa con la información contenida en los apéndices.

4.2.1. EL CUERPO HUMANO/EL COS HUMÀ

En relación con el primer centro de interés (tabla 3), el léxico del castellano se muestra estable con unos porcentajes de coincidencia del 67,6 % para el 2003 respecto a 2023, y del 58,08 en dirección inversa. Las formas coincidentes en los dos períodos son 194 (de los 287 y 334 que se obtienen los dos años, v. tabla 1), dato que implica una compatibilidad general del 45,43 % sobre la suma de los conjuntos (427 vocablos). Conviene tomar con cautela esta última cifra, puesto que cada corpus tiene variación intrínseca y no podemos saber *a priori* qué porcentaje de la variación

Tabla 3. Datos cuantitativos de los vocablos compartidos y exclusivos años 2003-2023:
El cuerpo humano/El cos humà

EL CUERPO HUMANO					
Año	Vocablos exclusivos	% exclusivos	% compartidos	Vocablos compartidos	% compatibilidad
2003	93	32,40	67,60	194	45,43
2023	140	41,92	58,08		
EL COS HUMÀ					
2003	100	34,72	65,28	188	48,83
2023	97	34,03	65,96		

se debe al cambio diacrónico o a la variación inherente en cada uno de los corpus. En lengua catalana, los datos son similares. La estabilidad del vocabulario alcanza índices muy semejantes, del 65,28 % y del 65,96 %; los vocablos compartidos son 188, cifra que representa una coincidencia del 48,83 % sobre el total de unidades (385).

Las cifras que acabamos de comentar se calculan sobre la totalidad de las unidades; no obstante, sabemos que los vocablos están sujetos a fenómenos idiolectales. Por este motivo, nos hemos fijado en las 50 primeras palabras según el índice de disponibilidad, representativas de las tendencias de grupo. Si ceñimos el análisis comparativo a estas palabras, 44 son idénticas en castellano y 45, en catalán (apéndice 1), lo que supone una compatibilidad del 88 % y del 90 %, respectivamente. Además, las unidades no compartidas dentro de este margen (6 en castellano y 5 en catalán, apéndice 2) sí aparecen en las listas de disponibilidad del otro periodo analizado, aunque en posiciones más allá de la 50. Por otro lado, es notable la coincidencia interlingüística en este centro de interés, tal como puede verse en el apéndice 1.

4.2.2. JUEGOS Y DISTRACCIONES / JOCS I DISTRACCIONS

Los datos son muy distintos en *Juegos y distracciones*, un campo léxico, como ya hemos comentado, mucho más ligado a los cambios sociales y culturales. Como se ve en la tabla 4, en este centro de interés, el porcentaje de coincidencia en lengua castellana del 2003 con respecto al 2023 es del 38,37 %; en sentido contrario es del 39,70 %. En catalán también se observa que el porcentaje es superior en el 2023, con índices del 37,44 % para el primer año y del 39,04 % para el segundo. Igual que sucedía con *El cuerpo humano*, las coincidencias entre la proporción de convergencia y exclusión son análogas en las dos lenguas. Como se puede observar, la cantidad de léxico compartido es ahora menor, con cifras cercanas al 40 %. Los vocablos compartidos son 160 en español, lo que arroja una cifra global de compatibilidad del 24,24 %; en catalán, 146 palabras son idénticas, este dato supone una convergencia del 23,62 %.

Tabla 4. Datos cuantitativos de los vocablos compartidos y exclusivos años 2003-2023:
Juegos y distracciones y Jocs i distraccions

JUEGOS Y DISTRACCIONES					
Año	Vocablos exclusivos	% exclusivos	% compartidos	Vocablos compartidos	% compatibilidad
2003	257	61,63	38,37	160	24,24
2023	243	60,30	39,70		
JOCS I DISTRACCIONS					
2003	244	62,56	37,44	146	23,62
2023	228	60,96	39,04		

Si nos fijamos de nuevo en las unidades representativas del conjunto, esto es, en las 50 primeras, los porcentajes de compatibilidad aumentan, aunque quedan lejos de los vistos en el centro anterior: las palabras coincidentes son 27 en castellano y 25 en catalán (apéndices 3 y 4), lo que implica una compatibilidad del 54 % y del 50 %. En este centro de interés, contrariamente a lo que sucedía en el anterior, hay una parte importante del vocabulario exclusivo que no aparece en el otro periodo temporal estudiado. Estas palabras se encuentran destacadas en negrita en los apéndices 3 y 4. En castellano, el 26,08 % de las palabras exclusivas de 2003 desaparecen en el léxico disponible del 2023 (6 de 23); en catalán esta cifra es del 12 % (3 de 25); dos de estas unidades coindicen (*game boy* y *scattergories*). Por el contrario, hay palabras que entran en la lengua con fuerza, ocupando posiciones centrales en la disponibilidad. Estas suponen el 69,56 % en castellano (16 de 23) y el 64 % (16 de 25) en catalán de las palabras exclusivas del 2023 situadas dentro de las 50 primeras posiciones, lo que nos da una idea clara de la variabilidad del léxico en este campo. En los apéndices 3 y 4, se ilustra claramente cómo la mayor parte de estos neologismos están relacionados con el uso de las pantallas, que parece haberse convertido en una de las principales formas de ocio adolescente.

En los listados generales, conviven *playstation*, “televisión”, “ordenador” y “videojuego” con diversiones tradicionales (“parchís”, “cartas”, “damas”, “esconde-te”, etc.). No obstante, en los neologismos de 2023 imperan, tanto en castellano como en catalán, los nombres comerciales de videojuegos (*FIFA*, *Fortnite*, *Brawl Stars*, *Clash Royale*, *Minecraft*...), las videoconsolas (*nintendo switch*, *Xbox*, *wii*...) y las redes sociales (*TikTok* e *Instagram*). La aparición de “tablet” entre las 50 primeras palabras es relevante en este sentido; “teléfono móvil” ostenta una posición privilegiada de tercer puesto en castellano y quinto en el léxico disponible catalán.

5. CONCLUSIONES

Tras haber realizado el análisis, estamos ahora en disposición de responder a las preguntas que nos planteamos en la introducción. Respecto a las diferencias en la producción de léxico disponible de 2003 y de 2023, ha quedado evidenciado que, con el paso de 20 años, la disponibilidad léxica parece disminuir en Lleida. A la espera de poder ampliar este estudio con el resto de los centros de interés, no podemos más que constatar la tendencia a la pérdida en la producción de léxico disponible; más notable y con significación estadística en *El cuerpo humano*, y más acusada en catalán que en castellano. Nuestros resultados contrastan con los de Alba (2014), quien detecta un incremento de vocabulario en su estudio diacrónico sobre el español en la República Dominicana.

Con respecto al porcentaje de léxico disponible que se muestra permeable al cambio en el eje histórico, hemos comprobado cómo se detectan dos tendencias opuestas en función del centro de interés analizado: en *El cuerpo humano* el porcentaje de

vocabulario compartido supera el 60 %; mientras que en *Juegos y distracciones* esta cifra desciende al 38 %. En las 50 primeras palabras se ve con claridad esta diferencia: más del 85 % coinciden en *El cuerpo humano* y, en cambio, alrededor del 50 % en *Juegos y distracciones*. El vocabulario de este último centro es mucho más cambiante, puesto que es susceptible a las modas y tendencias temporales. Además, su propia naturaleza es difusa, lo que implica una menor convergencia en las respuestas, tal como hemos visto al analizar el índice de cohesión.

El análisis cualitativo de estos dos campos léxicos revela diferencias entre los períodos temporales. Si nos fijamos en las palabras exclusivas de cada etapa dentro de las 50 primeras, en *El cuerpo humano* (apéndice 2) destacan en 2003 términos aprendidos en la escuela (“muñeca”, “tronco”, *artèria, ronyó...*), muchos de ellos relativos a partes del esqueleto (“fémur”, “tibia”, “peroné”/*peroné...*). En el 2023, en cambio, hallamos vocablos que designan músculos (*biceps, cuádriceps*”), tal vez en consonancia con la popularización de los gimnasios en los centros urbanos, al mismo tiempo que se normalizan términos relacionados con la función reproductora y sexual (*testicle, vulva...*). En las disimilitudes relativas a *Juegos y distracciones* (apéndices 3 y 4), llama la atención la desaparición del repertorio general de léxico disponible de algunos entretenimientos tradicionales (“salir”, “teatro”, *patinar...*) y, sobre todo, la aparición de nuevas formas de ocio, vinculadas al uso de dispositivos electrónicos, muchas de ellas coincidentes en castellano y en catalán (*FIFA, Fortnite, GTA, Call of Duty, Minecraft*).

En conclusión, las dos lenguas siguen esquemas muy similares y la disponibilidad se convierte, una vez más, en una herramienta adecuada para la descripción de la variación léxica, además de testimonio de los cambios culturales de una comunidad.

BIBLIOGRAFÍA

- Alba, O. (2013). Variación diacrónica del léxico disponible dominicano. *Lingüística española actual*, 35 (1), 149-180. <http://hdl.lib.byu.edu/1877/3531>.
- Alba, O. (2014). *Observación del cambio lingüístico en tiempo real: El nuevo léxico disponible de los dominicanos*. Santo Domingo: Banco de Reservas de la República Dominicana & Pontificia Universidad Católica Madre y Maestra.
- Delgado Fernández, R. (2024). *Edad y disponibilidad léxica: La variación social en el léxico disponible adulto de Salamanca* [tesis doctoral]. Salamanca: Universidad de Salamanca.
- Echeverría, M.S. (1991). Crecimiento de la disponibilidad léxica en estudiantes chilenos de nivel básico y medio. In H. López Morales (ed.), *La enseñanza del español como lengua materna* (pp. 61-78). Río Piedras: Universidad de Puerto Rico.
- Gómez-Devís, B. (2024). *Léxico disponible de los escolares valencianos: alcance, control y evolución en educación primaria*. Barcelona: Octaedro. <https://doi.org/10.36006/16444-0>.
- Gómez-Devís, M.B. & Herranz-Llácer, C.V. (2022). Léxico disponible de escolares de la etapa primaria o básica: bases y propuesta metodológicas. *Pragmalingüística*, 30, 183-204. <https://doi.org/10.25267/Pragmalinguistica.2022.i30.09>.
- Gougenheim, G., Michéa, R., Rivenc, P. & Sauvégot, A. (1964). *L'élaboration du français fondamental (I degréé). Étude sur l'établissement d'un vocabulaire et d'une grammaire de base*. París: Didier.
- Hernández, I. & Vivar, P. (2009). El léxico en jóvenes: una exploración diacrónica. *Interprettextos* (sección *Lengua labrada*), s. n., 118-130. www.ucol.mx/interprettextos/pdfs/406_inpret415.pdf.
- Hernández Muñoz, N. (2006). *Hacia una teoría integrada de la disponibilidad léxica: el léxico disponible de los estudiantes castellanos-manchegos*. Salamanca: Ediciones de la Universidad de Salamanca.
- Hernández Muñoz, N., Tomé, C., López, M. & Bartol, J.A. (2023). *Manual de LexPro*. <https://disprogram.usal.es/>.
- Hidalgo Gallardo, M. (2017). Evolución diacrónica del léxico disponible de estudiantes sinohablantes de ELE. *SinoEle*, 16, 1-16. https://www.sinoele.org/images/Revista/16/Articulos/Hidalgo_Sino-ELE_16_2017_1-16.pdf.
- Lago Caamaño, C. (2023). El procesamiento de datos en dos muestras comparables de léxico disponible. In C. Rodríguez Campo & B. Álvarez García (eds.), *Hacia una proyección del estudio de la lengua y la literatura: Textos y contextos en el Mundo Hispánico* (pp. 149-164). León: Servicio de Publicaciones de la Universidad de León.
- Lago Caamaño, C. (2024). Cambio en el léxico disponible del gallego: los peces y los productos del mar. *Variación y cambio lingüístico*, 1, 44-58. <https://doi.org/10.30827/3020.9854rvcl.1.1.2024.31105>.
- Llopis Rodrigo, F. & Gómez-Devís, B. (2010). *El lèxic disponible de València*. València: Denes.
- 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*. Lódz: Uniwersytet Lódzki.
- López Meirama, B. & Álvarez de la Granja, M. (2014). *Léxico disponible do galego*. Santiago de Compostela: Servizo de Publicacións e Intercambio Científico, Universidade de Santiago de Compostela.
- Mateo García, M.^a V. & García Marcos, F.J. (2019). Los estudios de disponibilidad léxica en la fundamentación de la moderna lingüística aplicada en España. *Ogigia*, 25, 31-55. <https://doi.org/10.24197/ogigia.25.2019.31-55>.
- Mateo García, M.^a V. & García Mateo, P. (2018). La percepción de la actividad física en el léxico disponible en la provincia de Almería. *Lengua y Habla*, 22, 55-75. <http://erevistas.saber.ula.ve/index.php/lenguayhabla/article/view/12830/21921923928>.
- Medina Peñate, I. (2018). *El léxico disponible en el marco del PRESEEA: investigación en Las Palmas de Gran Canaria y en Madrid* [tesis doctoral]. Alcalá de Henares: Universidad de Alcalá.

- Medina Peñate, I. & Samper Hernández, M. (2025). Evolución de las redes semánticas en el léxico disponible de ‘Partes del cuerpo’ en escolares de Gran Canaria. In A. Ávila Muñoz & I.C. Santos Díaz (eds.), *Avances y desarrollo de los estudios sobre el léxico disponible. Una aproximación desde la variación léxica* (pp. 93-120). Berlin: Peter Lang.
- Reyes Valdés, D., Reyes Valdés, J.R. & Lourdes Segarra, A. (2024). Pérdida de aspiración profesional y demérito en la autoimagen desde la disponibilidad léxica comparada, pre y postpandemia, en estudiantes de nivel secundaria. *Literatura y lingüística*, 49, 433-454. <http://erevistas.saber.ula.ve/index.php/lenguayhabla/article/view/12830/21921923928>.
- Samper Hernández, M., Hernández Cabrera, C.E. & Samper Padilla, J.A. (2020). Incidencia del factor sexo/género en la disponibilidad léxica de niños grancanarios de 6 años. In F. Gimeno Menéndez (coord.), *Lengua, sociedad y cultura* (pp. 287-312). Alacant: Publicaciones de la Universidad d'Alacant.
- Samper Padilla, J.A. (1998). Criterios de edición del léxico disponible. *Lingüística*, 10, 311-333.
- Samper Padilla, J.A. & Samper Hernández, M. (2006). Aportaciones recientes de los estudios de disponibilidad léxica. *Lynx. Panorámica de Estudios Lingüísticos*, 5, 5-95.
- Serrano Zapata, M. (2014). *Disponibilidad léxica en la provincia de Lleida: estudio comparado de dos lenguas en contacto* [tesis doctoral]. Lleida: Universitat de Lleida.
- Valencia, A. & Echeverría, M.S. (1999). *Disponibilidad léxica en estudiantes chilenos*. Santiago de Chile: Universidad de Chile & Universidad de Concepción.

APÉNDICES

APÉNDICE 1

Palabras compartidas dentro de las 50 primeras palabras según la disponibilidad en 2003 y 2023:
El cuerpo humano y El cos humà

EL CUERPO HUMANO	EL COS HUMÀ
ojو	ull
mano	peu
pie	nas
pierna	cama
brazo	dit
nariz	mà
dedo	braç
cabeza	boca
boca	orella
oreja	cap
uña	ungla
corazón	cabell
pelo	genoll
diente	dent
hueso	cor
rodilla	colze
cerebro	llengua
músculo	coll
pulmón	pit
lengua	turmell
hombro	cella
codo	pulmó
cuello	cara
ceja	os
espalda	cul
pecho	panxa
vena	llavi
estómago	esquena
cara	cervell
sangre	fetge
labio	múscul
tobillo	estòmac
hígado	vena
culo	pestanya
barriga	sang
piel	espatlla
pestaña	pèl
arteria	melic

cabello		penis
riñón		canell
ombligo		fèmur
intestino		costella
pene		hombro
bíceps		tíbia
		intestí

APÉNDICE 2

Palabras exclusivas dentro de las 50 primeras según la disponibilidad en 2003 y 2023:
El cuerpo humano y *El cos humà*

EL CUERPO HUMANO		EL COS HUMÀ	
2003	2023	2003	2023
muñeca	órgano	ronyó	pell
costilla	cadera	artèria	bíceps
fémur	gemelo	cuixa	òrgan
tibia	cuádriceps	front	testicle
peroné	célula	peroné	vulva
tronco (del cuerpo)	muslo		

APÉNDICE 3

Palabras compartidas y exclusivas dentro de las 50 primeras según la disponibilidad en
Juegos y distracciones (2003-2023)

COMPARTIDAS	EXCLUSIVAS 2003	EXCLUSIVAS 2023
parchís	trivial	(teléfono) (móvil)
fútbol	(video)consola	uno
cartas	hockey	FIFA
oca	solitario	Fortnite
básquet	natación	pelí(cula)
tele(visión)	butifarra	juego de mesa
play(station)	tabú	serie
monopoly	game boy	nintendo (switch)
ajedrez	comba	Brawl Stars
escondite	discoteca	cluedo
dominó	pasear	Clash Royale
ordenador	cuerda	amigos
pilla pilla	atletismo	Minecraft
leer	billar	GTA
videojuego	internet	libro

volei(bol)	teatro	quién es quién
tres en raya	handbol	Clash of Clans
cine	(juego de) rol	Call of Duty
música	teto	Catán
damas	salir	Xbox
tenis	scattergories	pica pared
deporte	radio	tablet
póquer	saltar	wii
baloncesto		
correr		
bailar		
cantar		

APÉNDICE 4

Palabras compartidas y exclusivas dentro de las 50 primeras según la disponibilidad en
Jocs i distraccions (2003-2023)

COMPARTIDAS	EXCLUSIVAS 2003	EXCLUSIVAS 2023
parxís	cine(ma)	uno
futbol	trivial	FIFA
oca	<i>escondite</i>	(telèfon) (mòbil)
bàsquet	tres en ratlla	Fortnite
cartes	handbol	joc de taula
tele(visió)	disco(teca)	pel·lícula
escacs	corda	esport
monopoly	ordenador	sèrie
play(station)	hoquei	cluedo
llegir	passejar	llibre
ordinador	game boy	GTA
dòmino	dormir	golf
tennis	amics	Call of Duty
videojoc	solitari	pica paret
<i>pilla pilla</i>	internet	Minecraft
música	escriure	TikTok
dames	billar	tablet
volei(bol)	parlar	Netflix
(video)consola	teatre	Mario Bros
botifarra	<i>teto</i>	Instagram
ballar	tabú	nintendo (switch)
córrer	scattergories	pàdel
cantar	nedar	LOL
póquer	ajedrez	Rummikub
fet i amagar	patinar	Valorant