diacrítica
Revista do Centro de Estudos Humanísticos
Vol. 40, n.º 1, 2026, pp. 99–117. doi: 10.21814/diacritica.6878

Interferencia, disponibilidad léxica y corpus en situación de contacto de lenguas. El caso del español en Lleida

Interferência, disponibilidade lexical e corpus em situações de contacto linguístico. O caso do espanhol em Lleida

Interference, lexical availability, and corpus in language contact situations. The case of Spanish in Lleida

Mª Ángeles Calero Fernández[1]
0000-0002-4019-6031

[1] Departamento de Filología y Comunicación, Universidad de Lleida, España.

Autor correspondiente: mariaangeles.calero@udl.cat

Resumen. En situaciones de contacto lingüístico son habituales las interferencias entre las lenguas que conviven en un mismo territorio. El nivel léxico es uno de los más productivos en este tipo de traspasos. En este trabajo, se compara la interferencia del catalán en el castellano en la disponibilidad léxica de la provincia de Lleida, por un lado, con, por el otro, la transferencia del catalán al castellano en las entrevistas del corpus PRESEEA-Lérida. Se comprueba que la presencia de las interferencias se ve condicionada por los factores sexo/género, nivel de instrucción y edad. Así, las mujeres y el sociolecto alto promueven menos transferencias y la segunda generación tiene un comportamiento discordante. Aunque no de manera concluyente, se demuestra que el corpus PRESEEA-Lérida contribuye a identificar el grado de integración que tienen las transferencias identificadas en el léxico disponible.

Palabras clave: Disponibilidad léxica. Variación léxica. Contacto de lenguas. Interferencia. Español de Cataluña.

Resumo. Em situações de contacto linguístico, são frequentes as interferências entre as línguas que coexistem num mesmo território. O nível léxico é um dos mais produtivos neste tipo de transferências. Neste trabalho, compara-se a interferência do catalão no castelhano na disponibilidade léxica da província de Lleida, por um lado, com, por outro, a transferência do catalão para o castelhano nas entrevistas do corpus PRESEEA-Lérida. Verifica-se que a presença das interferências é condicionada pelos fatores sexo/género, nível de instrução e idade. Assim, as mulheres e o socioleito alto promovem menos transferências e a segunda geração apresenta um comportamento discordante. Embora não de forma conclusiva, demonstra-se que o corpus PRESEEA-Lérida contribui para identificar o grau de integração das transferências identificadas no léxico disponível.

Palavras-chave: Disponibilidade lexical. Variação lexical. Contacto linguístico. Interferência. Espanhol da Catalunha.

Abstract. Interference between languages that coexist in the same territory is common in situations of language contact. The lexical level is one of the most productive in this type of transfer. In this paper, we compare the interference of Catalan with Spanish in the lexical availability of the province of Lleida, on the one hand, with, on the other, the transfer from Catalan to Spanish in the interviews from the PRESEEA-Lérida corpus. We find that the presence of interference is conditioned by sex/gender, educational level, and age. Women and those with a high sociolect promote fewer transfers, and the second generation exhibits a discordant behavior. Although not conclusive, we demonstrate that the PRESEEA-Lérida corpus contributes to identifying the degree of integration associated with the transfers identified in the available lexicon.

Keywords: Lexical availability. Lexical variation. Language contact. Interference. Catalan Spanish.

1. Introducción

El contacto de lenguas genera una serie de fenómenos sociales y lingüísticos de gran interés y, por ello mismo, se ha convertido en objeto de estudio ya desde la lingüística histórica –en el siglo XIX– en el contexto de procesos coloniales (Gómez Rendón, 2021). No obstante, se suele considerar que el punto de arranque de la preocupación por el contacto lingüístico y sus efectos en las lenguas y en las sociedades es la obra de Uriel Weinreich (1953), en la que se ocupa de cómo interactúan dos o más lenguas en situaciones de migración o de proximidad. Cuando el territorio se convierte en el espacio de cohabitación de dos o más sistemas lingüísticos por razones históricas, políticas o económicas, surgen cuestiones como cuál es el estatus de cada uno de dichos sistemas o el régimen de convivencia que existe entre ellos (armónica o conflictiva), cuándo y cómo se selecciona una u otra lengua, si se producen influencias interlingüísticas y en qué dirección e intensidad se dan, o el grado de tolerancia ante esa influencia.

Lo cierto es que, en las situaciones de contacto, es habitual que las lenguas que cohabitan se influyan entre sí. Las interferencias que se producen pueden afectar a cualquiera de los niveles del sistema (Payrató, 1985), aunque el nivel léxico es uno de los más productivos en este tipo de traspasos. Los trabajos sobre disponibilidad léxica[1] en territorios bilingües han recogido numerosos casos de transferencias[2] de vocabulario entre las lenguas convivientes y han intentado determinar qué factores intervienen en el proceso y en qué medida las interferencias detectadas se están consolidando (Álvarez de la Granja, 2011; Casanova, 2013; Gómez Devís, 2020; Martínez Olmos, 2008; Serrano Zapata, 2022; 2024a; 2024b; Tabernero, 2008). A la hora de evaluar el grado de integración de las transferencias en el léxico disponible se tiene en cuenta tanto la frecuencia de

[1] La disponibilidad léxica se ocupa del léxico disponible, esto es, el que aflora en nuestra mente cuando surge un tema concreto (cuerpo, casa, trabajo, ciudad, campo…). En los inicios de estos estudios ya se afirmó que la disponibilidad es un instrumento apropiado para identificar los préstamos lingüísticos (Mackey, 1971, vol. 2).

[2] En este trabajo se utilizará *interferencia* y *transferencia* de manera indistinta, atendiendo a los límites semánticos que López Morales (1989) y Granda (1994) proponen entre estos dos términos junto con el de *convergencia*, tras analizar las distintas definiciones que las tres designaciones han recibido de diversos autores.

aparición como el desplazamiento que experimenta la designación patrimonial correspondiente de la lengua que recibe la interferencia (Gimeno & Gimeno, 2003). Sin embargo, los cuestionarios que se emplean para recoger el léxico disponible[3] promueven reflexiones sobre la lengua que no están conectadas con el uso, de modo que la evocación de un término por asociaciones mentales no garantiza un empleo real y efectivo en el discurso (Hernández, 2006). Ello reduce la fiabilidad del método que utilizan los estudios de disponibilidad léxica en zonas bilingües para atribuir un grado mayor o menor de integración de la interferencia en la lengua receptora. Recurrir a otro tipo de corpus lingüístico contribuiría mejor a discernirlo, como también lo haría el estudio de la eventual estabilidad de estas transferencias en el eje diacrónico (Serrano Zapata, 2025).

Un corpus lingüístico es una colección estructurada de textos (orales o escritos) producidos en situaciones reales, recogidos siguiendo criterios lingüísticos explícitos destinados a garantizar una muestra representativa, e informatizados (Calero, 2024; Rojo, 2021). Su objetivo es facilitar un mejor conocimiento de cómo es, cómo funciona y cómo se usa la lengua. El *Proyecto de Estudio Sociolingüístico del Español de España y América* (PRESEEA) busca ofrecer un corpus de español hablado que contemple sus variedades diatópicas y diastráticas. Reúne a numerosos equipos de investigación que trabajan con una metodología común en la construcción de subcorpus locales o regionales pertenecientes a comunidades tanto monolingües como bilingües o plurilingües[4]. Uno de esos subcorpus es el PRESEEA-Lérida[5], elaborado en una comunidad lingüística en la que se encuentran en contacto el español y el catalán. El español hablado en Lleida se circunscribe a la variedad norteña peninsular. En la época en la que se recogieron los cuestionarios de disponibilidad y las entrevistas sociolingüísticas de esta zona de contacto –a finales de los 90 y principios de los 2000 (como se verá)–, el 75% de la población había nacido en Lleida, la incidencia de población extranjera era testimonial (en torno a un 5%) y el resto de habitantes no necesariamente era castellanohablante, sino que también incluía personas nacidas en otros territorios de habla catalana. Al mismo tiempo, la diversidad dialectal de la población castellanohablante afincada en Lleida ha impedido que alguna variedad diatópica tenga el peso suficiente para influir en la forma de hablar español, siendo la variedad estándar –la que se aprende en las escuelas y se oye en los medios de comunicación– la que sienta las bases del modo de hablar en esta provincia catalana. Algo similar sucede en el resto de Cataluña. De ahí que podamos utilizar indistintamente la designación de *español* o de *castellano* para referirnos a la variedad que se utiliza en tierras que tienen el catalán como lengua autóctona, si no hay implícitas cuestiones identitarias o políticas que decanten el uso de uno u otro término. Los rasgos de esta variedad hispana no son uniformes en todo el territorio, ya que dependen del peso de la población autóctona, de la cantidad de lenguas en contacto, de las actitudes lingüísticas, de la lengua materna y de su variedad diatópica, de la lengua habitual, o de

[3] Los sujetos que actúan como informantes tienen que registrar en una lista todas aquellas palabras que les vienen a la mente sobre un campo léxico concreto que se ofrece como estímulo (el *centro de interés*). Para ello disponen de dos minutos en cada centro.

[4] Cf. https://preseea.uah.es/corpus-preseea

[5] El nombre que se dio al corpus en sus inicios fue este, con el topónimo patrimonial *Lérida*, y se ha mantenido así desde entonces por haber sido mencionado de este modo en diversos foros.

la identidad lingüística (Fernández & López, 2020). No obstante, existen elementos comunes que permiten describir esta variedad (Calero, 2006; Poch, 2019; Sinner, 2004; Sinner & Wesh, 2008).

En el trabajo que presentamos aquí, pretendemos analizar la interferencia léxica del catalán en el castellano hablado en Lleida. Para ello, estudiaremos la disponibilidad léxica de la lengua española en esa provincia, comparándola con la interferencia que se produce en el uso real de la lengua en las entrevistas del sociolecto alto –establecido a partir del nivel de estudios– del corpus PRESEEA-Lérida. Ya nos hemos aproximado a este particular en Calero y Serrano (2017), en donde el cotejo se hizo con los datos proporcionados por el sociolecto bajo, por ser el estrato previsiblemente más permeable a la producción de fenómenos de contacto de lenguas, al haber estado menos expuesto al español normativo a través de la escolarización. En esta ocasión, el análisis de los informantes con estudios superiores del que nos ocuparemos nos permitirá observar las eventuales diferencias de comportamiento lingüístico con respecto al sociolecto bajo, así como aquellas interferencias que podemos considerar normalizadas por ser usadas al margen del grado de instrucción y de la edad. La tesis de Casanovas (2000)[6] no es concluyente en los fenómenos morfosintácticos, puesto que solo algunas interferencias de este nivel de lengua parecían estar determinadas por la variable grado de instrucción; sus trabajos sobre el léxico son cualitativos y exponen solo ejemplos diversos extraídos de las 36 entrevistas que recogió dentro del marco del PRESEEA-Lérida (Casanovas, 2008). Gómez Devís (2013), en cambio, al analizar la relación entre los fenómenos de contacto lingüístico castellano-catalán (en la variedad de valenciano) y la variable nivel sociocultural, constata "una proporción decreciente de fenómenos inversa al sociolecto de los hablantes" (Gómez Devís, 2013, p. 370), de modo que el sociolecto bajo es el que produce un mayor número de interferencias y de cambios de código. La autora basa su análisis en las entrevistas del corpus PRESEVAL (PRESEEA-Valencia).

Nuestras hipótesis de partida son las siguientes:

a) el grado de integración lingüística de los fenómenos de contacto detectados en las encuestas de disponibilidad léxica puede medirse a partir del uso real de la lengua, de modo que el corpus del PRESEEA-Lérida va a permitirnos identificarlo;

b) el nivel de estudios actuará como un factor determinante en el trasvase del catalán al castellano porque comporta un grado diferente de exposición a la lengua normativa;

c) la intervención del factor sexo o género también será concluyente en el mayor o menor uso de interferencias, dada la relación que se ha demostrado entre sexo o género y éxito escolar[7].

[6] Parte de esta tesis está publicada en Casanovas (2005).

[7] Según el informe *Igualdad en cifras* (Ministerio de Educación, Formación Profesional y Deportes, 2025), la tasa de abandono temprano y la de repetición de curso es menor en las chicas; en cambio es mayor entre ellas el índice de finalización de los estudios en todos los niveles educativos, aunque no en todos los ámbitos del conocimiento.

2. El corpus PRESEEA-Lérida versus el léxico disponible de Lleida

El corpus PRESEEA-Lérida ha tenido una construcción dilatada en el tiempo, entre 1997 y 2007. Está constituido por 54 entrevistas semiinformales orales[8] cuya duración oscila entre 45 minutos y 1 hora y que pertenecen a informantes bilingües nacidos en Lleida o que llegaron antes de cumplir 10 años. Las entrevistas se recogieron en la capital de la provincia. Recorren diversos temas: familia, trabajo, ocio, la casa, la infancia, cambios sufridos en la ciudad, fiestas locales, situaciones hipotéticas futuras y argumentos de películas. El muestreo se hizo por cuotas a partir de los factores sociales que se tuvieron en cuenta, quedando las muestras distribuidas como se ve en la Tabla 1. Luego se aplicó una postestratificación considerando otro factor social, a saber, la lengua materna, que contemplaba, como variantes, el catalán (44 informantes, el 81,5% de la muestra) y el castellano (10 informantes, el 18,5% de la misma).

Tabla 1. Distribución de las variables sociales en la muestra del PRESEEA-Lérida (ciudad)[9]

Sexo/Género[a]	Edad[b]	Nivel de estudios
varones: 27	1.ª generación: 18	primarios: 18
mujeres: 27	2.ª generación: 18	secundarios: 18
	3.ª generación: 18	universitarios: 18

[a] Utilizamos conjuntamente los términos sexo y género porque, aunque las diferencias lingüísticas y comunicativas entre mujeres y varones no son debidas a diferencias biológicas, sino a los constructos sociales sobre qué es ser mujer y qué es ser varón (por lo tanto, lo que se entiende por género), no se tuvo en cuenta la identidad de género.

[b] Los grupos etarios establecidos en el PRESEEA son: 1.ª generación, de 20 a 34 años; 2.ª generación, de 35 a 54 años; y 3.ª generación, de 55 años en adelante.

Por su parte, el diccionario del léxico español disponible en Lleida fue confeccionado por Serrano Zapata (2014) dentro del *Proyecto Panhispánico de Estudio del Léxico disponible*[10]. El vocabulario fue recogido mediante encuestas pasadas a alumnado preuniversitario de centros de secundaria de la provincia de Lleida, incluida la capital, entre 2002 y 2003. Presenta los 16 centros de interés tradicionales en estos estudios[11], a los que se suma el campo asociativo *Los colores*. La muestra está formada por 240 informantes bilingües, etiquetados en función del sexo/género, tipo de centro

[8] Estas 54 entrevistas proceden de un fondo total de 73 entrevistas, entre las que se incluyen 36 recogidas y utilizadas por Casanovas (2000) en su tesis doctoral y 37 recogidas por Mª Ángeles Calero en el marco de un proyecto coordinado financiado por la DGICYT en el trienio 2005–2007, en el que participaban equipos de investigación de las Universidades de Valencia, Málaga, Alcalá, Las Palmas de Gran Canaria, Granada y Lleida, bajo la dirección de José Ramón Gómez Molina.

[9] Salvo que se indique lo contrario, todas las tablas han sido elaboradas por la propia autora a partir del corpus estudiado.

[10] Cf. http://www.dispolex.com/info/el-proyecto-panhispanico

[11] Estos son: *1. El cuerpo humano, 2. Ropa y complementos, 3. Partes de la casa, 4. Muebles de la casa, 5. Alimentos y bebidas, 6. Objetos colocados en la mesa para la comida, 7. La cocina y sus utensilios, 8. La escuela: muebles y materiales, 9. Iluminación, calefacción y medios para airear un recinto, 10. La ciudad, 11. El campo, 12. Medios de transporte, 13. Trabajos del campo y del jardín, 14. Animales, 15. Juegos y distracciones, 16. Profesiones y oficios*.

(público o privado), tipo de localidad (urbana o rural), nivel sociocultural y lengua materna (Tabla 2). El tipo de muestreo empleado fue por racimos o conglomerados.

Tabla 2. Distribución de las variables sociales en la muestra del léxico disponible en Lleida (provincia)

Sexo/Género	Tipo de centro	Zona geográfica	Nivel sociocultural	Lengua materna
varones: 93 mujeres: 147	público: 120 privado: 60	urbana: 120 rural: 120	alto: 36 medio: 135 bajo: 69	castellano: 66 catalán: 157 ambas: 10 aranés: 7

Como puede verse, ambas bases de datos fueron recogidas en época similar, por lo que nos ocupamos del mismo corte sincrónico. La información que ofrecen es complementaria. Por un lado, la lengua hablada en un contexto semiinformal nos permite ver el léxico en acción, por tanto, nos ayuda a determinar el grado de integración de las transferencias léxicas en el discurso real espontáneo. Por otro lado, podemos conocer el porcentaje de aparición de cada fenómeno de contacto y el lugar que ocupa con respecto al resto de elementos evocados en el léxico disponible; si bien, al estar condicionados por el estímulo que supone cada centro de interés, pueden no haberse recogido otros fenómenos usuales pero que no son inducidos por los campos asociativos empleados.

3. Fenómenos de contacto en el sociolecto alto del corpus PRESEEA-Lérida

El corpus PRESEEA-Lérida está transliterado y digitalizado, pero no ofrece informaciones cuantitativas, como el total de palabras, ni tiene etiquetadas las transferencias del catalán al castellano, por lo que deben extraerse manualmente. Sin el cómputo total de palabras, no puede establecerse el grado de integración social de los fenómenos de contacto a través de la frecuencia relativa de los casos encontrados. Tampoco es posible analizar el desplazamiento del elemento lingüístico patrimonial (fónico, gramatical, léxico o pragmático) si no se contabiliza, al mismo tiempo, las veces que dicho elemento ha sido utilizado frente a las ocurrencias de la interferencia. No obstante, podemos obtener el promedio de transferencias por informante y otros datos cuantitativos de interés.

Como pretendemos comparar las interferencias obtenidas en el corpus PRESEEA-Lérida y las identificadas en el léxico español disponible en Lleida, debemos intentar mantener la uniformidad del vocabulario que tenemos en cuenta en cada tipología de transferencia. Por ello mismo, hemos excluido del análisis del corpus PRESEEA-Lérida, tanto en el sociolecto alto como en el bajo, las muletillas y exclamaciones (tipo *oi?*, *clar!*, *home*, *Déu n'hi do!*, *i tant!*, *pam!*, *porta!* o *(això) rai!*), así como las locuciones (como *ir de bólido* ['actuar con mucha prisa por tener muchas cosas que hacer'], *día de cada día* ['día laborable'], *encontrar a faltar* ['echar en falta'], *no venir de aquí* ['carecer de importancia'] o *a dojo* ['a trote y moche']), ni las construcciones con el verbo "hacer" (como *hacer un café* ['tomarlo'], *hacer vacaciones* ['tenerlas o tomarlas'], *hacer fútbol*

[‘jugar a fútbol’] o *hacer la siesta* [‘dormirla’]). Igualmente, se han desestimado adverbios, preposiciones y conjunciones, habituales en la conversación, porque nunca aparecen en los cuestionarios de disponibilidad.

Las transferencias léxicas en las categorías del sustantivo, adjetivo y verbo que hemos hallado en el corpus PRESEEA-Lérida las hemos agrupado, por un lado, en palabras dichas en catalán, aunque se hablara en castellano y, por otro, en palabras catalanas adaptadas al español en la pronunciación o en la morfología o aquellas voces castellanas que toman prestada alguna acepción de su correspondiente término catalán.

En cuanto a las palabras catalanas en su forma original que se han usado en el sociolecto alto de este corpus, hemos recogido un total de 37 ocurrencias o concordancias (esto es, todos los términos que aparecen en la Tabla 4 junto con las veces que se han repetido, lo que se señala con un número entre paréntesis) que se reparten entre 34 transferencias distintas. Si tenemos en cuenta que se trata de 18 informantes, la media de interferencias léxicas por informante en el sociolecto alto es de 2,1. No obstante, el comportamiento lingüístico entre los grupos no es homogéneo. La distribución por sexo/género y grupo etario de los 34 casos de este tipo de traspaso léxico se presenta en la Tabla 3. Observamos que el número de catalanismos aumenta con la edad y que se da más en los varones que en las mujeres, en una proporción de casi el doble a favor del colectivo masculino.

Tabla 3. Cantidad de transferencias léxicas distintas que conservan la forma original catalana en el sociolecto alto del corpus PRESEEA-Lérida por sexo/género y edad

	1.ª generación		**2.ª generación**		**3.ª generación**		**intergeneraciones**	**Total**		
	M	**V**	**M**	**V**	**M**	**V**	**M y V**	**M**	**V**	**M y V**
N.º	3	0	4	8	5	13	1	**12**	**21**	**1**
Total	**3**		**12**		**18**		**1**	**34**		

En la Tabla 4 aportamos la información cualitativa. En ella inventariamos todos los casos que hemos documentado, con indicación entre paréntesis –si procede– de la cantidad de veces que el término catalán ha sido utilizado en cada subgrupo de la muestra y con su traducción al castellano. Como se observa, no hay coincidencias léxicas en los distintos estratos sociales. Encontramos un único caso de interferencia emitida por más de un grupo etario y por ambos sexos/géneros (en la columna “intergeneraciones”). Se trata del sustantivo *col·legi*, identificado en el colectivo de mujeres jóvenes y en el de varones de tercera generación.

Tabla 4. Palabras catalanas usadas en el sociolecto alto del corpus PRESEEA-Lérida por sexo/género y edad

Mujeres		
1.ª generación	**2.ª generación**	**3.ª generación**
canaller [kana'ʎe] ('niñero') *col·legi* ('colegio') *padrinet*[12] ('padrino de bautismo') *xerrar* [ʃa'ra] ('charlar', 'conversar')	*braçal* [bra'sal] ('brazo de río o acequia') *cavallot* ('alborotado, revoltoso', especialmente una chica) *esvaït* ('disipado', 'debilitado') *puntilletes* ('puntillas')	*caixa* ('caja bancaria') *capgrossos* (sing. capgrós: 'cabezudos') *gegants* ('gigantes') *panellets* ('dulces de almendra') *regidoria* [reʒido'ɾiɛ] ('concejalía')
Varones		
1.ª generación	**2.ª generación**	**3.ª generación**
	agafar ('coger') *aplecs* ('reunión, encuentro') *brodadores* ('bordadoras') *escorxador* ('matadero') *mitger* ('agricultor que va a medias con el propietario') *pagès* [pa'ʒɛs] ('campesino') *pocavergonya* ('desvergonzado/a') *telenotícies* [teleno'tisies] ('telediario') (2)	*castell* ('castillo') *catalans* *col·legi* ('colegio') *escolanets* ('monaguillos') *estudis* ('estudios') *enamorat* ('enamorado') *espelma* ('vela') (2) *espluga* ('cueva') *poble* ('pueblo') *professors* [profes'sos] ('profesores') *respectar* ('respetar') *resta* ('resto arqueológico) *rovellons* (2)

Como puede observarse, el vocabulario catalán insertado en el discurso en castellano es disperso. Lo mismo veremos que sucede más adelante con las adaptaciones del catalán y con las transferencias semánticas de dicha lengua al español. La razón es de carácter estructural. Las entrevistas del PRESEEA, si bien se hilvanan siguiendo un guion concreto para promover el uso de exponentes lingüísticos similares entre todos los informantes y hacer posible los estudios comparados, pueden derivar a temas particulares en cada ocasión. Dado que se pretende recoger muestras de habla extensas, resulta imprescindible incitar a los individuos entrevistados a responder con intervenciones lo más largas posibles, y una estrategia es identificar los intereses del informante y explotarlos en la interacción con él. La consecuencia directa es que se producen desequilibrios entre las entrevistas tanto en los temas tratados como en el vocabulario utilizado. Por otra parte, esta dispersión nos dificulta reconocer eventuales patrones de transferencia léxica del catalán al español.

Las palabras en catalán que hemos incluido en la Tabla 4 (y luego en la Tabla 6) no constituyen un cambio de código, es decir, no hay una voluntad de pasarse al catalán

[12] Se trata de un diminutivo afectivo de *padrí*, que en el catalán de Lleida significa 'abuelo', probablemente porque los abuelos debían de actuar como padrinos de bautismo. *Padrí* convive con *avi* i *iaio*, que son patrimoniales en catalán. La forma diminutiva (*padrinet*) se ha creado para distinguirlo de *padrí* ('abuelo') contribuyendo a ello, tal vez, el que los padrinos de bautismo fueran más jóvenes que los abuelos.

como parece indicar la autocorrección que sigue, en algunos casos, al empleo del término catalán. En las entrevistas queda muy claro cuándo el informante alterna conscientemente las lenguas, porque advierte de ello, porque usa las pausas de manera expresiva, o porque se produce una interacción rápida en catalán con la persona que entrevista

Si pasamos a las interferencias del catalán en el sociolecto alto del corpus del PRESEEA-Lérida consistentes en una adaptación de la palabra catalana a la fonética o morfología del español o resultantes del traspaso de una acepción del término catalán a su correspondiente español, nos enfrentamos a 38 transferencias distintas que se actualizan en un total de 62 ocasiones. En este caso, la media por informante es significativamente superior al promedio de palabras usadas en catalán; en concreto, es de 3,4. La distribución de las 38 interferencias en función de las variables sexo/género y grupo etario se presenta en la Tabla 5.

Tabla 5. Cantidad de transferencias léxicas distintas que son una adaptación del término catalán o la importación de un sentido de la palabra catalana en el sociolecto alto del corpus PRESEEA-Lérida por sexo/género y edad

	1.ª generación		2.ª generación		3.ª generación		intergeneraciones		Total		
	M	V	M	V	M	V	solo M	M y V	M	V	M y V
N.º	5	7	5	5	9	2	3	2	22	14	2
Total	12		10		11		5		38		

En este tipo de transferencias no se produce un patrón lineal entre los grupos etarios, como hemos visto que se daba en las palabras catalanas usadas en su forma original, sino que aquí observamos un patrón curvilíneo, puesto que la segunda generación es la que promueve menos interferencias, y las generaciones primera y tercera tienen un número mayor de casos. Por otra parte, el comportamiento lingüístico en función de la variable sexo/género es el opuesto al que hemos identificado en las palabras catalanas engarzadas en los discursos en español. Aquí son las mujeres las que sufren más la interferencia.

Que la variable sexo dé resultados contrarios en uno y otro caso nos informa de que socialmente no está bien vista la inclusión de palabras catalanas en las intervenciones en castellano, pero que, en cambio, no está tan estigmatizada la interferencia que responde a una adaptación o a una extensión semántica. López Morales (1989) afirma que las mujeres son más conscientes de la valoración social de los fenómenos lingüísticos, lo que explica que su comportamiento se alinee con lo prestigioso –o, como mínimo, con lo neutro– y rehúya lo que recibe una consideración negativa. Podría ser que este tipo de transferencia no fuera percibida como tal, es decir, que la comunidad lingüística no sea consciente de que se trata de elementos foráneos, no patrimoniales del español, por lo que no cabría ningún posicionamiento ni a favor ni en contra. En esta circunstancia, la actitud de las mujeres se alinearía con un vocabulario que se interpreta como vernáculo y normativo.

La información cualitativa que arroja el sociolecto alto para este tipo de interferencias puede verse en la Tabla 6.

Tabla 6. Palabras catalanas adaptadas al español o palabras españolas que importan un sentido del catalán recogidas en el sociolecto alto del corpus del PRESEEA- Lérida por sexo/género y edad

Mujeres		
1.ª generación	**2.ª generación**	**3.ª generación**
aparador (‘escaparate’) *costado* (‘lado’) *chopa* (‘mojada’) FAENA (‘trabajo’) *fuego a tierra* (‘chimenea’) *maja* (‘bonita’) TORRE (‘casa de campo’)	CHALET [ʧa’let] (5) *charrar* (col. ‘charlar’.) *estadas* (‘actividades extraescolares en período vacacional’) *pasturar* (usual: ‘comer’) PLEGAR[13] (‘acabar de trabajar’) *recordarse* (‘acordarse’) *retirar* (‘irse a casa’) VENIR (‘ir’)[14] (4)	CHALET [ʧa’let] *colla* (‘grupo de personas’) FAENA (‘trabajo’) (3) *firetas* (‘ferias recreativas’) *fregaplatos* (‘lavavajillas’) (2) *mosén* [mo’sen] (‘cura’) *padrina* (‘abuela’) PLEGAR (‘cerrar la empresa’) *recordarse* (‘acordarse’) *restaurant* [restau’ɾant] *sentir* (‘oír’) *tocino* (‘cerdo’) TORRE (‘casa rural, de campo’) VENIR (‘ir’) *vigilar* (‘tener cuidado, cuidar vigilando de posibles peligros’)
Varones		
1.ª generación	**2.ª generación**	**3.ª generación**
básquet[15] (‘baloncesto’) *chafardería* (‘cotilleo’) *compasado* (‘acomodado’) *conversa*[16] (‘conversación’) *destinación* (‘destino’) FAENA (‘trabajo’) (5) *pedalar* (‘pedalear’)	*acordarse* (algo) (‘recordar’) *aliniado* (del cat. ‘línia’) *aniversario* (‘cumpleaños’) (en plan) *panching* (‘arte de desconectar y no hacer nada’). *pastilleo* (‘acción de drogarse con pastillas’) VENIR (‘ir’)	FAENA (‘trabajo’) *pote* (‘bote’) *travesar* (‘cruzar’)

[13] El verbo *plegar*, en catalán, tiene más acepciones que la palabra *plegar* española. Entre esas otras acepciones de más están ‘acabar de trabajar’, ‘irse a casa después de realizar una tarea’, ‘cesar o cerrar, por ejemplo, un negocio o una tienda’. Tienen en común el sentido de que algo se acaba y de que, antes de ello, se recoge todo lo que se ha estado usando; por esta razón, consideramos estas acepciones como unidad.

[14] En catalán, el verbo *venir* se utiliza para referir tanto el movimiento hacia la persona que habla (en castellano, *venir*), como el movimiento hacia la persona que escucha (en castellano, *ir*), mientras que el verbo *anar* se emplea para movimientos hacia un lugar en el que no está ni la persona que habla ni la que escucha (en castellano, *ir*). Las ocurrencias de *venir* contabilizadas aquí son las que se han dado en un contexto en el que, en español, se usaría *ir*.

[15] En español existe esta forma, aunque la más usual en España es *baloncesto* (cf. *baloncesto* y *básquet* en CORPES XXI). En cambio, *bàsquet* es el único término que tiene el catalán, lo que influye en un mayor empleo de esta palabra en el español de Cataluña. Por eso lo hemos incluido aquí.

[16] El DLE marca esta palabra como coloquial, pero en el español de Cataluña se da en todos los registros por influencia de *conversa* en catalán, que es la traducción de *conversación*; por ello hemos incluido el término en el análisis.

petar (‘explotar’, ‘romperse’, ‘crujir’)	(2)

Constatamos, de nuevo, la dispersión de las interferencias léxicas, aunque se ha encontrado un mayor número de transferencias que aparecen más de una vez y en diversos colectivos. Volvemos a indicar entre paréntesis el número de ocurrencias que ha tenido el ítem léxico en el grupo social en el que lo hemos anotado. Al tratarse de más casos de los vistos en las palabras catalanas usadas en su forma original y al distribuirse de forma diferente según la edad o el sexo/género, reservamos las versalitas para las interferencias que se dan en más de un grupo etario (5: *faena*, *torre*, *chalet*, *plegar* y *venir*) y subrayamos aquellas que se repiten entre mujeres y varones (2: *faena* y *venir*). Este mismo proceder se seguirá en las tablas siguientes. Las 5 transferencias contabilizadas en la columna *intergeneraciones* de la Tabla 5, se distribuyen del siguiente modo: entre las voces que solo han emitido las mujeres, *chalet* y *plegar* se oyen en la segunda y la tercera generaciones, y *torre* ha sido recogida en entrevistas a mujeres de primera y tercera generaciones; en cuanto a las palabras que comparten tanto géneros como generaciones, se ha encontrado *faena*, que aparece en mujeres y varones de primera y tercera generaciones, y *venir*, documentada en mujeres y varones de segunda generación y mujeres de tercera generación.

4. Comparación entre los sociolectos alto y bajo del corpus PRESEA-Lérida en los fenómenos de contacto de lenguas

En el sociolecto bajo se han detectado 60 palabras distintas en catalán, pertenecientes a las categorías de sustantivo, adjetivo y verbo, que son utilizadas en su forma original, con un total de 88 ocurrencias[17]. La media de empleo de este tipo de transferencias por informante es de 4,9, más del doble que el promedio obtenido en el sociolecto alto, lo que nos confirma que el nivel de estudios es determinante en la presencia de fenómenos de contacto de lenguas. La distribución de los 60 ejemplos según las variables sociales se presenta en la Tabla 7.

[17] Estos datos –y los relativos a las interferencias por adaptación o resignificación en el sociolecto bajo, de los que se habla más adelante– se han calculado a partir de la revisión que hemos aplicado a los listados de fenómenos de transferencia ofrecidos por Calero y Serrano (2017), ya que los contemplados allí abarcan más categorías gramaticales, así como exclamaciones y unidades fraseológicas, que aquí no consideramos.

Tabla 7. Cantidad de transferencias léxicas distintas que conservan la forma original catalana en el sociolecto bajo del corpus PRESEEA-Lérida por sexo/género y edad

	1.ª generación		2.ª generación		3.ª generación		intergeneraciones	Total		
	M	**V**	**M**	**V**	**M**	**V**	**M y V**	**M**	**V**	**M y V**
N.º	6	3	8	17	12	12	2	**26**	**32**	**2**
Total	**9**		**25**		**24**		**2**	**60**		

En la variable edad, se observa un comportamiento más afecto a la norma castellana entre las personas menores de 35 años, quienes presentan menos de la mitad de las interferencias distintas que emergen en la segunda y tercera generaciones; en este caso, el patrón curvilíneo es tenue entre los grupos de más edad. El mayor índice de escolarización entre la juventud en Cataluña podría explicar que la primera generación sea la que menos palabras catalanas usa en sus interacciones en español.

En cuanto a la variable sexo/género, vuelven a ser los varones los que patrocinan este tipo de interferencia, aunque la diferencia con los usos lingüísticos de las mujeres se reduce respecto a lo que veíamos en el sociolecto alto. Esta circunstancia nos ratifica en la idea de que la inclusión de palabras catalanas en mensajes construidos en castellano no está bien considerada y las mujeres lo perciben. Los varones desatienden esta estigmatización por la acción del prestigio encubierto (Trudgill, 1972).

El análisis cualitativo de este tipo de transferencia nos lleva a la Tabla 8 (que revisa la Tabla 6 de Calero y Serrano [2017]), en donde obtenemos dos palabras en catalán que sobrevuelan las diferentes generaciones y ambos sexos/géneros (columna *intergeneraciones* de la Tabla 7): *escamot* y *mosso(s) de esquadra*. No hay ninguna coincidencia léxica con el sociolecto alto en este tipo de transferencias.

Tabla 8. Palabras en catalán recogidas en el sociolecto bajo del corpus del PRESEEA-Lérida por sexo y edad

Mujeres		
1.ª generación	**2.ª generación**	**3.ª generación**
boixets (’encaje de bolillos’) *cadira* (‘silla’) de ruedas *castellers* (‘personas que construyen castillos humanos’) *crosses* (‘muletas’) *enxaneta* (‘persona que corona un castillo humano’) *escudellada* (<escudella, plato típico)	*carregar* (‘cargar’) *celobert* (‘patio de luces’) *descarregar* (‘descargar’) *desplegament* (despliegue’) (3) *dir-se* (‘llamar-se) (3) *dòmino* (‘dominó’) *eix* (‘eje’) ESCAMOT (‘grupo de personas’) MOSSO D’ESQUADRA (‘mozo de escuadra’) *xiquet* (‘niño’)	*bufet* (de abogados) *canalla* (‘niños’) *carrer* (‘calle’) *cassets* (‘casetes’) *farinetes* (juego infantil’) *glassilla* (‘tela de forro’) (5) *iaia* (‘abuela’, afectivo) (2) *mare* (’madre’) *mossèn* [mo’sɛn] (‘sacerdote’) (2) *padrí* (‘abuelo’) *sastressa* (‘sastra’) *teulada* (‘tejado’)
Varones		
1.ª generación	**2.ª generación**	**3.ª generación**

bloc (‘bloque’) ESCAMOT (‘pequeño grupo de personas’) MOSSOS D’ESQUADRA *paleta* (‘albañil’) (2) *sofregit* (‘sofrito’) (2)	*autoritat* (‘autoridad’) (2) *baldufes* (‘peonzas’) *diners* (‘dinero’) *disquet* (‘disquete’) *feina* (‘trabajo’) *foguera* de Sant Joan (‘hoguera de S. Juan’) *golfa* (desván’) *ics* (‘equis’) *institut* (2) MOSSOS DE ESQUADRA *neboda* (‘sobrina’) *nen(s)* (‘niño(s)’) (2) *palet* (juego de niños) (2) *poti-poti* (‘desorden confusión’) *pous de gel* (‘pozos de hielo’) *ral* (sust. ‘real’) *teatre* (‘teatro’) *vaixell* (‘barco’)	*ballar* (‘bailar’) *ballets* (8)[18] *chalet* *colla* (8) sardanista (‘grupo de personas que se juntan para bailar sardanas) *colles* (‘grupos de gente’) *infantesa* (‘infancia’) *mallots* *pare* (‘sacerdote’) *poguer* (dialectal ‘poder’) *queviures* (víveres’) *sanitat* (‘sanidad’) *sardanes* (‘sardanas’)

Fuente: Elaboración propia a partir de los datos de Calero y Serrano (2017).

Si pasamos a las interferencias que se manifiestan como una adaptación a las estructuras del español o como una exportación de un sentido inexistente en la palabra castellana, hallamos 111 ocurrencias en el sociolecto bajo, repartidas entre 58 ejemplos diferentes (Tabla 9). La media de este tipo de transferencias por informante es 6,2, la más alta obtenida hasta ahora.

Tabla 9. Cantidad de transferencias léxicas distintas que son una adaptación del término catalán o la importación de un sentido de la palabra catalana en el sociolecto bajo del corpus PRESEEA-Lérida por sexo/género y edad

	1.ª generación		**2.ª generación**			**3.ª generación**		**intergeneraciones**			**Total**		
	M	**V**	**M**	**V**	**M y V**	**M**	**V**	**M**	**V**	**M y V**	**M**	**V**	**M y V**
N.º	6	1	8	16	1	12	3	2	4	5	**28**	**24**	**6**
Total	**7**		**25**			**15**		**11**			**58**		

De nuevo encontramos un patrón curvilíneo en la variable edad, pero, mientras que en el sociolecto alto la segunda generación era la menos propensa a este tipo de interferencias, en el sociolecto bajo es precisamente la que promueve este uso de adaptaciones del catalán o de traspasos semánticos, y lo hace con profusión, especialmente en comparación con la primera generación. Si hablamos del factor sexo/género, el comportamiento lingüístico visto en el sociolecto alto se repite: las mujeres son más propensas a utilizar este tipo de transferencias más próximas a la norma.

La Tabla 10 (que reformula la Tabla 5 de Calero & Serrano [2017]), evidencia 11 interferencias que han sido actualizadas por más de un grupo generacional (columna

[18] Se ha incluido como palabra catalana por la forma en la que ha sido pronunciada, a saber, [ba’ʎets], tal y como se hace en catalán, así como por la forma que adopta el plural.

intergeneraciones de la Tabla 9): *padecer* y *subir* se dan entre mujeres de distintas generaciones; *almorzar*, *plegar*, *regidor* y *torre* aparecen entre varones de diferentes grupos etarios; y *básquet*, *cortejar*, *faena*, *nena* y *venir* se reparten entre ambos sexos/géneros y entre diferentes grupos de edad. Al compararla con la Tabla 6, se comprueba que los dos sociolectos comparten 6 transferencias léxicas por adaptación o por extensión semántica de la palabra castellana: *básquet*, *faena*, *padrina*, *plegar*, *torre* y *venir*, a las que hay que añadir dos voces de la misma familia léxica, a saber, *chafardear* (en el sociolecto alto) y *chafardería* (en el sociolecto bajo).

Tabla 10. Palabras catalanas adaptadas al español o palabras españolas que importan un sentido del catalán recogidas en el sociolecto bajo del corpus del PRESEEA- Lérida por sexo y edad

Mujeres		
1.ª generación	**2.ª generación**	**3.ª generación**
afilerar (‘enderezar’) *aparejado* (‘emparejado’) *criaturadas* (‘niñerías’) *expolsar* (‘sacudir el polvo’) *fuego a tierra* (‘hogar donde cocinar’) *regadora* (‘regadera’)	*charrar* (‘hablar, charlar’) CORTEJAR (‘noviar tanto el chico como la chica’) *decir* (‘llamar’) (3) *dejadas* (‘abandonadas’) *despiedados* (‘sin piedad’) habitación de los enredos (‘trastos’) *mudarse* (‘vestirse de manera elegante’) (2) FAENA (‘trabajo’) (4) *no gozar* (‘no atreverse’) NENA (‘niña hasta la adolescencia’) PADECER (‘sufrir’) (2) *pulidas* (‘limpias’) *recordarse* (‘acordarse’) *repulsa* (‘bronca’) SUBIR (a alguien) (‘educarlo’, criarlo’) (2)	*antigas* (‘antiguas’) BÁSQUET (‘baloncesto’) (6) *chafarderías* (‘cotilleos’) *desparejadas* (‘desemparejadas’) *dineros* (‘dinero’) *escala* (‘escalera’) *esclatar* la guerra (‘estallar’) FAENA (7) *fregar cristales* (‘lavar cristales’) (2) *malmeter* (‘estropear’) *mosaico* (‘baldosa’) (2) NENA (2) *nuvias* (‘novias’) PADECER (‘sufrir’) (3) *padrina* (‘abuela’) *placeta* *señora grande* (‘señora mayor’) SUBIR (a alguien) (‘educarlo, criarlo’) VENIR (‘ir’)
Varones		
1.ª generación	**2.ª generación**	**3.ª generación**
ALMORZAR (‘desayunar’) (3) *envenado* (‘vendado’) FAENA (‘trabajo’) (5) girarse (‘darse la vuelta’) (5) PLEGAR (‘acabar de trabajar’) REGIDORA (‘concejala’) TORRE (‘casa de campo’) (2)	*aparador* (‘escaparate’) decirse (‘llamarse’) (2) BÁSQUET (‘baloncesto’) (2) *enfrontadas* (enfrentadas’) *enfrontamiento* (‘enfrentamiento’) *enmurallado* (‘amurallado’) *escobar* (‘barrer’) *esportivas* (‘deportivas’) FAENA (‘trabajo’) (7) *faldilla* (‘falda’)	ALMORZAR (‘desayunar’) *arrencar* (‘arrancar’) *caja de cabales* (‘caja de caudales’) CORTEJAR (‘noviar’) *padrino* (‘abuelo’) *tocinos* (‘cerdos’)

VENIR (‘ir’) (2)	*llogar* (‘alquilar’)
	NENE(S) (‘niño hasta la adolescencia’) (12)
	orquestras (‘orquestas’)
	PLEGAR (‘acabar de trabajar’)
	puesto (‘lugar’)
	rascaculos (‘tobogán’) (2)
	REGIDOR (‘concejal’)
	relliscar (‘resbalar’) (2)
	rellogado (‘realquilado’)
	sotano (‘zutano’)
	TORRE (‘casa de campo’)
	traviesas (‘vigas’)
	vieja (‘antigua’)

Fuente: Elaboración propia a partir de los datos de Calero y Serrano (2017).

5. Fenómenos de contacto presentes en los dos corpus (disponibilidad y PRESEEA)

En una primera comparación entre las transferencias obtenidas en el léxico español disponible en la provincia de Lleida[19] y las identificadas en el sociolecto bajo del corpus PRESEEA-Lérida, Calero y Serrano (2017) hallaron 15 fenómenos comunes, que son las interferencias siguientes: *baldufa* (‘peonza’), *bufet* (‘bufete’), *cadira* (‘silla’), *celobert* (‘patio de luces’), *escala* (‘escalera’), *(es)golfa* (‘desván, buhardilla’), *esportiva* (‘[zapatilla] deportiva’), *faldilla* (‘falda’), *fuego a tierra* (‘hogar, chimenea’), *mosso (d’esquadra)*, *pagès* (‘campesino’), *palet* (‘palé’), *paleta* (‘albañil’), *regadora* (‘regadera’) y *teulada* (‘tejado’).

Si la comparación la hacemos con el sociolecto alto del corpus PRESEEA-Lérida, las coincidencias con el léxico disponible disminuyen drásticamente; solo se dan dos casos: *fuego a tierra* y *pagès*.

Paleta es etiquetado por Serrano Zapata (2014, 2021) como préstamo integrado por tener una frecuencia de aparición superior al 15% en las encuestas de disponibilidad. Por su parte, *(es)golfa*, *fuego a tierra*, *mosso (d’esquadra)* y *pagès*, entran dentro de lo que Serrano Zapata califica como préstamos en vías de integración, por tener una frecuencia de aparición que va entre el 5% y el 15%. El resto se agrupa entre las transferencias cuyas ocurrencias no alcanzan el 5%; sin embargo, en el corpus PRESEEA-Lérida estas mismas transferencias se actualizan en las entrevistas, aunque solo se den entre los hablantes del sociolecto bajo, salvo *fuego a tierra* y *pagès*, que concurren también en las muestras del sociolecto alto. Su presencia en el corpus del PRESEEA-Lérida es un indicio de que su extensión social tiene un alcance mayor que el que revela la disponibilidad léxica.

No obstante, no se debe olvidar que la escasa presencia en los listados de disponibilidad léxica no puede considerarse un indicio de escasa integración, en la medida

[19] Los fenómenos de contacto de lenguas en el léxico disponible de la provincia de Lleida son consultables en Serrano Zapata (2014). Entre las páginas 504 y 537, los ofrece clasificados por tipos dentro de cada centro de interés. La taxonomía que utiliza baraja dos criterios: la frecuencia de aparición y la gramaticalidad. En las páginas 562 y 563 presenta, además, información cuantitativa (el rango, porcentaje de aparición y el índice de disponibilidad) de las transferencias que Serrano Zapata considera préstamos en vías de integración y préstamos integrados, cotejándolos con las voces españolas patrimoniales a las que equivalen.

en que puede responder simplemente a la escasa centralidad del concepto en el campo asociativo en cuestión. Para obtener conclusiones más válidas, debe tenerse en cuenta también el nivel de presencia del mismo concepto expresado en castellano. Por ejemplo, las interferencias *paleta*, *fuego a tierra* y *mosso d'esquadra* presentan índices de disponibilidad superiores a los de sus equivalente en español: 0.21888, frente a 0.09072 de *albañil*; 0.04004, frente a 0.02705 de *chimenea*, y 0.03702, frente a 0.00711 de *mozo de escuadra*, respectivamente.

Por otra parte, en la comparación entre los sociolectos alto y bajo, se han obtenido coincidencias en 5 transferencias léxicas por adaptación o por extensión semántica de los que no hay ningún testimonio en el léxico disponible: *faena*, *padrina*, *plegar*, *torre* y *venir*, además de *chafardear* (en el sociolecto alto) y *chafardería* (en el sociolecto bajo). Los campos asociativos de las encuestas de disponibilidad no facilitan la evocación de vocablos como estos, pero el corpus PRESEEA-Lérida da fe de su extensión social.

6. Conclusiones

Para descubrir cómo operan los fenómenos de contacto de lenguas resulta necesaria una aproximación multidisciplinar. Un acercamiento desde la disponibilidad léxica puede complementarse con el uso de corpus lingüísticos. Los estudios sobre léxico disponible en territorios bilingües han intentado determinar qué transferencias pueden considerarse integradas o en proceso de integración, por tanto, si han saltado ya la frontera interlingüística y los hablantes ya no las ven como elementos foráneos o extraños a la lengua que las acoge, alcanzando así el estatus de préstamo. Sin embargo, las limitaciones metodológicas del estudio de la disponibilidad exigen echar mano de otras fuentes para certificar el grado de aceptación y de uso de las transferencias lingüísticas,

Por ello, hemos acometido aquí el estudio comparado de la interferencia en el corpus PRESEEA-Lérida y en el léxico del español disponible en este territorio lingüístico. Por un lado, hemos podido ver que las transferencias del catalán al español de Lleida están condicionadas por factores sociales. Hemos comprobado que el nivel de instrucción es decisivo en el grado de permeabilidad a la interferencia, de modo que el sociolecto bajo introduce más palabras en catalán y transfiere otras que adapta formal o semánticamente al español en mayor medida que el sociolecto alto, con lo que queda confirmada la segunda hipótesis. Asimismo, hemos certificado que las mujeres se ajustan más a la norma lingüística, incorporando menos transferencias que los varones; en especial, ellas rehúyen utilizar palabras en catalán mientras hablan en castellano: confirmamos, pues, la tercera hipótesis. La edad, que inicialmente no habíamos considerado como un factor que pudiera promover la presencia o ausencia de transferencias del catalán al castellano, ha demostrado tener cierta incidencia en estos fenómenos, pero el comportamiento difiere según el sociolecto y el tipo de interferencia de la que se trate (palabras catalanas que se cuelan en el discurso en castellano, o palabras del catalán que se han adaptado a la forma y la fonética del castellano o que han ampliado el espectro semántico de palabras castellanas con las que coinciden).

Los datos cuantitativos y cualitativos arrojados en el análisis de los sociolectos alto y bajo del PRESEEA evidencian que la dimensión de la interferencia del catalán en el

español de Lleida es mucho mayor que la que arroja el léxico disponible, aunque solo estemos hablando de este nivel de lengua, el vocabulario. La diferencia entre los temas de conversación en las entrevistas, por un lado, y, por el otro, los centros de interés en las pruebas asociativas de la disponibilidad, conduce a una discrepancia entre las interferencias que se recogen en el corpus PRESEEA-Lérida y el léxico disponible del español de esta zona. No obstante, las interferencias compartidas, aun siendo pocas, nos ayudan no solo a refrendar el grado de integración imputado en el léxico disponible sobre la base de la frecuencia de aparición, sino también nos permiten ver que otros fenómenos de contacto que han aparecido testimonialmente en los cuestionarios de disponibilidad tienen una mayor difusión de la constatada aplicando esta metodología. En consecuencia, la primera hipótesis queda también confirmada.

Financiación: Este artículo se ha realizado dentro del proyecto *Estudio de los condicionantes sociales del español actual en el centro y norte de España: nuevas identidades, nuevos retos, nuevas soluciones (ECOS-C/N)*, (N.º de Referencia PID2023-148371NB-C42), financiado por el Ministerio de Ciencia, Innovación y Universidades de España y cuyos investigadores principales son Adrián Cabedo (Universidad de Valencia) y Cristina Illamola (Universidad de Barcelona).

Referencias

Álvarez de la Granja, M. (2011). La presencia del gallego en el léxico disponible del español de Galicia. Análisis formal y funcional. In B. López Meirana (Ed.), *Estudios sobre disponibilidad léxica en el español de Galicia* (pp. 17–102). Universidade de Santiago de Compostela.

Calero Fernández, M. Á. (2006). El español hablado en Cataluña: una revisión sociolingüística. In A. M. Cestero, I. Molina & F. Paredes (Eds.), *Estudios sociolingüísticos del español de España y América* (pp. 205–211). Arco/Libros.

Calero Fernández, M. Á. (2024). Análisis de errores y etiquetado en los corpus de aprendices de ELE informatizados. In M. Serrano Zapata (Ed.), *Corpus lingüísticos y enseñanza de lenguas* (pp. 31–48). Universidad de Lleida.

Calero Fernández, M. Á. & Serrano Zapata, M. (2017). Transferencias léxicas del catalán al español de Lleida. In A. M. Cestero Mancera & F. Moreno Fernández (Eds.), *Procesos de variación y cambio en el español de España. Estudios sobre el corpus PRESEEA*, Monográfico de *LinRed. Lingüística en la red*, *15*, 1–17. https://dialnet.unirioja.es/servlet/articulo?codigo=6555910

Casanova Ávalos, M. (2013). Valencianismos en el léxico disponible de Castellón. In E-Casanova & C. Calvo (Eds.), *Actas del XXVI Congreso internacional de Lingüística y Filología Románicas (Valencia 6-11 de septiembre de 2010)* (pp. 531–543). De Gruyter.

Casanovas Català, M. (2000). *Análisis cualitativo y cuantitativo de la morfosintaxis de una segunda lengua: el caso del español en contacto con el catalán* [Tesis doctoral, Universidad de Lleida]. https://hdl.handle.net/10803/8179

Casanovas Català, M. (2005). *Español y catalán en contacto. La expresión deíctica en el castellano hablado en Lleida*. Shaker Verlag.

Casanovas Català, M. (2008). Patrones léxicos en el español de los catalanohablantes. In C. Sinner & A. Wesch (Eds.), *El castellano en las tierras de habla catalana* (pp. 181–198). Iberoamericana / Vervuert.

Fernández García, A. & López Izquierdo, M. (Eds.) (2020). Lenguas en contacto con el español y desafíos sociales. Presentación. *HispanismeS, 16.*

Gimeno Menéndez, F. & Gimeno Menéndez, M. V. (2003). *El desplazamiento lingüístico del español por el inglés*. Cátedra.

Gómez Devís, M. B. (2013). Fenómenos frecuentes del contacto lingüístico en el castellano-catalán (variedad valenciano). In J. R. Gómez Molina (Coord.), *El español de Valencia. Estudio sociolingüístico* (pp. 347–370). Peter Lang.

Gómez Devís, M. B. (2020). Léxico disponible de estudiantes bilingües: presencia e integración del español en el catalán de Valencia. In F. Gimeno Menéndez (Coord.), *Lengua, sociedad y cultura: Estudios dedicados a Alberto Carcedo González* (pp. 185–204). Universidad de Alicante.

Gómez Rendón, J. (2021). El contacto lingüístico: más allá de lenguas, identidades y territorios. *Revista Argentina de Historiografía Lingüística*, *13*(1), 25–50.

Granda, G. de (1994). El español del Paraguay. Distribución. uso y estructuras. In *Español de América, español de África y hablas criollas. Cambios, contactos y contextos* (pp. 288–313). Gredos.

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.

López Morales, H. (1989). *Sociolingüística*. Gredos.

Mackey, W. F. (1971). *Le vocabulaire disponible du Français* (Vols. 1–2). Didier.

Martínez Olmos, E. (2008). *Disponibilidad léxica en las comunidades de habla alicantinas* [Tesis doctoral, Universidad de Alicante]. http://hdl.handle.net/10045/14307

Ministerio de Educación, Formación Profesional y Deportes. (2025). *Igualdad en cifras. MEFD 2025*. Subdirección General de Atención al Ciudadano, Documentación y Publicaciones. https://www.libreria.educacion.gob.es/libro/igualdad-en-cifras-mefd-2025-aulas-por-la-igualdad_185371/

Payrató, Ll. (1985). *La interferència lingüística (Comentaris i exemples catala-castellà)*. Curial-Publicaciones de la Abadia de Montserrat.

Poch Olivé, D. (Ed.) (2019). *El español de Cataluña en los medios de comunicación*. Iberoamericana/Vervuert. https://doi.org/10.31819/9783964568854

Rojo, G. (2021). *Introducción a la lingüística de corpus en español*. Routledge. http://doi.org/10.4324/9781003119760

Serrano Zapata, Maribel (2014). *Disponibilidad léxica en la provincia de Lleida: estudio comparado de dos lenguas en contacto* [Tesis doctoral, Universidad de Lleida]. http://hdl.handle.net/10803/285008

Serrano Zapata, M. (2021). Transferències lingüístiques castellà-català: un estudi de disponibilitat lèxica a Lleida. *Treballs de Sociolingüística Catalana*, *31*, 149–164. https://doi.org/10.2436/20.2504.01.183

Serrano Zapata, M. (2022). La variable lengua materna en la disponibilidad léxica de una comunidad bilingüe. *Pragmalinguistica*, *30*, 395–412. https://doi.org/10.25267/Pragmalinguistica.2022.i30.18

Serrano Zapata, M. (2024a). Dels sostens* a la "pica": lengua materna y transferencias castellano-catalán. In L. Mariottini & M. Palmerini (Eds.), *Avances en los estudios de lingüística hispánica: perspectivas teóricas y aplicadas entre lengua y sociedad* (pp. 339–356). Dykinson.

Serrano Zapata, M. (2024b). *El castellano y el catalán en contacto. Efectos sobre el vocabulario*. Peter Lang.

Serrano Zapata, M. (2025). Anàlisi diacrònica de les transferències lèxiques castellà-català a les primeres dècades del segle XXI. *Caplletra. Revista Internacional de Filologia, 78*, 51–71. https://doi.org/10.7203/caplletra.78.30265

Sinner, C. (2004). *El castellano de Cataluña. Estudio empírico de aspectos léxicos, morfosintácticos, pragmáticos y metalingüísticos*. Marx Niemeyer Verlag.

Sinner, C. & Wesch, A. (Eds.) (2008). *El castellano en las tierras de habla catalana*. Iberoamericana/Vervuert.

Tabernero, C. (2008). Disponibilidad léxica y contacto de lenguas. *Oihenart: Cuadernos de Lengua y Literatura*, *23*, 545–565.

Trudgill, P. (1972). Sex, covert prestige and linguistic change in the urban British English of Norwich. *Language in Society, 1*(2), 179–195. https://doi.org/10.1017/S0047404500000488

Weinreich, U. (1953). *Languages in contact*. Linguistic Circle of New York.

[recibido el 1 de octubre de 2025 y aceptado para su publicación el 20 de enero de 2026]