

Corpus y contraste pragmático de interjecciones
Limitaciones metodológicas en el estudio de *o* (polaco) y *oh* (español)

Corpus and pragmatic contrast of interjections
Methodological limitations in the study of *o* (Polish) and *oh* (Spanish)

Raúl Fernández Jódar

Uniwersytet im. Adama Mickiewicza w Poznaniu (Poznań, Polonia)

raul.fernandez.jodar@amu.edu.pl

<https://orcid.org/0000-0002-5173-5897>

Recibido el 27/5/2025, aceptado el 3/11/2025, publicado el 17/4/2026

Creative Commons Attribution 4.0 International (CC BY 4.0)

© 2026 Raúl Fernández Jódar

Cómo citar este artículo

Fernández Jódar, Raúl 2026. Corpus y contraste pragmático de interjecciones. Limitaciones metodológicas en el estudio de *o* (polaco) y *oh* (español). *Studia linguistica romanica* 2026.15, 1-24. <https://doi.org/10.25364/19.2026.15.1>.

Resumen

Este artículo examina el comportamiento pragmático de las interjecciones *o* en polaco y *oh* en español a partir de corpus orales y escritos, con especial atención a sus funciones expresivas y discursivas en contextos reales de comunicación. Desde una perspectiva contrastiva, se identifican tanto valores compartidos como divergentes, particularmente en el modo en que ambas interjecciones se integran en la interacción espontánea. El análisis se basa en datos procedentes de corpus del español como lengua extranjera (*CAES* y *Corpus oral de ELE*) y del corpus oral polaco *Spokes*, y permite observar una marcada asimetría en la frecuencia, distribución y anotación de estas formas. A partir de estos datos, se reflexiona sobre las limitaciones metodológicas que presentan los corpus actuales para el estudio de fenómenos pragmáticos sensibles al contexto, como las interjecciones, y se subraya la necesidad de avanzar hacia modelos de corpus más representativos, anotados y funcionales para la investigación contrastiva y la didáctica de lenguas.

Palabras clave

Español, interjección, limitación metodológica, lingüística de corpus, polaco, pragmática contrastiva.

Abstract

This article examines the pragmatic behavior of the interjections *o* in Polish and *oh* in Spanish, based on data from oral and written corpora, with a special focus on their expressive and discursive functions in real communicative contexts. From a contrastive perspective, both shared and divergent values were identified, particularly in the way how these interjections are integrated into spontaneous interaction. The analysis draws on data from L2-corpora of Spanish (*CAES* and *Corpus oral de ELE*) and from the Polish oral corpus *Spokes*, revealing a marked asymmetry in the frequency, distribution and annotation of these forms. Based on these findings, the article discusses the methodological limitations of current corpora for the study of context-sensitive pragmatic phenomena, such as interjections, and highlights the need for representative, annotated and functionally oriented corpora for contrastive research and language teaching.

Keywords

Contrastive pragmatics, corpus linguistics, interjection, methodological limitation, Polish, Spanish.

Índice

1	Introducción.....	4
2	Las interjecciones en la teoría lingüística.....	5
2.1	Definición y rasgos formales.....	5
2.2	Clasificación gramatical.....	5
2.3	Clasificación semántico-pragmática.....	6
2.4	Funciones expresivas y discursivas.....	7
2.5	Relevancia para el análisis contrastivo.....	7
3	Valores pragmáticos de <i>oh</i> (español) y <i>o</i> (polaco): análisis contrastivo.....	7
3.1	Valores compartidos.....	8
3.2	Valores divergentes.....	8
3.3	Consideraciones sintácticas y prosódicas.....	12
4	Corpus y análisis pragmático: potenciales y limitaciones.....	12
4.1	Representatividad de las interjecciones en corpus escritos vs. orales.....	13
4.2	Ejemplos contrastivos: <i>CAES</i> , <i>Corpus oral de ELE</i> , <i>Spokes</i>	13
4.3	Riesgos de generalización desde corpus desequilibrados.....	16
4.4	El papel del <i>data-driven learning</i> y la conciencia pragmática.....	16
5	Hacia corpus más funcionales para la investigación y la didáctica.....	17
5.1	Criterios para la construcción de corpus pragmáticamente fiables.....	17
5.2	Corpus pedagógicos: propuestas desde la literatura reciente.....	18
5.3	La necesidad de etiquetado pragmático y metainformación contextual.....	19
5.4	Proyectos emergentes y líneas futuras.....	19
6	Implicaciones didácticas.....	19
6.1	La interjección como componente de la competencia pragmática.....	20
6.2	Integración de corpus en la enseñanza del componente pragmático.....	20
6.3	Sensibilización contrastiva y prevención de interferencias.....	20
6.4	Formación docente y conciencia metapragmática.....	21
7	Conclusiones.....	21
	Abreviaturas y referencias bibliográficas.....	23

1 Introducción

[1] Las interjecciones son unidades lingüísticas autónomas y altamente expresivas que permiten al hablante reaccionar ante la realidad de forma sintética e inmediata. Por su estrecha vinculación con la subjetividad, la oralidad y la interacción, constituyen un objeto de estudio privilegiado para el análisis pragmático. Entre ellas, las interjecciones *oh* en español y *o* en polaco presentan cierto nivel de paralelismo formal y funcional, lo que permite interrogarnos por sus valores comunicativos compartidos y divergentes. Esta comparación, sin embargo, no puede abordarse únicamente desde el plano gramatical o semántico, sino que exige una aproximación empírica que tenga en cuenta los contextos reales de uso.

[2] En este marco, el presente trabajo se inscribe en una doble perspectiva: contrastiva y corpus-lingüística. Por un lado, se propone un análisis pragmático-comparativo entre *oh* y *o* a partir de sus funciones en enunciados auténticos extraídos de corpus del español y del polaco. Por otro lado, se plantea una reflexión crítica sobre las posibilidades y limitaciones metodológicas que conlleva este tipo de estudio, especialmente en relación con la disponibilidad de datos orales en los corpus existentes.

[3] La elección del enfoque corpus-lingüístico no responde únicamente a razones descriptivas, sino también a un interés metodológico más amplio: evaluar en qué medida los corpus actuales permiten representar con fidelidad los usos interaccionales de fenómenos altamente contextuales, como las interjecciones. En efecto, tal como se observa en § 4, los corpus de español como lengua extranjera (ELE), como el *CAES* o el *Corpus oral de ELE*, ofrecen una presencia muy reducida de la interjección *oh*, lo que contrasta con la frecuencia y variedad de usos de *o* en el corpus oral polaco *Spokes*. Esta asimetría plantea interrogantes no solo sobre los sistemas lingüísticos comparados, sino también sobre los criterios de diseño, etiquetado y representatividad de los corpus disponibles.

[4] Por ello, el análisis empírico se articula en dos bloques complementarios. En primer lugar (§ 3), se examinan los valores pragmáticos de las interjecciones *oh* española y *o* polaca, atendiendo a sus coincidencias y divergencias funcionales en distintos contextos comunicativos. En segundo lugar (§ 4 y 5), se evalúan las implicaciones metodológicas del trabajo con corpus, subrayando la necesidad de avanzar hacia recursos más equilibrados y adecuados para la investigación de fenómenos pragmáticos. Esta parte se apoya en la literatura reciente sobre corpus pedagógicos y *data-driven learning*, así como en estudios que abogan por un etiquetado más fino de las unidades con carga pragmática.

[5] En conjunto, el artículo busca no solo describir el comportamiento pragmático de las interjecciones *oh* y *o*, sino también examinar críticamente los supuestos metodológicos que subyacen al uso de corpus en el análisis del discurso oral. Se plantea así un marco que permite avanzar en la comprensión contrastiva de fenómenos pragmáticos y en su integración en la enseñanza de lenguas.

2 Las interjecciones en la teoría lingüística

[6] Las interjecciones constituyen una categoría tradicionalmente marginal dentro de la gramática, debido a su difícil encaje en los modelos sintácticos y semánticos convencionales. Sin embargo, su presencia constante en la interacción oral, así como su alta carga expresiva y su potencial discursivo, las convierte en unidades clave para el estudio de la pragmática del habla.

2.1 Definición y rasgos formales

[7] Desde el punto de vista formal, las interjecciones se caracterizan por su independencia sintáctica, su invariabilidad morfológica y su entonación marcada (Alarcos Llorach 1994). A menudo constituyen enunciados completos por sí solas y no establecen relaciones de concordancia ni de subordinación con otros elementos del discurso. En español, ejemplos típicos son *ay*, *oh*, *uf* o *eh*, mientras que en polaco se encuentran formas como *o*, *ach*, *oj* o *no*.

[8] Estas unidades suelen aparecer en posición periférica, bien al inicio del enunciado, bien como inserciones parentéticas, y a menudo van acompañadas de pausas y modulaciones prosódicas significativas. Su función primaria no es referencial, sino expresiva o relacional, lo que explica la limitación de los mecanismos de análisis tradicionales basados en la predicación.

2.2 Clasificación gramatical

[9] Las gramáticas descriptivas suelen considerar las interjecciones como una clase cerrada y periférica, no siempre integrada en la oración. Esta visión está presente también en Alarcos Llorach (1994), quien subraya su carácter marginal desde el punto de vista gramatical. La *NGLE* (§ 32.1) las define como «palabras o expresiones que constituyen enunciados independientes y que expresan sentimientos súbitos o apelaciones breves». En este sentido, se las considera una de las pocas unidades con capacidad para constituir un acto de habla completo sin necesidad de estructura proposicional.

[10] No obstante, esta clasificación ha sido revisada en estudios más recientes que cuestionan la homogeneidad del conjunto. Autores como Cuenca (2006) proponen distinguir entre interjecciones prototípicas, de carácter emocional, y unidades procedentes de otras clases gramaticales que han adquirido funciones interjectivas en el discurso, como *bueno*, *mira* u *oye*. Estas formas gramaticalizadas cumplen funciones conversacionales y actúan como marcadores del discurso, más que como expresiones emotivas o fáticas.

[11] Por su parte, la tradición gramatical polaca (véanse Nagórko 2010; Polański 2000; Derwojedowa, Karaś & Kopcińska 2005), a diferencia de la hispánica, considera las onomatopeyas (*wyrazy dźwiękonaśladowcze* u *onomatopeje*) como un subgrupo semántico de las interjecciones (*wykrzykniki* o *interiekcje*), que, a su vez, se describen como una clase de palabras invariables en cuanto a género y número y sintácticamente independientes. Se suelen escribir con un signo de ex-

clamación (!). Semánticamente, pueden expresar una emoción (*oj, brr*), un deseo, una orden o una voluntad de provocar una reacción (*hej, halo, pst*), o también pueden ser una imitación fónica de los sonidos producidos por diversos objetos, acciones o movimientos (*szur, ciach, chlup, bzz*).

2.3 Clasificación semántico-pragmática

[12] Desde una perspectiva funcional, las interjecciones se han clasificado en función del tipo de acto de habla o actitud que expresan. Wilkins (1992) y Ameka (1992) proponen una tipología universal, basada en actos primitivos comunicativos, que incluye:

1. Interjecciones emotivas, que expresan estados afectivos del hablante: *oh* (español), *ay* (español), *ach* (polaco), *oj* (polaco).
2. Interjecciones conativas, que buscan una reacción del oyente: *eh* (español), *oye* (español), *no* (polaco), *ej* (polaco).
3. Interjecciones fáticas, destinadas a iniciar, mantener o cerrar el contacto: *hola* (español), *adiós* (español), *cześć* (polaco), *na razie* (polaco).
4. Interjecciones descriptivas o imitativas, que representan sonidos o acciones: *pum* (español), *zas* (español), *bam* (polaco), *buch* (polaco).

[13] Cabe señalar que las categorías 3 y 4 de esta tipología – interjecciones fáticas y descriptivas o imitativas – han sido objeto de debate. En el primer caso, fórmulas como *hola* o *adiós* se describen a menudo como expresiones de cortesía o marcadores conversacionales más que como interjecciones en sentido estricto, lo que explica que algunas gramáticas no las incluyan dentro de esta clase (cf. *NGLE*). En el segundo, la inclusión de onomatopeyas (*zas, pum* en español; *buch, bam* en polaco) como interjecciones plantea problemas categoriales: mientras que en la tradición polaca se consideran un subgrupo semántico de las interjecciones (Nagórko 2010; Pawlak 2017), en la tradición hispánica a menudo se tratan como signos icónicos independientes. Estas discrepancias muestran que la tipología de Ameka (1992) y Wilkins (1992), aunque útil, debe entenderse como un marco de referencia flexible y no exento de controversias.

[14] Ahora bien, muchas formas, como *oh* u *o*, pueden tener usos polisémicos que dependen del contexto comunicativo, la entonación y el tipo de interacción. Por ello, algunos autores insisten en la necesidad de considerar no solo la forma, sino también la función pragmática y el entorno discursivo de cada aparición concreta (Cuenca 2006; Blakemore 2002). Entre ellos, Cueto Vallverdú & López Bobo (2003) destacan la doble naturaleza, emocional y discursiva, de muchas interjecciones en uso real, mientras que Wierzbicka (1992, 1996) añade que muchas interjecciones codifican marcos semánticos profundamente culturales y no son directamente trasladables entre lenguas, incluso cuando coinciden formalmente.

2.4 Funciones expresivas y discursivas

[15] Una distinción fundamental para este estudio es la que se establece entre interjecciones expresivas y discursivas. Las primeras constituyen reacciones inmediatas a un estímulo, de carácter emocional o evaluativo (*oh, ay, ach*), mientras que las segundas actúan como instrucciones pragmáticas que orientan al oyente en la interpretación del discurso (*bueno, mira, no*). En muchos casos, una misma forma puede cumplir ambas funciones según el contexto.

[16] *Oh* y *o*, foco de este estudio, pertenecen prototípicamente al grupo de interjecciones expresivas, pero pueden adquirir funciones discursivas de apertura, de encuadre narrativo o de énfasis interpretativo. Así, en español, *oh* puede introducir una valoración afectiva o marcar una transición discursiva con tono emocional, mientras que en polaco *o* puede anunciar la aparición de un referente narrativo inesperado o activar la atención.

[17] Desde una perspectiva cognitivo-pragmática, Blakemore (2002) ha propuesto que algunas interjecciones funcionan como expresiones procedimentales, es decir, no aportan contenido conceptual, sino que instruyen al oyente sobre cómo procesar la información que sigue. Esto refuerza la necesidad de analizarlas no solo desde su valor léxico, sino también como operadores inferenciales dependientes del contexto.

2.5 Relevancia para el análisis contrastivo

[18] El análisis de interjecciones como *oh* y *o* requiere una aproximación que combine la perspectiva gramatical con el análisis contextual y funcional. Su comportamiento no puede explicarse adecuadamente sin considerar la modalidad de producción (oral o escrita), el tipo de interacción, el tono prosódico y el conocimiento compartido entre interlocutores. Por ello, el estudio contrastivo que se presenta en este artículo se basa en corpus que reflejan, con diferentes grados de fiabilidad, estos aspectos del uso real. Además, en la línea de trabajos contrastivos como Baran (2010), Pawlak (2017) o Wicherek & Balches Arenas (2012), se observa que las interjecciones ocupan un lugar problemático desde el punto de vista categorial en la comparación español-polaco.

3 Valores pragmáticos de *oh* (español) y *o* (polaco): análisis contrastivo

[19] Las interjecciones *oh* en español y *o* en polaco presentan notables paralelismos formales: ambas son formas breves, vocálicas, sin inflexión, que tienden a ocupar una posición periférica en el enunciado y a acompañarse de entonación exclamativa. Sin embargo, su comportamiento pragmático no es enteramente coincidente. En este apartado se analizan sus valores desde una perspectiva contrastiva, distinguiendo usos compartidos y usos divergentes, tal como se manifiestan en corpus orales y escritos. La clasificación funcional adoptada en este apartado combina las tipologías propuestas por Ameka (1992), Wilkins (1992) y Cuenca (2006) con el análisis empírico de los datos de corpus. Con el fin de mantener la claridad expositiva, se presentan ejemplos tanto auténticos – extraídos de los cor-

pus consultados (*CAES*, *Corpus oral de ELE*, *Spokes*) – como construidos con fines ilustrativos.

3.1 Valores compartidos

[20] Tanto *oh* como *o* pueden funcionar como interjecciones expresivas de reacción inmediata ante un estímulo emocional. En este sentido, comparten los siguientes valores:

- Sorpresa o asombro, tanto ante lo inesperado como ante lo extraordinario:
Oh, no lo sabía.
O, naprawdę? '¿En serio?'
- Admiración o entusiasmo, ante un estímulo positivo:
¡Oh, qué maravilla!
O, jakie to piękne! '¡Qué bonito!'
- Pena o tristeza:
Oh, lo siento mucho.
O, to przykre. 'Eso es triste.'

En estos casos, ambas interjecciones introducen un juicio afectivo, cuya intensidad puede ser modulada por la entonación o por recursos léxicos adyacentes. La entonación ascendente o descendente, así como la duración vocálica, juega un papel clave en la interpretación emocional, especialmente en el discurso oral.

[21] Asimismo, en ambos idiomas puede observarse el uso de la interjección como marcador empático dirigido al interlocutor, en contextos de escucha activa o refuerzo afectivo:

- *Oh, entiendo lo que dices.*
- *O, rozumiem.*

[22] Estos usos compartidos sugieren un núcleo pragmático común, anclado en la dimensión afectiva de la interacción, y permiten postular una cierta equivalencia funcional básica entre *oh* y *o*.

3.2 Valores divergentes

[23] Los numerosos valores compartidos provocan que las diferencias entre ambas interjecciones queden difuminadas y sean difíciles de detectar para un aprendiz polaco de ELE, independientemente de su nivel. Sin embargo, existen ciertos valores no compartidos que merecen especial atención por su potencial interferencia en la producción espontánea del español y su repercusión en la naturalidad discursiva.

[24] Uno de los valores no equivalentes más significativos entre *o* en polaco y *oh* en español se observa cuando la interjección polaca se emplea para introducir un ejemplo que refuerza o confirma una afirmación previa del hablante. En estos casos, *o* no solo cumple una función ilustrativa, sino que adquiere un matiz pragmático adicional, que aquí proponemos denominar *satisfacción discursiva*. Con esta etiqueta nos referimos al momento en que el hablante encuentra una ocurrencia concreta que valida o confirma con éxito su argumento, produciendo un efecto

de refuerzo positivo y de autoafirmación en el discurso. A modo de ejemplo ilustrativo, en una conversación espontánea podría aparecer la secuencia *O, na przykład w moim mieście ...* 'Oh, por ejemplo, en mi ciudad ...', donde la interjección introduce un ejemplo que confirma la afirmación precedente y refuerza la posición del hablante. Se trata, por tanto, de una propuesta terminológica propia, destinada a describir un valor pragmático no registrado en la tradición gramatical polaca ni en los estudios contrastivos previos, pero observable en los datos analizados. Tal como han observado Bańko (2001) y Grochowski (1991), la forma *o* ocupa una posición ambigua entre el estatus de interjección plena y partícula discursiva, lo que refleja su especialización pragmática. Así pues, esta interjección no solo señala una ejemplificación, sino también la seguridad subjetiva del hablante respecto a la pertinencia de lo que va a decir o ha dicho, lo que genera una carga afectiva positiva.

[25] Este uso es natural y frecuente en el polaco oral espontáneo, pero no encuentra un correlato funcional claro en español. En esta lengua, *oh* rara vez introduce ejemplos, y mucho menos cumple la función de señalar satisfacción discursiva. Como ejemplo ilustrativo, una frase como *¡Oh, basta con oír lo que dicen mis vecinos!* no funcionaría como ejemplificación validante, sino que se interpretaría como exclamación de sorpresa o énfasis emocional, lo que evidencia la falta de equivalencia semántica directa con *o*. Su uso se asocia más habitualmente con la manifestación de emociones como sorpresa, admiración, miedo o dolor, lo que condiciona su comportamiento en el discurso. Esta caracterización coincide, en gran medida, con las definiciones ofrecidas por los diccionarios normativos del español, que describen *oh* como una interjección usada para denotar asombro, sorpresa o admiración (*DLE*, s.v. *oh*; *DUE*, s.v. *oh*; Seco, Andrés & Ramos 1999, s.v. *oh*). En contraste, los diccionarios de referencia del polaco – como el *PWN* (s.v. *o*) o Dubisz (2003, s.v. *o*) – destacan para *o* valores expresivos próximos a la sorpresa, la admiración o la compasión, lo que confirma una base semántica compartida con *oh*, aunque con una proyección discursiva más amplia.

[26] Como resultado, cuando se traduce un enunciado polaco con *o* en esta función al español, es frecuente que la interjección se omita por completo, lo que refleja la falta de una correspondencia semántica exacta con *oh*. Sin embargo, esta omisión no debe interpretarse como ausencia de equivalencia pragmática, sino más bien como evidencia de que la equivalencia debe buscarse en otras unidades interjectivas del español – tales como *bueno*, *vaya*, *ah* o *eh* –, que, según el contexto, pueden desempeñar funciones de refuerzo discursivo, introducción de ejemplos o implicación afectiva. En este sentido, la traducción puede optar por eliminar la interjección o por sustituirla con un recurso interjectivo alternativo, de modo que el ajuste pragmático no se entienda como imposibilidad de uso de interjección, sino como ausencia de equivalencia formal y semántica estricta entre *o* y *oh*.

[27] Así, por ejemplo, un hablante polaco podría decir:

- (1) O, ja mówię to ja się już nawet cieszę no bo przecież to są ludzie tacy wiesz milionerami można ich nazwać.
'Oh, yo digo que hasta me alegro, porque, claro, son gente que, ya sabes, se podría llamar millonarios.'

Aquí, *o* introduce un ejemplo concreto que corrobora una afirmación general. El hablante no solo quiere ilustrar, sino también subrayar que su afirmación no es abstracta: hay evidencia palpable. La interjección marca ese 'momento de prueba' con una actitud de afirmación casi enfática.

[28] Otro ejemplo sería:

- (2) Coraz więcej ludzi nie ufa politykom. O, wystarczy posłuchać, co mówią moi sąsiedzi.
'Cada vez más gente desconfía de los políticos. Ø Basta con oír lo que dicen mis vecinos.'

De nuevo, *o* introduce no solo un ejemplo, sino una confirmación argumentativa, algo que refuerza el discurso con implicación personal. En la versión española, la interjección desaparece porque *oh* no puede cumplir esa función de manera idiomática ni natural.

[29] De hecho, insertar *oh* en español en contextos análogos puede resultar forzado o artificioso:

- (3) Cada vez más gente desconfía de los políticos. ¡Oh! Basta con oír lo que dicen mis vecinos.

En este caso, *oh* no actúa como marcador de satisfacción o refuerzo, sino que se interpreta más bien como una exclamación de sorpresa, con lo cual no cumple la función deseada. Por ello, la traducción más natural consiste en omitir la interjección y confiar en el contenido proposicional y en el contexto para transmitir el efecto discursivo.

[30] Esta diferencia de valor pragmático representa una fuente potencial de transferencia negativa en el aprendizaje del español por parte de hablantes polacos. Al percibir una aparente equivalencia formal entre *o* y *oh*, el aprendiente puede sentirse tentado a trasladar directamente estructuras discursivas del polaco al español, generando enunciados gramaticalmente correctos pero poco idiomáticos o estilísticamente marcados. De ahí la importancia de explicitar esta divergencia en el aula desde un enfoque contrastivo.

[31] Desde el punto de vista sintáctico, la interjección *o* en polaco presenta también una mayor flexibilidad posicional. Puede aparecer antes o después del

segmento que introduce, lo que le permite adaptarse a distintas configuraciones discursivas. En cambio, en español, aunque construcciones como *Tenemos libre mañana, ¡oh!* son gramaticalmente posibles, resultan inusuales en el habla cotidiana. La forma más natural sería *¡Oh!, mañana tenemos libre*, lo que refleja una preferencia clara por la posición inicial en el enunciado.

[32] Esta diferencia queda ilustrada en los siguientes ejemplos extraídos del corpus de lengua oral espontánea *Spokes*:

- (4) Łódź nowych miastem nowych technologii. O!
'Łódź, la nueva ciudad de las nuevas tecnologías. Ø.'
- (5) No i Magda mówiła że dzwonił że robi tam ale na razie no ni ma tak żeby rozszerzona robota tylko un się załapał, o.
'Y Magda dijo que llamó, que está trabajando allí, pero de momento no es un trabajo ampliado, solo él consiguió algo, Ø.'
- (6) No to oni tam się jakoś zostali szczęśliwie no to do nich te dzieci niektóre chodzili ja tam chodziłam też to takie tajne komplety się nazywały, o.
'Pues ellos se quedaron allí de alguna manera, con suerte, y algunos niños iban a su casa. Yo también iba. Se llamaban *tajne komplety*, Ø.'

En todos estos casos, *o* aparece como cierre de un segmento discursivo, con función enfática o de cierre argumentativo. Esta colocación final es perfectamente natural en polaco, pero no lo sería en español, donde la interjección pospuesta tiende a parecer arcaica, teatral o ajena al uso cotidiano.

[33] Así, los ejemplos (1) y (2) anteriores podrían presentarse de la siguiente forma:

- (7) O, ja mówię to ja się już nawet cieszę no bo przecież to są ludzie tacy wiesz milionerami można ich nazwać, o.
'Oh, yo digo que hasta me alegro, porque, claro, son gente que, ya sabes, se podría llamar millonarios, Ø.'
- (8) Coraz więcej ludzi nie ufa politykom. Wystarczy posłuchać, co mówią moi sąsiedzi, o.
'Cada vez más gente desconfía de los políticos. Basta con oír lo que dicen mis vecinos, Ø.'

[34] Desde una perspectiva pragmática, el contraste es aún más visible en la manera en que ambas lenguas gestionan la ejemplificación y la validación del discurso. En polaco, *o* se integra con fluidez para señalar la satisfacción del hablante ante la pertinencia de lo dicho. En español, en cambio, *oh* no desempeña esta fun-

ción, y su inclusión en tales contextos resultaría marcada o poco idiomática.

[35] Este desajuste se refleja con frecuencia en la traducción, donde la interjección *o* desaparece sin merma comunicativa, siendo sustituida en ocasiones por gestos, entonación o acotaciones contextuales. Así, frases como *O! A le jak znajde jajko szpaka ... tez będzie fajnie* 'Ø. Pero si encuentro un huevo de estornino ... también me sirve' no requieren una interjección equivalente en español, ya que el contexto y la intención comunicativa se recuperan por otros medios.

[36] En conclusión, la interjección *o* en polaco presenta un valor específico de refuerzo discursivo, ejemplificación validante y satisfacción epistémica que no tiene correspondencia directa en *oh* en español. Además, su distribución sintáctica más flexible y su integración natural en el flujo conversacional representan un reto para la producción adecuada en español por parte de aprendientes polacos. Estas divergencias refuerzan la necesidad de abordar el estudio pragmático de las interjecciones desde una perspectiva contrastiva, tanto en la investigación como en el ámbito de la enseñanza de ELE.

3.3 Consideraciones sintácticas y prosódicas

[37] Desde el punto de vista sintáctico, ambas interjecciones tienden a ocupar posiciones periféricas: encabezando el enunciado, intercaladas o como cláusulas independientes. Sin embargo, la interjección *o* en polaco se integra con mayor frecuencia en estructuras complejas, mientras que *oh* en español tiende a mantenerse aislada.

[38] Prosódicamente, tanto *oh* como *o* se asocian con entonaciones marcadas y pausas delimitadoras. En polaco, es habitual una entonación descendente o suspendida, que anuncia la introducción de un contenido relevante. En español, *oh* suele ir acompañada de una entonación ascendente-descendente típica de las exclamaciones emocionales.

[39] Estas diferencias sugieren que *o* ha desarrollado en polaco un uso más discursivo y menos exclusivamente afectivo, mientras que *oh* en español mantiene una mayor dependencia de la carga emocional y de la prosodia para su interpretación.

4 Corpus y análisis pragmático: potenciales y limitaciones

[40] El análisis de unidades lingüísticas con alto componente contextual y afectivo, como las interjecciones, plantea numerosos desafíos metodológicos cuando se realiza a partir de corpus. Aunque los corpus constituyen una herramienta imprescindible para estudiar el uso real del lenguaje, su aplicabilidad al análisis pragmático, y especialmente al de fenómenos orales espontáneos, está condicionada por factores como el tipo de datos incluidos, el grado de transcripción, el etiquetado y la representatividad comunicativa. Este apartado examina dichas limitaciones a partir del caso concreto de las interjecciones *oh* y *o*.

4.1 Representatividad de las interjecciones en corpus escritos vs. orales

[41] Los corpus escritos tienden a infrarrepresentar fenómenos propios de la interacción espontánea, como las interjecciones, los marcadores discursivos o las estrategias de cortesía. Aunque algunas de estas formas aparecen también en textos literarios o periodísticos, su frecuencia y función no son equivalentes a las del habla real.

[42] Esto es especialmente evidente en el caso de *oh*, cuya presencia en corpus escritos se asocia con usos estilizados, poéticos o literarios, como exclamaciones enfáticas o marcas de arcaísmo, pero no con los usos empáticos, dubitativos o de escucha activa que caracterizan su función pragmática en la conversación.

[43] Los corpus orales, por su parte, ofrecen un acceso más directo al uso espontáneo de *oh* y *o*, pero presentan limitaciones en cuanto a extensión, equilibrio y anotación pragmática. Muchos corpus orales de acceso público incluyen grabaciones didácticas o entrevistas semidirigidas, lo que restringe la aparición de formas no planificadas o reactivas.

4.2 Ejemplos contrastivos: *CAES*, *Corpus oral de ELE*, *Spokes*

[44] Los corpus consultados confirman esta asimetría. En el *CAES*, orientado a ELE, la interjección *oh* aparece en general en muy contadas ocasiones, y casi exclusivamente en textos escritos o tareas con una carga emocional explícita.

Palab. ortográficas

oh

Volver Limpiar Buscar

Nivel de español			L1				
	Resultados	Estudiantes	Frec. norm.		Resultados	Estudiantes	Frec. norm.
A1	1 / 175 502	1 / 729	6 / millón	Alemán	0 / 24 466	0 / 53	0 / millón
A2	2 / 249 069	2 / 705	8 / millón	Árabe	1 / 152 691	1 / 502	7 / millón
B1	2 / 211 613	2 / 538	9 / millón	Chino mandarín	1 / 93 998	1 / 230	11 / millón
B2	0 / 177 053	0 / 384	0 / millón	Francés	0 / 84 841	0 / 243	0 / millón
C1	0 / 115 155	0 / 188	0 / millón	Griego	0 / 14 286	0 / 34	0 / millón
Sexo			Inglés	0 / 128 790	0 / 300	0 / millón	
	Resultados	Estudiantes	Frec. norm.	Italiano	0 / 35 035	0 / 87	0 / millón
Hombre	3 / 324 211	3 / 917	9 / millón	Japonés	2 / 105 636	2 / 419	19 / millón
Mujer	2 / 604 181	2 / 1627	3 / millón	Polaco	0 / 68 552	0 / 121	0 / millón
				Portugués	1 / 161 737	1 / 388	6 / millón
				Ruso	0 / 58 360	0 / 167	0 / millón

Figura 1: Resultados de búsqueda de *oh* en *CAES*

[45] De hecho, en polaco como L1 no aparece en ninguna ocasión, como se observa en la siguiente imagen:

The screenshot shows the CAES search interface. At the top, there is a navigation bar with 'CAES', 'Datos', 'Búsqueda', 'Listados', 'Ayuda', and 'Contacto'. The main search area is divided into 'Búsqueda' and 'Resultado'. Under 'Búsqueda', there is a 'Tipo' dropdown set to 'Palab. ortográficas'. Under 'Resultado', there are dropdowns for 'Tipo de resultado' (set to 'Frecuencias simples'), 'Ordenación', 'Tamaño de página' (set to '50'), and 'Agrupación'. Below these are 'Filtros' including 'Sensibilidad' (with checkboxes for 'Acentos' and 'Mayúsculas'), 'Nivel de español' (set to 'Cualquiera'), 'L1' (set to 'Polaco'), 'Pais' (set to 'Cualquiera'), 'Sexo' (set to 'Cualquiera'), 'Edad desde' (set to 'Cualquiera'), 'Edad hasta' (set to 'Cualquiera'), 'Tipología textual' (set to 'Cualquiera'), 'Tema' (set to 'Cualquiera'), and 'Estudiante'. The search input field contains 'oh'. At the bottom right, there are buttons for 'Volver', 'Limpiar', and 'Buscar'. A light blue message box at the bottom states 'No se encontraron resultados.'

Figura 2: Ausencia de *oh* con L1 polaco en el CAES

[46] Lo mismo ocurre en el *Corpus oral de ELE*, donde *oh* apenas figura entre las interjecciones empleadas por los hablantes no nativos y aparece solo en tres ocasiones en los más de 90 minutos de grabación de las cuatro entrevistas realizadas (dos con aprendices de nivel A2 y dos con aprendices de nivel B1).

Corpus Oral de Español como Lengua Extranjera (ELE)

Información Entrevistas Errores Búsqueda Ayuda

Lema de la(s) palabra(s): oh Categoría: interjección Lengua materna del estudiante: Polaco

Resultados de *oh* (interjección) en estudiantes cuya lengua materna es el polaco:

ENT: no hay < sitio libre> ...
 ADZ: [<] < & s = iah! ioh!> no@g hay@g sitios@g < & eh libres@g //

oh

ENT: [<] < hhh (%act: assent)> ///
 ADZ: ~ en este lugar ...

ENT: ¿ por qué no vamos a → la hamburguesería < o / qué> +
 ADZ: [<] < por = ioh! sí / sí / por qué & n [/] no [/] no vamos a < cervecería (%pho: [seβese'ria]) / sí eso xxx> ///

oh

ENT: [<] < hhh (%act: assent) / ¿ qué te parece si> vamos a una hamburguesería ?
 ADZ: iah! es = ioh! & eh / yo también xxx he visto hhh (%act: laugh) ///

oh

Figura 3: Ejemplo de resultados parciales de *oh* sin filtrar por L1 en *Corpus oral de ELE*

[47] Por el contrario, en el corpus *Spokes*¹ del polaco oral contemporáneo, la forma *o* aparece con una frecuencia considerable, aunque esta cifra debe interpretarse con cautela, ya que incluye usos no interjectivos, principalmente como conjunción disyuntiva. En una búsqueda simple por forma ortográfica, *o* se registra en más de un 6% de los enunciados del corpus, lo que indica una amplia presencia en el habla espontánea. No obstante, dado que la interjección *o* no está etiquetada de forma diferenciada, es decir, no se distingue sistemáticamente de sus homónimos gramaticales mediante categorías como *qub* o *intj*, la identificación de sus usos interjectivos exige una inspección cualitativa del contexto.

[48] Una estrategia complementaria para delimitar sus funciones pragmáticas consiste en analizar las coocurrencias léxicas inmediatas. Los datos obtenidos muestran una serie de combinaciones altamente frecuentes en las que *o* precede expresiones inequívocamente interjectivas o expresivas, como *o jezu* (22 casos), *o*

¹ Las búsquedas se realizaron en el portal *Spokes* mediante inspección manual y consulta por forma ortográfica, sin aplicar filtros por categorías gramaticales, ya que la forma *o* no aparece claramente anotada como interjección en el corpus. La atribución de valor interjectivo se basó en el análisis contextual de las coocurrencias léxicas, el orden de aparición y el comportamiento prosódico presunto (por ejemplo, aparición aislada, uso exclamativo o combinación con marcadores discursivos, nombres propios religiosos o expresiones emocionales típicas).

jezus (16), *o matko* (14), *o jejku* (9), *o ja* (16), *o nie* (15), *o właśnie* (3) u *o teraz* (3). También es significativa su aparición ante verbos de percepción o cognición en primera o segunda persona, como en *wiesz, o* (15), *wiem, o* (7), *wiedzieli, o* (2) o *wiedziala, o* (1), donde *o* introduce una valoración emocional o una reacción empática. Finalmente, los contextos en los que *o* va seguido de signos de puntuación enfáticos (*o!* = 19 casos) o partículas discursivas como *i* (*o i* = 9 casos) refuerzan su interpretación como interjección.

[49] En conjunto, estos datos ilustran que, aunque el etiquetado del corpus no permita aislar automáticamente todos los usos interjectivos de *o*, sí es posible recuperar un número significativo de ocurrencias interpretables como tales mediante un análisis combinado de coocurrencias léxicas y entorno sintáctico. Esta evidencia refuerza la hipótesis de que *o* goza de una alta integración en la dinámica emocional y discursiva del polaco oral, en contraste con la escasa presencia y funcionalidad de *oh* en los corpus de ELE en español. Sin embargo, esta disparidad cuantitativa no puede interpretarse únicamente como una diferencia lingüística entre ambas lenguas, sino también como una limitación estructural de los corpus disponibles en español, donde la interacción espontánea y los fenómenos no normativos suelen estar subrepresentados.

4.3 Riesgos de generalización desde corpus desequilibrados

[50] Trabajar con corpus que no reflejan adecuadamente los usos conversacionales auténticos puede inducir a errores de interpretación, especialmente en el ámbito de la pragmática. Si un corpus no incluye suficientes muestras de habla espontánea, las interjecciones como *oh* aparecen como excepcionales, arcaicas o puramente literarias, lo que distorsiona su valor comunicativo real.

[51] Conviene subrayar que la crítica a los corpus que aquí se plantea no pretende generalizar de manera absoluta a partir del estudio de únicamente dos interjecciones. Más bien, se trata de un caso paradigmático que ilustra limitaciones ya señaladas en la bibliografía sobre corpus (Cheng 2010; McEnery & Xiao 2011; Barron 2020). En efecto, fenómenos pragmáticos como los marcadores discursivos (*bueno, pues, mira*), las expresiones de cortesía (*gracias, por favor*), o incluso elementos prosódicos (*mm, ajá*), presentan patrones de infrarrepresentación similares en muchos corpus generales. Por ello, *oh* y *o* constituyen un ejemplo específico de un problema más amplio: la dificultad de los corpus actuales para reflejar con fidelidad unidades altamente sensibles al contexto y a la interacción.

[52] Como advierten McEnery & Xiao (2011), muchos corpus fueron diseñados con fines lexicográficos o gramaticales, y no contemplan los parámetros discursivos que interesan a la lingüística pragmática o a la didáctica comunicativa. Este desfase entre tipo de corpus y objetivo de análisis exige una lectura crítica de los resultados y la consideración de datos complementarios.

4.4 El papel del *data-driven learning* y la conciencia pragmática

[53] Desde una perspectiva aplicada, estas limitaciones inciden también en

el uso didáctico de los corpus. En este sentido, resulta pertinente mencionar el enfoque de *data-driven learning*, entendido como una metodología de enseñanza que coloca al aprendiente en contacto directo con datos lingüísticos auténticos extraídos de corpus, de modo que pueda descubrir patrones de uso mediante la observación, la inducción y la reflexión (Johns 1991; Boulton 2010). Este enfoque, relativamente novedoso en la enseñanza de lenguas, se fundamenta en la idea de que el análisis autónomo de ejemplos reales promueve una conciencia metalingüística y pragmática más sólida que la exposición exclusiva a materiales elaborados. Tal como señala Cheng (2010), el *data-driven learning* puede resultar muy eficaz para sensibilizar a los estudiantes sobre patrones reales del uso lingüístico, pero su efectividad depende de que los datos sean representativos de la lengua hablada y de sus funciones pragmáticas.

[54] En el caso de las interjecciones, su interpretación depende no solo de la forma lingüística, sino también de factores como la prosodia, la gestualidad, la dinámica conversacional y el conocimiento compartido entre interlocutores. Sin un etiquetado adecuado, prosódico, situacional o discursivo, el corpus pierde gran parte de su utilidad como herramienta para el desarrollo de la competencia pragmática.

[55] Esta reflexión enlaza con propuestas recientes sobre la creación de corpus pedagógicos específicamente diseñados para la enseñanza de la competencia discursiva, con anotaciones que incluyan funciones comunicativas, actos de habla, marcas prosódicas y tipos de interacción (McEnery & Xiao 2011). En este marco, las interjecciones pueden ser tratadas no como excentricidades gramaticales, sino como estrategias comunicativas centrales.

5 Hacia corpus más funcionales para la investigación y la didáctica

[56] Los resultados del análisis precedente subrayan la necesidad de avanzar hacia corpus que reflejen con mayor fidelidad la riqueza pragmática del discurso oral. Las interjecciones como *oh* y *o*, altamente sensibles al contexto y a la modalidad de producción, son un ejemplo paradigmático de las limitaciones de muchos corpus actuales. En este apartado se abordan propuestas concretas para el diseño de corpus más adecuados desde el punto de vista funcional, investigativo y pedagógico.

5.1 Criterios para la construcción de corpus pragmáticamente fiables

[57] Para que un corpus resulte útil en el estudio de fenómenos pragmáticos, este trabajo propone una serie de criterios mínimos, formulados a partir de las limitaciones observadas en la interjección *oh* en los corpus de ELE (*CAES, Corpus oral de ELE*) y en la interjección *o* en el corpus oral polaco *Spokes*, y en consonancia con planteamientos previos de McEnery & Xiao (2011) y Barron (2020). Dichos criterios incluyen:

- Modalidad oral auténtica: priorizar interacciones espontáneas y no inducidas, con el fin de recoger fenómenos como interjecciones y marcadores dis-

- cursivos en su contexto natural.
- Diversidad de géneros discursivos: incorporar conversaciones informales, narraciones, discusiones y llamadas telefónicas, en las que el uso interjetivo se manifiesta con mayor densidad funcional.
- Anotación contextual: especificar relación entre interlocutores, finalidad de la interacción y marco situacional, elementos imprescindibles para interpretar valores pragmáticos.
- Transcripción detallada y multilínea: reflejar segmentación prosódica, pausas, solapamientos y emisiones incompletas, que resultan fundamentales para la interpretación de interjecciones.

[58] La pertinencia de estos criterios se confirma en los datos analizados. Así, en el corpus *Spokes*, la ocurrencia *O, jakie to piękne!* '¡Oh, qué bonito!' aparece como reacción expresiva en un contexto espontáneo, pero no está anotada como interjección emotiva, lo que dificulta su recuperación automática. En el *CAES*, en cambio, se documenta la secuencia escrita *¡Oh, qué sorpresa!*, sin anotación prosódica ni situacional, lo que impide identificarla como marcador empático. Estos ejemplos muestran cómo la ausencia de información pragmática y contextual reduce el valor analítico y didáctico de los corpus.

5.2 Corpus pedagógicos: propuestas desde la literatura reciente

[59] La literatura sobre enseñanza de lenguas ha insistido en la necesidad de crear *pedagogical corpora* específicamente diseñados para desarrollar la competencia comunicativa, tanto receptiva como productiva. Según Barron (2020), un corpus con fines didácticos debería incorporar etiquetas pragmáticas (acto de habla, función comunicativa, tono emocional), metadatos contextuales accesibles al estudiante, ejemplos breves y contextualizados, y anotaciones graduadas por niveles de competencia.

[60] Un ejemplo de este tipo es el corpus *ESLORA* (Vázquez Rozas & Blanco 2023), que integra anotaciones pragmáticas detalladas y metadatos contextuales. No obstante, se trata de un corpus de hablantes nativos, por lo que no refleja la producción de aprendientes de ELE. En este sentido, la incorporación de interjecciones como *oh* u *o* en corpus pedagógicos contrastivos permitiría tratarlas no como excentricidades gramaticales, sino como estrategias comunicativas centrales, ilustrando con datos reales cómo se expresan sorpresa, refuerzo discursivo o implicación afectiva en diferentes lenguas.

[61] Por otra parte, Cheng (2010) y McEnery & Xiao (2011) insisten en que los corpus diseñados exclusivamente con fines lexicográficos no son necesariamente adecuados para el aula, especialmente en lo que respecta al componente pragmático. En este sentido, corpus como el *CAES* o el *Corpus oral de ELE*, aunque útiles en otros ámbitos, no resultan suficientes para trabajar interjecciones como *oh* desde una perspectiva comunicativa.

5.3 La necesidad de etiquetado pragmático y metainformación contextual

[62] Una de las carencias más significativas de los corpus generales es la falta de anotación pragmática sistemática. El análisis de *oh* y *o* muestra cómo estas interjecciones, pese a su alta frecuencia en la interacción oral, quedan diluidas en etiquetas morfosintácticas generales, invisibilizando su valor expresivo. Para superar esta limitación, es necesario contar con etiquetas que indiquen:

- Función ilocutiva (evaluativa, empática, exhortativa);
- Orientación (hacia el oyente, hacia el discurso, hacia el contenido);
- Grado de implicación afectiva;
- Posición enunciativa (inicio, inserción, cierre);
- Correlatos prosódicos relevantes.

[63] El etiquetado pragmático permitiría no solo afinar el análisis cuantitativo y cualitativo, sino también facilitar la explotación didáctica de los corpus en entornos de *data-driven learning*. Así, un estudiante de ELE podría observar, por ejemplo, cómo *o* en polaco introduce un ejemplo con valor de refuerzo discursivo, mientras que en español este efecto suele lograrse mediante interjecciones alternativas como *bueno* o *vaya*. De este modo, la inclusión de interjecciones en corpus pedagógicos permitiría tratarlas no como meros adornos literarios, sino como componentes pragmáticos esenciales de la competencia comunicativa.

5.4 Proyectos emergentes y líneas futuras

[64] El desarrollo de corpus equilibrados, anotados y accesibles para el estudio pragmático es una tarea aún en construcción. Proyectos recientes como el *Spokes* en Polonia demuestran que es posible integrar interacciones reales, metadatos contextuales y herramientas de búsqueda sofisticadas. Por su parte, el proyecto *Visegrad corpus of Spanish as a foreign language* (en preparación), en el que participan investigadores de varios países centroeuropeos, plantea, en su fase inicial, la posibilidad de recoger datos orales espontáneos, lo que podría abrir una vía para subsanar algunas de las carencias señaladas en corpus como el *CAES* o el *Corpus oral de ELE*.

[65] Además, iniciativas de menor escala, como la compilación de minicorpus especializados en fenómenos pragmáticos concretos, podrían ofrecer soluciones viables y más inmediatas. En el caso de las interjecciones, bastaría con reunir una muestra representativa de fragmentos de habla espontánea, debidamente segmentada y etiquetada, para facilitar tanto la investigación como su aplicación en el aula.

6 Implicaciones didácticas

[66] El estudio pragmático-contrastivo de interjecciones como *oh* en español y *o* en polaco no solo tiene valor teórico, sino que plantea consecuencias relevantes para la enseñanza de ELE, especialmente en contextos como el polaco, donde existen estructuras interjectivas funcionalmente similares. La comprensión y producción adecuada de estas unidades depende no solo del dominio léxico, sino tam-

bién de la conciencia pragmática, de la sensibilidad al contexto y de la exposición a datos reales. A continuación se exponen algunas líneas de acción didáctica derivadas del presente análisis.

6.1 La interjección como componente de la competencia pragmática

[67] A pesar de su brevedad formal, las interjecciones desempeñan funciones clave en la interacción, como expresar emociones, manifestar empatía, gestionar turnos o reforzar la cortesía. Ignorar su enseñanza conduce a una competencia comunicativa incompleta, en la que el aprendiz puede dominar estructuras complejas, pero carecer de recursos básicos para reaccionar de forma natural y adecuada en contextos cotidianos.

[68] En este sentido, las interjecciones no deben tratarse como elementos ornamentales o literarios, sino como actos de habla autónomos, con un valor comunicativo pleno. Su incorporación en el aula de ELE contribuiría a mejorar la fluidez, la expresividad y la naturalidad de los aprendientes. En esta línea, Edeso Natalías (2005, 2007, 2009) ha defendido reiteradamente la necesidad de integrar las interjecciones como componentes funcionales del discurso oral en contextos de ELE.

6.2 Integración de corpus en la enseñanza del componente pragmático

[69] El uso didáctico de corpus permite exponer a los estudiantes a ejemplos auténticos y variados, pero requiere una selección crítica del material. En el caso de las interjecciones, los corpus tradicionales ofrecen muy pocos ejemplos contextualizados y orales, por lo que se recomienda:

- Utilizar fragmentos de corpus orales anotados (cuando estén disponibles);
- Crear corpus propios a partir de transcripciones de materiales audiovisuales realistas (series, documentales, grabaciones espontáneas);
- Seleccionar ejemplos en los que la interjección tenga una función comunicativa clara y contextualizada.

Este enfoque favorece el aprendizaje inductivo de las interjecciones, su clasificación funcional y la comparación con estructuras equivalentes en la lengua materna.

6.3 Sensibilización contrastiva y prevención de interferencias

[70] La existencia de formas paralelas en la lengua materna, como *o* en polaco, puede ser una ventaja para el aprendizaje de *oh*, pero también un factor de interferencia. El contraste permite identificar semejanzas, pero también divergencias funcionales, prosódicas y discursivas que deben tratarse explícitamente en clase.

[71] Por ejemplo, mientras que en polaco *o* puede introducir un nuevo tópico narrativo con naturalidad, en español el uso de *oh* con la misma función resultaría forzado o literario. Esta diferencia debe ser tematizada mediante ejemplos reales, análisis guiado y ejercicios de transformación contextual. Como advierte

Krzempek (2016), la enseñanza del polaco como L2 suele tratar las interjecciones de forma periférica o descontextualizada, lo que complica su transferencia pragmática a otras lenguas.

6.4 Formación docente y conciencia metapragmática

[72] Para incorporar con eficacia las interjecciones en la enseñanza, es necesario que el profesorado de ELE desarrolle una conciencia metapragmática, entendida como la capacidad de reconocer, analizar y explicar el funcionamiento pragmático de los recursos lingüísticos en distintos contextos. En este caso, implica que el docente pueda identificar cuándo una interjección cumple una función expresiva, apelativa o discursiva, y cómo su uso depende de factores situacionales, prosódicos y culturales. Esto incluye:

- Familiaridad con los corpus y sus limitaciones;
- Conocimiento de las funciones discursivas de las interjecciones;
- Capacidad para generar ejemplos contextualizados y didácticamente pertinentes.

[73] Esta conciencia metapragmática supone ir más allá del conocimiento gramatical: requiere la habilidad de explicar a los aprendientes no solo qué significa una interjección, sino también cómo y cuándo se usa en la interacción real. De este modo, el profesorado puede generar ejemplos contextualizados, anticipar posibles interferencias con la lengua materna y guiar a los estudiantes en la comparación contrastiva. En esta línea, Barron (2020) subraya que el trabajo con corpus pragmáticamente anotados permite al profesorado y al alumnado observar no solo los actos de habla, sino también los matices expresivos, interculturales y situacionales, constituyendo así una base sólida para el desarrollo de esta conciencia metapragmática.

7 Conclusiones

[74] El presente trabajo ha abordado el análisis pragmático-contrastivo de las interjecciones *oh* en español y *o* en polaco desde una doble perspectiva: funcional y metodológica. Por un lado, se han examinado los valores expresivos y discursivos de ambas formas en corpus escritos y orales; por otro, se ha reflexionado críticamente sobre los límites actuales de los corpus disponibles para este tipo de análisis, en especial cuando se trata de fenómenos estrechamente ligados a la interacción oral y la subjetividad del hablante.

[75] Desde el punto de vista funcional, *oh* y *o* comparten un núcleo de valores expresivos (sorpresa, admiración, pena, empatía, etc.) que permite postular una cierta equivalencia pragmática básica. Sin embargo, también se han identificado diferencias significativas: *o* en polaco presenta una mayor variedad de usos discursivos. Así, aparece con frecuencia en contextos narrativos, como recurso de tematización o apertura tópica, mientras que *oh* en español mantiene un perfil más marcadamente emocional y literario. Estas divergencias se reflejan tanto en la dis-

tribución sintáctica como en la entonación y en el grado de gramaticalización discursiva.

[76] Desde el punto de vista metodológico, el artículo ha mostrado cómo las limitaciones de los corpus actuales, especialmente en el caso de ELE, afectan a la posibilidad de describir de forma adecuada unidades pragmáticas como las interjecciones. La escasa presencia de *oh* en corpus como *CAES* o el *Corpus oral de ELE* no refleja necesariamente su marginalidad funcional, sino las carencias estructurales de esos recursos, que no fueron concebidos para el análisis discursivo o la enseñanza de la competencia comunicativa.

[77] En consecuencia, se propone avanzar hacia el diseño de corpus más representativos, anotados y orientados a la pragmática, tanto para la investigación como para la enseñanza. Esto implica incluir datos orales espontáneos, mejorar la anotación funcional y prosódica, y desarrollar subcorpus específicos para fenómenos discursivos poco representados. Asimismo, se ha destacado el valor del enfoque *data-driven learning* y de los corpus pedagógicos como herramientas eficaces para la enseñanza de interjecciones y otros elementos pragmáticos en contextos de ELE.

[78] Finalmente, se subraya la necesidad de una formación docente sólida en análisis pragmático y uso de corpus, que permita al profesorado identificar, interpretar y enseñar fenómenos como *oh* y *o* de manera contextualizada y contrastiva. El estudio de estas formas, aparentemente marginales, ofrece una vía privilegiada para explorar la dimensión afectiva, interpersonal y discursiva del lenguaje, tanto en la investigación como en la práctica didáctica.

Abreviaturas y referencias bibliográficas

- Alarcos Llorach 1994 = Emilio Alarcos Llorach 1994. *Gramática de la lengua española*. Espasa Calpe.
- Ameka 1992 = Felix Ameka 1992. Interjections: The universal yet neglected part of speech. *Journal of Pragmatics* 18.2-3, 101-118.
- Bańko 2001 = Mirosław Bańko 2001. *Z pogranicza leksykografii i językoznawstwa. Studia o słowniku jednojęzycznym*. Nakł. Wydziału Polonistyki Uniw. Warszawskiego.
- Baran 2010 = Marek Baran 2010. *Emotividad y convención sociopragmática. Una contribución al estudio del ethos comunicativo de la comunidad hispanohablante peninsular*. Wydawnictwo Uniwersytetu Łódzkiego.
- Barron 2020 = Anne Barron 2020. Teaching pragmatic competence with corpora: Intensification in expressions of gratitude across varieties. Holger Limberg, Karen Glaser (eds.). *Pragmatische Kompetenzen im schulischen Fremdsprachenunterricht*. Lang, 105-137.
- Blakemore 2002 = Diane Blakemore 2002. *Relevance and linguistic meaning. The semantics and pragmatics of discourse markers*. Cambridge University Press.
- Boulton 2010 = Alex Boulton 2010. Data-driven learning: Taking the computer out of the equation. *Language Learning* 60.3, 534-572.
- CAES = Guillermo Rojo, Ignacio Palacios (eds.) 2013-. *Corpus de aprendices de español*. <https://galvan.usc.es/caes>.
- Cheng 2010 = Winnie Cheng 2010. What can a corpus tell us about language teaching? Anne O'Keeffe, Michael J. McCarthy (eds.). *The Routledge handbook of corpus linguistics*. Routledge, 319-332.
- Corpus oral de ELE* = Leonardo Campillos Llanos (ed.) 2013. *Corpus oral de español como lengua extranjera (ELE)*. http://cartago.illf.uam.es/exist/rest/db/corpus/home_es.html.
- Cuenca 2006 = María J. Cuenca 2006. *La connexió i els connectors. Perspectiva oracional i textual*. Eumo Editorial.
- Cueto Vallverdú & López Bobo 2003 = Natalia Cueto Vallverdú, María J. López Bobo 2003. *La interjección. Semántica y pragmática*. Arco Libros.
- Derwojedowa, Karaś & Kopcińska 2005 = Magdalena Derwojedowa, Halina Karaś & Dorota Kopcińska (eds.) 2005. *Język polski. Kompendium*. Świat Książki.
- DLE = Real Academia Española 2014. *Diccionario de la lengua española*. 23a edición. <https://dle.rae.es/>.
- Dubisz 2003 = Stanisław Dubisz (ed.) 2003. *Uniwersalny słownik języka polskiego*. Wydawnictwo Naukowe PWN.
- DUE = María Moliner 1998. *Diccionario de uso del español*. Gredos.
- Edeso Natalías 2005 = Verónica Edeso Natalías 2005. Las interjecciones en la enseñanza de ELE: Significados expresivos de la interjección *oh*. María A. Castillo Carballo, Olga Cruz Moya, Juan M. García Platero, Juan P. Mora Gutiérrez (eds.). *Las gramáticas y los diccionarios en la enseñanza del español como segunda lengua. Deseo y realidad. Actas del XV Congreso internacional de ASELE. Sevilla, 22-25 de septiembre de 2004*. Editorial Universidad de Sevilla, 326-333. <https://dialnet.unirioja.es/descarga/articulo/1421285.pdf>.
- Edeso Natalías 2007 = Verónica Edeso Natalías 2007. ¿Qué hacer con las interjecciones en una clase de ELE? Enrique Balmaseda Maestu (ed.). *Las destrezas orales en la enseñanza del español L2-LE. XVII Congreso internacional de la Asociación del español como lengua extranjera (ASELE). Logroño, 27-30 de septiembre de 2006*. Universidad de La Rioja, 495-508. <https://dialnet.unirioja.es/descarga/articulo/2450508.pdf>.
- Edeso Natalías 2009 = Verónica Edeso Natalías 2009. *Contribución al estudio de la interjección en español*. Lang.

- ELE = Español como lengua extranjera.
- ESLORA = Victoria Vázquez Rozas (ed.) 2007-. *Corpus para el estudio del español oral*. <http://es-lora.usc.es>.
- Grochowski 1991 = Maciej Grochowski 1991. Status semantyczny wykrzykników właściwych. *Prace filologiczne* 37, 155-163.
- Johns 1991 = Timothy F. Johns 1991. Should you be persuaded? Two samples of data-driven learning materials. *English Language Research Journal* 4, 1-16.
- Krzempek 2016 = Monika Krzempek 2016. Ach, ten język polski! Wykrzykniki w podręcznikach do nauczania języka polskiego jako obcego. Violetta Próchniak, Magdalena Smoleń-Wawrzusin (eds.). *Nauczanie języka polskiego. Tradycje i innowacje*. Wydawnictwo Katolicki Uniwersytet Lubelski Jana Pawła II, 91-107.
- McEnery & Xiao 2011 = Tony McEnery, Richard Xiao 2011. What corpora can offer in language teaching and learning. Eli Hinkel (ed.). *Handbook of research in second language teaching and learning*. Vol. 2. Erlbaum, 364-380.
- Nagórko 2010 = Alicja Nagórko 2010. *Podręczna gramatyka języka polskiego*. Wydawnictwo Naukowe PWN.
- NGLE = Real Academia Española, Asociación de Academias de la Lengua Española 2009-2011. *Nueva gramática de la lengua española*. Espasa.
- Pawlak 2017 = Aneta Pawlak 2017. La interjección. Wiaczesław Nowikow (ed.). *Gramática contrastiva español-polaco*. Wydawnictwo Uniwersytetu Łódzkiego, 671-702.
- Polański 2000 = Kazimierz Polański (ed.) 2000. *Encyklopedia języka polskiego*. 2a edycja. Wydawnictwo Ossolineum.
- PWN = Mieczysław Szymczak (ed.) 1988-. *Słownik języka polskiego PWN*. Wydawnictwo Naukowe PWN. <https://sjp.pwn.pl/>.
- Seco, Andrés & Ramos 1999 = Manuel Seco, Olimpia Andrés, Gabino Ramos (eds.) 1999. *Diccionario del español actual*. Aguilar.
- Spokes = Polish and English language corpora for research and applications (PELCRA) (ed.) 2024-2025. *Spokes conversational data search*. <https://spokes.clarin-pl.eu/>.
- Vázquez Rozas & Blanco 2023 = Victoria Vázquez Rozas, Marta Blanco 2023. El corpus ESLORA de español oral: un recurso para la investigación en ELE. *TEISEL. Tecnologías para la investigación en segundas lenguas* 2, 1-25. <https://doi.org/10.1344/teisel.v2.41757>.
- Wicherek & Balches Arenas 2012 = Marta Wicherek, R. Sergio Balches Arenas 2012. Sobre el carácter polidiscursivo de la interjección polaca NO y su posible relación con actos de habla y conversacionales en polaco y en español: contextos pragmáticos de uso comparados en ambas lenguas. *Romanica Cracoviensia* 11.1, 457-471. <https://doi.org/10.4467/20843917RC.11.050.0118>.
- Wierzbicka 1992 = Anna Wierzbicka 1992. The semantics of interjection. *Journal of Pragmatics* 18.2-3, 159-192.
- Wierzbicka 1996 = Anna Wierzbicka 1996. Między modlitwą a przekleństwem: *O Jezu!* i podobne wyrażenia na tle porównawczym. *Etnolingwistyka. Problemy języka i kultury* 8, 25-39.
- Wilkins 1992 = David P. Wilkins 1992. Interjections as deictics. *Journal of Pragmatics* 18.2-3, 119-158.