



Estudios y aplicaciones de PRESEEA. Fonética, gramática y léxico

Studies and applications of PRESEEA. Phonetics, grammar, and lexicon

Sara Engra Minaya
Universidad de Alcalá

Elisa de las Fuentes Gutiérrez
Universidad de Alcalá

Florentino Paredes García
Universidad de Alcalá

Abstract

In this monograph titled Studies and Applications of PRESEEA: Phonetics, Grammar, and Lexicon, ten articles are presented containing research conducted within the framework of the PRESEEA macro-project and based on data collected from the project's corpora —although other sources may have been used occasionally—. The volume is divided into three sections. The first one is dedicated to phonetics and includes two works, one focusing on segmental units and the other on suprasegmental units. The second section focuses on aspects of grammatical variation and comprises five articles addressing topics such as periphrastic gerunds, pronominal subjects, clitic doubling, and periphrases. The final section contains three works on lexicon, one of them focusing on quantifying expressions, another on anglicisms, and a third on units with certain lexicographic marks. The data and results presented correspond to cities located on both shores: Colombia (Barranquilla, Cartagena, Valledupar, Cali, and Medellín), Cuba (Havana), Spain (Granada, Alcalá de Henares, Valencia, and Palma de Mallorca), and Mexico (Tijuana, Puebla, Mexicali, and Mexico City).

Keywords: PRESEEA, linguistic variation, sociolinguistics, phonetics, grammar, lexicon.

Resumen

En el presente monográfico, Estudios y aplicaciones de PRESEEA. Fonética, gramática y léxico, se ofrecen diez artículos que contienen investigaciones realizadas en el marco del macro-proyecto PRESEEA y a partir de datos correspondientes a los corpus recogidos en el seno del proyecto —si bien eventualmente pueden haberse usado otras fuentes—. El volumen se compone de tres bloques. El primero está dedicado a la fonética e incluye dos trabajos, uno sobre unidades segmentales y otros sobre unidades suprasegmentales. El segundo bloque se centra en aspectos de variación gramatical y se compone de cinco artículos, donde se abordan temas como el gerundio perifrástico, el sujeto pronominal, la duplicación de clíticos o la asperífrasis. El último bloque contiene tres trabajos sobre léxico, uno de ellos sobre expresiones cuantificadoras, otro sobre anglicismos y un tercero sobre unidades que tienen ciertas marcas lexicográficas. Los datos y resultados que se ofrecen corresponden a ciudades ubicadas en las dos orillas: Colombia (Barranquilla, Cartagena, Valledupar, Cali y Medellín), Cuba (La Habana), España (Granada, Alcalá de Henares, Valencia y Palma de Mallorca) y México (Tijuana, Puebla, Mexicali y Ciudad de México).

Palabras clave: PRESEEA, variación lingüística, sociolingüística, fonética, gramática, léxico.

Citar como: Engra Minaya, Sara; de las Fuentes Gutiérrez, Elisa; Paredes García, Florentino (2024). Estudios y aplicaciones de PRESEEA. Fonética, gramática y léxico. Normas, 14(1), 1-6, doi: 10.7203/Normas.v14i1.28640.

1. Introducción al monográfico

Uno de los principales objetivos de la investigación en sociolingüística es el estudio de la variación lingüística con objeto de determinar los factores lingüísticos y extralingüísticos que

condicionan los fenómenos variables o que operan sobre ellos en una determinada comunidad de habla. Las investigaciones sobre la lengua hablada se pueden realizar desde perspectivas teóricas y metodológicas diversas, pero los resultados más fiables son los que parten de datos obtenidos mediante una metodología rigurosa, compartida y homologable. De este modo, es posible observar y explicar la forma en que se configuran sociolingüísticamente las comunidades de habla y, a partir de ahí, comparar unas con otras y determinar las relaciones de convergencia y divergencia que se establecen entre las variedades geolectales y sociolectales de la lengua.

El *Proyecto para el estudio sociolingüístico del español de España y América* (PRESEEA) es un proyecto internacional que nació con el propósito de crear corpus de lengua hablada en distintas ciudades hispanohablantes de América y de España que permitiesen realizar investigaciones a partir de datos comparables y homologados. Uno de los fines que persigan los estudios realizados a partir de las investigaciones del PRESEEA ha de ser “analizar el desarrollo de la variación geolingüística y sociolingüística en el conjunto del dominio hispanohablante” (Moreno Fernández, 2021b: 7). En última instancia, las investigaciones, realizadas a partir de muestras sociológicas controladas y representativas de cada territorio, servirán para identificar los patrones sociolingüísticos y geolectales de las comunidades de habla junto a los procesos de convergencia y/o divergencia entre ellas y, en última instancia, permitirán extraer conclusiones teóricas y reformular teorías (Moreno Fernández, 2005, 2006, 2009, 2021a; Cestero Mancera, 2012; Moreno y Cestero, 2020). En la página electrónica del proyecto (<https://preseea.uah.es>) se han puesto a disposición de cualquier investigador interesado en la variación geolingüística y sociolingüística del español una serie de documentos de investigación, que sirven de ayuda para estudiar y conocer el estado de varios fenómenos lingüísticos en distintas zonas y el dinamismo interno de la lengua (véase la colección [*Documentos PRESEEA de investigación*](#))

Entre los hitos del PRESEEA, aparte de los corpus de más de 40 comunidades de habla urbanas hispanas, hay que señalar las investigaciones específicas que se han ido realizando desde finales del siglo XX, época en que nació el proyecto. En la actualidad se están llevando a cabo estudios coordinados sobre fenómenos lingüísticos variables de los distintos niveles lingüísticos, que han sido descritos en diversos puntos del español —por ejemplo, la *-s* en coda, el yeísmo, la expresión del sujeto pronominal, el leísmo, el diminutivo, los mecanismos de atenuación, los marcadores discursivos o el tabú—. Algunos de estos resultados comparados pueden consultarse en los monográficos de Cestero y Moreno (2017, 2023), Molina, Paredes y Cestero (2020) o el de Cestero y Núñez (2023), que complementa al que ahora ofrecemos.

En el presente monográfico, *Estudios y aplicaciones de PRESEEA. Fonética, gramática y léxico*, se ofrecen diez artículos que contienen investigaciones realizadas en el marco del macroproyecto PRESEEA y a partir de datos correspondientes a los corpus recogidos en el seno del proyecto —si bien eventualmente pueden haberse usado otras fuentes—. El volumen se compone de tres bloques. El primero está dedicado a la fonética e incluye dos trabajos, uno sobre unidades segmentales y otros sobre unidades suprasegmentales. El segundo bloque se centra en aspectos de variación gramatical y se compone de cinco artículos, donde se abordan temas como el gerundio perifrástico, el sujeto pronominal, la duplicación de clíticos o las perífrasis. El último bloque contiene tres trabajos sobre léxico, uno de ellos sobre expresiones cuantificadoras, otro sobre anglicismos y un tercero sobre unidades que tienen ciertas marcas lexicográficas. Los datos y resultados que se ofrecen corresponden a ciudades ubicadas en

las dos orillas: Colombia (Barranquilla, Cartagena, Valledupar, Cali y Medellín), Cuba (La Habana), España (Granada, Alcalá de Henares, Valencia y Palma de Mallorca) y México (Tijuana, Puebla, Mexicali y Ciudad de México).

El volumen se abre con «La frontera del habla “norteña” en México. Comparación entre los corpus PRESEEA de Tijuana y Mexicali». Los autores llevan a cabo un estudio para confirmar si la noción del subdialecto bajacaliforniano septentrional que distingue Juan Manuel Lope Blanch (1970) en su Atlas Lingüístico de México continúa en el presente. Para ello, emplean las entrevistas de los corpus de Mexicali y Tijuana y, de cada una de ellas, seleccionan las 10 primeras apariciones de los fonemas velar fricativo sordo, africado palatal sordo y alveolar fricativo sordo; analizan la realización de cada fonema por cada hablante; determinan la posición del fonema; y correlacionan las variables lingüísticas con la edad y el género de los participantes. Los resultados demuestran que los rasgos se encuentran presentes en la actualidad y parecen confirmar la noción del bajacaliforniano septentrional.

En su trabajo sobre «Relaciones entre prosodia y sintaxis: un estudio piloto sobre el uso de las pausas en el fraseo del habla de Granada», Cecilia González Medina presenta los primeros resultados de un análisis prosódico del español de Granada. Para su estudio, emplea la «Guía de estudios de prosodia basada en el uso en los corpus PRESEEA» de Velásquez Upegui y Martín Butragueño (2018) y, a partir de una muestra de doce informantes del corpus PRESEEA-Granada estudia qué variables sociales y lingüísticas (sintácticas y prosódicas) pueden influir en el uso de las pausas como potenciales delimitadoras de lindes prosódicos. Los resultados permiten concluir que la edad y la longitud del enunciado pueden afectar significativamente a la presencia o ausencia de pausas en el fraseo del habla granadina, y que existen indicios de que ciertas estructuras bivalentes resultan privilegiadas para la interpolación de pausa.

El boque dedicado a los estudios de variación gramatical se inicia con el artículo “El gerundio no perifrástico en el habla del español de La Habana”, de Ailyn Figueroa González y Milagros Alfonso Vega. Los autores analizan los usos de gerundio no perifrástico en entrevistas de PRESEEA-La Habana, incluyendo una construcción novedosa integrada por el verbo *ser* + gerundio. Para estudiar esta estructura en concreto, las autoras elaboraron un corpus ampliado con las *Muestras del habla culta de La Habana* y *Ameresco-La Habana*. El análisis aborda la función de la oración no finita, la clase de sujeto de los verbos involucrados, el aspecto léxico y gramatical y la estructura informativa. Los resultados demuestran la interconexión entre diversos fenómenos gramaticales. Concluyen que la oración de gerundio desempeña la función de atributo, que hay un alto grado de genericidad tanto en el sujeto del verbo copulativo como en el de la forma no personal y que la imperfectividad intrínseca del gerundio se refuerza sintácticamente: por este motivo, se propone que este gerundio atributivo se comporta como un focalizador aspectual.

La investigación de Rafael Orozco, Margareth Marmolejo y Róbinson Grajales sobre sujetos pronominales en el español andino colombiano consiste en un análisis preliminar de la expresión variable de este fenómeno lingüístico en el corpus PRESEEA-Cali, Colombia. Se exploran los efectos de cinco variables predictoras (2 sociales y 3 lingüísticas) en 4425 muestras: los resultados revelan la tasa pronominal (29,7%) más alta encontrada en una comunidad de habla latinoamericana continental. Se encuentra que las variables lingüísticas más poderosas son la persona y el número gramatical, especialmente es el pronombre *uno* el que más favorece los sujetos expresos. En cuanto a las variables sociales, se observa que son las mujeres quienes

promueven los sujetos expresos, tal como ocurre en la región Caribe, pero contrario a lo encontrado en Medellín. A su vez, pudo comprobarse que los jóvenes favorecen los sujetos tácitos y los mayores favorecen los sujetos expresos. Este hallazgo puede tener implicaciones cognitivas, puesto que se espera normalmente que los hablantes jóvenes tengan tasas pronominales más altas.

En su trabajo sobre “Factores sociolingüísticos determinantes en la posposición del pronombre personal sujeto del costeño”, Alder Luis Pérez Córdoba analiza la posposición del pronombre personal sujeto en el español del Caribe colombiano o costeño desde la perspectiva sociovariacionista. Para ello, se utilizó una muestra de 18 informantes por cada corpus sociolingüístico de las ciudades del Caribe colombiano (Barranquilla, Cartagena y Valledupar), recogidos en el marco del PRESEEA, y se codificaron 150 verbos finitos variables por hablante de acuerdo con la Guía del PRESEEA para el estudio de la variación de la ESP. Tras analizar estadísticamente los datos con Rbrul, los resultados confirman la coherencia interna del costeño y su inclusión en el español caribeño, pese a la ligera divergencia de Valledupar. Ciertas variables clásicas determinantes en la posposición resultan significativas, pero otras no, sugiriendo la necesidad de más análisis sobre este aspecto en el español general.

María del Mar Saeta Marcos investiga la duplicación de clíticos en función de complemento directo en el español de Granada (España). La presencia de la reduplicación facultativa del complemento directo se ha estudiado en el español de América y se explica mediante la asociación de la conjugación objetiva de las lenguas indígenas de frontera y la influencia de algunas variables lingüísticas. No obstante, en España este rasgo lingüístico no se ha analizado todavía en profundidad utilizando corpus lingüísticos que consideren variables sociales: de ahí la relevancia académica del presente trabajo. Saeta Marcos selecciona una muestra de casos reduplicados con complemento directo léxico pospuesto en el corpus PRESEEA-Granada, sin descartar léismos, para después cruzarlos estadísticamente con aquellos casos no duplicados en los que, presentando clítico, la concordancia objetiva era posible. Los resultados demuestran la relevancia de las variables lingüísticas, tales como el carácter no etimológico del complemento directo átono, el género del referente y determinados verbos, mientras que la variable social del nivel de estudios de los informantes necesita más investigaciones que confirmen su conexión con los resultados.

Cierra el bloque el artículo “Algunas contribuciones de los corpus de la red PRESEEA al estudio del español en contacto con el catalán”, de Andrés Enrique-Arias, que ejemplifica las posibilidades del empleo de los corpus de la Red PRESEEA para estudiar los efectos del contacto con el catalán en el castellano a partir de estudios que contrastan datos de Alcalá de Henares, Valencia y Palma de Mallorca, entre otras comunidades. A partir de la variación en el uso de los futuros (*cantaré* frente a *voy a cantar*) y en la distribución de las perífrasis de obligación *haber de* y *tener que*, Enrique-Arias lleva a cabo un análisis contrastivo que revela algunos efectos de la influencia del catalán en las comunidades bilingües. Los resultados muestran que, en los corpus de territorios catalanoparlantes, se observa una mayor retención del futuro morfológico *cantaré*, así como un menor grado de diferenciación de las funciones de modalidad deóntica y epistémica en la distribución de las perífrasis de obligación *haber de* y *tener que*.

El tercer y último bloque del volumen está dedicado al léxico. Juan José Rodríguez y Margareth Marmolejo Caicedo presentan en su trabajo una investigación exploratoria para analizar el uso de la expresión cuantificadora “un poco de” con los significados de *poca canti-*

dad y *mucha cantidad* en los corpus PRESEEA-Cali y PRESEEA-Medellín. Tras encontrar 20 ocurrencias en Cali y 14 en Medellín, analizan cualitativa y cuantitativamente cuatro variables lingüísticas (mucho/poca cantidad, nombres abstractos/concretos, nombres contables/no contables y nombres plurales/singulares) en relación con el sexo, la edad y el nivel educativo de los participantes. Los resultados muestran una mayor ocurrencia del uso no tradicional de la expresión cuantificadora en ambos corpus; revelan que en la mayoría de los casos en que “un poco de” denota mucha cantidad se acompaña de un nombre contable plural o colectivo y que cuando denota poca cantidad se emplea la fórmula “un poco de” + nombre singular abstracto; y permiten concluir que puede existir un mayor uso del significado de “mucha cantidad” en los hombres, la primera generación y el nivel educativo bajo, aunque “un poco de” con el significado de mucha cantidad está presente en todos los grupos sociales estudiados.

En «Aproximación al estudio de los anglicismos en dos comunidades de PRESEEA-México: Puebla y Mexicali», Erik Franco y Niktelol Palacios presentan un estudio comparado de los anglicismos léxicos encontrados en 36 entrevistas extraídas de los corpus mencionados. Para ello, analizan las características gramaticales y referenciales de esos anglicismos y contrastan su presencia y frecuencia de aparición en las entrevistas, partiendo de la idea de que es posible que Mexicali sea más anglicista por encontrarse cerca de la frontera de Estados Unidos. Los resultados permiten comprobar sus hipótesis y destacar la importancia de los préstamos en los nombres propios.

Por último, En este trabajo, los autores Montiel Pérez y Serrano Morales realizan un estudio de panel para determinar cambios léxico-semánticos en el español de la Ciudad de México. Para ello, empleando 19 entrevistas del corpus PRESEEA-Ciudad de México, analizan el vocabulario que suele ser marcado lexicográficamente en los diccionarios y llevan a cabo pruebas estadísticas a fin de relacionar las categorías léxicas *popular*, *grosero*, *tradición culta* y *préstamos* con la edad, el sexo y el nivel de instrucción de los participantes. Los resultados permiten concluir que algunas unidades léxicas mutan de manera significativa respecto a su correlación con rasgos sociales: estratificación por edad, cambio individual, nivelación sociolingüística y un cambio de actitud hacia determinadas categorías léxicas.

2. Referencias Bibliográficas

Cestero Mancera, A. M. (2012). El Proyecto para el estudio sociolingüístico del español de España y América. *Español Actual*, 98, 227-234.

Cestero Mancera, A. M., & Moreno Fernández, F. (2017). Procesos de variación y cambio en el español de España. Vol. *Monográfico XV. Lingüística en la Red*. http://www.linred.es/numero15_monografico.html

Cestero Mancera, A. M., & Moreno Fernández, F. (2023). Patrones sociolingüísticos y geolectales del español (Vol. 94). *Círculo de Lingüística Aplicada a la Comunicación*. <https://revistas.ucm.es/index.php/CLAC>

Cestero Mancera, A. M., & Núñez Bayo, Z. (2023). Estudios y aplicaciones de PRESEEA. sociopragmática y enseñanza de español. *Normas* Vol. 13/1.

Molina Martos, I., Paredes García, F., & Cestero Mancera, A. M. (2020). Sociolinguistic patterns and processes of convergence and divergence in Spanish. *Spanish in Context*, 17(2).

Moreno Fernández, F. (1996). Metodología del 'Proyecto para el estudio sociolingüístico del Español de España y de América' (PRESEEA). *Lingüística*, 8, 257-287.

Moreno Fernández, F. (2005). Corpus para el estudio del español en su variación geográfica y social: El corpus PRESEEA. *Oralia: Análisis del discurso oral*, 8, 123-140.

Moreno Fernández, F. (2006). Información básica sobre el "Proyecto para el estudio sociolingüístico del español de España y América"-PRESEEA (1996-2010). *Revista Española de Lingüística*, 36, 385-391.

Moreno Fernández, F. (2009). El estudio coordinado de las hablas hispánicas. En M. V. Camacho, J. J. Rodríguez, & J. Santana (Eds.), *Estudios de lengua española: Descripción, variación y uso: Homenaje a Humberto López Morales* (pp. 548-566). Iberoamericana.

Moreno Fernández, F. (2021a). Guía PRESEEA para la investigación lingüística. Proyecto PRESEEA/ Editorial Universidad de Alcalá.

Moreno Fernández, F. (2021b). Metodología del "Proyecto para el estudio sociolingüístico del español de España y de América (PRESEEA). Documentos PRESEEA de investigación, Documentos de trabajo, 1. <https://doi.org/10.37536/preseea.2021.doc1>

Moreno Fernández, F., & Cestero Mancera, A. M. (2020). El proyecto PRESEEA: desarrollos analíticos. *Verba: Anuario Galego de Filoloxía*, Anexo 80, 119-138.