

Una propuesta de texto fonológicamente equilibrado: *El ratón Arturo, la adaptación al español de Arthur the Rat*

Mario Casado-Mancebo

Universidad Nacional de Educación a Distancia (España) ✉ 

Nuria Polo Cano

Universidad Nacional de Educación a Distancia (España) ✉ 

Celia Teira Serrano

Universidad Pontificia de Salamanca (España) ✉ 

<https://dx.doi.org/10.5209/clac.81669>

Enviado: 25 de abril de 2022 • Aceptado: 20 de mayo de 2022

ES Resumen: Para recabar datos lingüísticos se suelen usar textos fonéticamente (fonológicamente) equilibrados. Una búsqueda de este tipo de textos en español descubre pocos textos con estas características. El objetivo del presente trabajo es proponer un texto nuevo, *El ratón Arturo*, a partir de la adaptación de un texto original en inglés. La principal motivación para confeccionar un texto nuevo es que los que existen, y de verdad están compuestos según la distribución de fonemas de esta lengua, son demasiado cortos. La lectura de los existentes es inferior a los dos minutos, lo que dificulta calcular medidas estadísticamente fiables para un análisis vocal. La ventaja de usar este texto frente a los existentes en español es que es más largo, supera los dos minutos, y presenta un menor número de repeticiones de palabras que los que ya existen. Además, está compuesto por palabras frecuentes y oraciones sencillas. Por lo que, a pesar de la longitud mayor, no es un texto difícil de leer o de procesar porque su complejidad es similar a la de los textos anteriores. Además, puede servir para recoger datos de población adulta e infantil. Cumple con los dos requisitos de este tipo de textos: presenta al menos una ocurrencia de cada uno de los fonemas del español y estos reproducen la frecuencia de distribución de esta lengua. Además, se han recogido ejemplos de los grupos consonánticos y las codas más frecuentes. Asimismo, se ha confeccionado con palabras simples y derivadas, con acento contrastivo, con oraciones con presencia/ausencia de consonantes nasales y con oraciones con diversas modalidades oracionales. Las características específicas con las que se ha construido permiten que pueda ser utilizado tanto en la investigación fonética como para evaluar problemas motores del habla, como disartrias y apraxias, y problemas de la función vocal.

Palabras clave: texto fonéticamente equilibrado; medida para la producción del habla y de la voz; fragmento leído; comparación de hablantes.

ENG A proposal for a phonologically balanced passage: *El ratón Arturo, the Spanish adaptation of Arthur the Rat*

Abstract: Phonetically (phonologically) balanced texts are often used to collect linguistic data. However, there are few texts in Spanish with these characteristics. The aim of this paper is to propose a new text, *El ratón Arturo*, adapted from its original English version. The rationale for creating a new text is that the existing ones, Spanish truly phonetically-balanced-texts, are too short, less than two minutes when reading. This short length hinders their use for calculating statistically accurate vocal metrics. *El ratón Arturo* is longer and contains fewer repetitions of words than the existing ones, high frequency words, and simpler sentences. Spanish phonetically-balanced-texts reading lasts more than two minutes. Although longer, it is neither hard to read nor to process; its complexity is similar to previously existing texts. Moreover, it can be used to collect data from adult or infant populations. It meets two essential requirements for this type of text: it presents at least one occurrence of each Spanish phonemes, following Spanish phoneme distribution. In addition, examples of most frequent consonant clusters and codas are represented. Likewise, it includes simple and complex words, contrastive accent samples, sentences with and without nasal consonants and several modality sentences. These specific characteristics allow the use of this text not only in phonetic research but also in the evaluation of speech motor disorders, such as dysarthrias or apraxias, and vocal function.

Keywords: phonetically balanced text; voice and speech production measurement; reading passage; speaker comparison.

Sumario: Índice: 1. Introducción. 1.1. El ratón Arturo. 2. Metodología. 2.1. La adaptación del texto al español. 2.2. Transcripción y análisis fonológico. 2.3. Análisis estadístico. 3. Resultados. 3.1. Propuesta de texto fonológicamente equilibrado. 3.2. Estadística descriptiva del texto. 3.3. Análisis de correlación y de concordancia. 4. Discusión. 5. Conclusiones. Agradecimientos. Contribución de autoría CREDiT. Referencias bibliográficas.

Cómo citar: Casado-Mancebo, M., Polo Cano, N., Teira Serrano, C. (2024). Una propuesta de texto fonológicamente equilibrado: *El ratón Arturo*, la adaptación al español de *Arthur the Rat*. *Lingüística Aplicada a la Comunicación* 98 (2024). 289-299. <https://dx.doi.org/10.5209/clac.81669>

1. Introducción

En la investigación sobre cómo los hablantes usan el lenguaje existen diversas formas de conseguir datos de producción que, a grandes rasgos, se resumen en recoger muestras de habla oral espontánea o leída (Podesva y Zsiga, 2016). Las primeras se obtienen mediante lenguaje dirigido (por ejemplo, describir una imagen), entrevistas o conversaciones y las segundas mediante la compleción de tareas y cuestionarios o la lectura de palabras, oraciones o textos estandarizados. El investigador elige la que más le convenga según el objetivo de su investigación.

Asimismo, para la evaluación clínica de los trastornos motores de la producción del habla (disartrias y apraxias) se utilizan tanto las tareas descritas en el párrafo anterior como tareas específicas (Duffy, 2013; Hegde y Freed, 2017; Freed, 2020). Entre estas últimas se encuentran los test auditivos, los exámenes orofaciales y diacocinésicos (motores), la administración de test estandarizados y las evaluaciones de la conversación y del habla natural (Hegde y Freed, 2017). La evaluación motora del habla se realiza mediante la producción lo más rápida posible de sílabas /pa/, /ta, /ka/ para evaluar la acción labial (/pa/), la elevación de la punta de la lengua (/ta/) y de la parte posterior de la lengua (/ka/); mediante la producción de palabras con distinta complejidad silábica y morfológica (como palabras simples y derivadas), palabras muy frecuentes (los días de la semana) o habla automática (como contar) y mediante el análisis de la cadena hablada (Freed, 2020).

No obstante, las decisiones sobre cómo recabar los datos para la investigación o para evaluar a personas con trastornos del habla influyen en la espontaneidad y el control de la muestra, y en la replicabilidad de datos entre hablantes o situaciones y su posterior generalización tanto al resto de la población como para el seguimiento de la evaluación de un mismo hablante en el tiempo. Por ejemplo, utilizar la lectura de fragmentos de texto escrito puede reducir la espontaneidad de la muestra frente a utilizar una conversación espontánea (Laan, 1997), aunque existe controversia entre los autores sobre eso (Boyd et al., 2015). De hecho, al pedir a los participantes que lean el mismo fragmento se consigue que todos reproduzcan el mismo contenido, lo que facilita su comparación (Nolan et al., 2009). La ventaja de utilizar fragmentos de texto leído sobre oraciones independientes es que permite al hablante exhibir los patrones entonativos naturales de su lengua (Abberton, 2005) y favorece además la producción de fenómenos poco frecuentes que no suelen aparecer en lo que dura la emisión de una oración simple (Boyd et al., 2015). De hecho, a partir de la modulación prosódica, un fragmento de texto permite observar el control motor del habla, algo que no es posible con tareas que usan exclusivamente sílabas o palabras aisladas (Patel et al., 2013).

Un subgrupo de los fragmentos de texto dentro de la investigación son los *textos fonéticamente equilibrados* (Gibbon et al., 1997). Estos se caracterizan por: (i) presentar el número completo de fonemas de la lengua en cuestión, al menos un ejemplo de cada fonema, y (ii) hacerlo con la misma frecuencia de aparición que en el habla real. Además, contienen ejemplos de alofonía y reglas fonotácticas que operan en la lengua del texto (Baird et al., 2021) y están contruidos con palabras frecuentes (Jesus et al., 2015). De hecho, lo deseable es que se cumplan estos requisitos con el número menor de palabras posible (Coloma, 2015). Sin embargo, si el texto es demasiado corto no se suele conseguir que sea fonológicamente equilibrado (Baird et al., 2021). Por ello, para que se consigan estos dos requisitos los textos fonológicamente equilibrados suelen contener entre 100 y 300 palabras (Lammert et al., 2020). Adicionalmente, para que pueda ser utilizado para la evaluación vocal y cumpla con los requisitos del análisis fonético, la lectura de este tipo de materiales debe durar al menos dos minutos (Howard, 1998; Abberton, 2005). Si en términos de velocidad de elocución se leen en voz alta alrededor de 160-180 palabras por minuto (Rodero, 2012), para cumplir con este requisito de longitud los textos no podrían tener menos de 320-360 palabras.

En resumen, al construirlos con estos requisitos se consigue que este tipo de textos presenten validez ecológica, pues son representativos de lo que realmente ocurre en la lengua natural. Esto permitirá la generalización de los datos fonético-fonológicos (o de otro tipo) que se recojan a toda la población. Por ello, se utilizan para conseguir muestras de discurso continuado, de procesos fonológicos cuyo alcance va más allá de la palabra morfológica o para documentar rasgos suprasegmentales (Podesva y Zsiga, 2016). Si además se precisa hacer uso de dichos fragmentos para recoger datos clínicos o de poblaciones con trastornos vocales, donde el control motor se encuentra afectado, resulta necesario además controlar otras características fonéticas de los textos como, por ejemplo, la sonoridad de los sonidos, la complejidad de las sílabas, la complejidad de las palabras, etc.

Existe gran cantidad de ejemplos de estos tipos de textos en lengua inglesa. Quizás el primero y más conocido es el *The Northwind and the Sun* («El viento norte y el sol»), una fábula de Esopo que se empezó a utilizar en 1949 (Asociación Internacional de Fonética, 1999; 2010) y que ha sido adaptada a más de 150 lenguas (Baird et al., 2021). Junto a este texto, se han ido publicando otros similares (cf. Powell, 2006, para la revisión y referencias de los textos ingleses): *If* (Kipling, 1940), *Your Rate of Oral Reading* (Fairbanks, 1940), *The Rainbow Passage* (Fairbanks, 1960), *My Grandfather* (Van Riper, 1963; Darley et al., 1975, citados en Reilly y Fisher, 2012), *Arthur the Rat* (Abercrombie, 1964), *The Limpy Passage* (Wilson y Rice, 1977), *The Farm Passage* y *The Hunter Passage* (Crystal y House, 1982), *The Picnic* y *The Trip to the Zoo* (Wilson, 1987) y *Comma Gets a Cure* (Honorof et al., 2000), además de otros no recogidos en la clasificación de Powell (2006) como *Towne-Heuer Reading Passage* (Heuer et al., 2000), *The Boy who Cried Wolf* (Deterding, 2006) y *The Caterpillar* (Patel et al., 2013).

Frente a esta prolífica creación de materiales en inglés, en español apenas se reconocen seis textos que cumplen con las características descritas en párrafos anteriores. Un ejemplo es la versión del español centronorte peninsular de *El viento norte y el sol* (Martínez-Celdrán, Fernández y Carrera, 2003), con su versión del español murciano (Monroy y Hernández, 2015) y la adaptada de Coloma (2015). La versión tradicional en español de *El viento norte y el sol*, aunque sea uno de los textos más utilizados, no está exenta de limitaciones. No tiene ejemplos de uso de todos los fonemas del español, ya que /ɲ/ y /j/ no aparecen, presenta muchas palabras repetidas y no reproduce la frecuencia de fonemas del español. Por ejemplo, /o, r, b, x/ estarían sobrerrepresentados y /d, θ/ infrarrepresentados (Coloma, 2015). En España, también son conocidos dos textos más: *El joyero Federico Vanero... y Hay algo ahí, en el aire...*, concebidos en el marco de proyectos de investigación concretos, disponibles en: https://joaquimllisterri.cat/phonetics/fon_esp/Textos_equilibrio_fonetico_espanol.html (Bruyninckx et al., 1994; Ortega et al., 2000). Para el español de Chile son muy conocidos *El abuelo* (González, y Bevilacqua, 2012) y *La familia* (Toledo, 2017). Sin embargo, aunque supuestamente son textos fonológicamente equilibrados, no lo están: aunque presentan el inventario fonológico propio de esta variedad del español, no reproducen la distribución de los fonemas de esta variedad para un corpus de noticias orales de referencia. *La familia* sobrerrepresenta los fonemas /a/ y /m/ e infrarrepresenta la /e/ y *El abuelo* sobrerrepresenta los fonemas /u/ y /b/ (Martínez-Cifuentes et al., 2020). Además, la versión española de *The Rainbow Passage* «*El arco iris*» usada en Colombia tampoco sigue la distribución de fonemas en esa variedad de español (Camargo y Marín, 2013). Finalmente, los textos españoles tienen en común con la mayoría de los textos ingleses arriba recogidos (Lammert et al., 2020) que son demasiado cortos (menos de 200 palabras) para conseguir reproducir todos los fonemas de la lengua de forma equilibrada y en contexto, ya que su lectura apenas dura menos de un minuto. Consecuentemente, hace cuestionable su utilización para tareas de evaluación logopédica (Melle, 2007).

1.1. El ratón Arturo

Un texto más largo es *Arthur the Rat*. Esta fábula fue recomendada ya en el siglo XIX para llevar a cabo análisis fonéticos en inglés por Paul Passy (Passy, 1886) y Henry Sweet (Sweet, 1890), citados en MacMahon (1991). No obstante, su autoría se atribuye a Gertrude Sellon, la cual publicó el texto original en 1876 bajo el título *I Don't Know* («No lo sé») (MacMahon, 1991). Desde entonces se ha utilizado ampliamente para documentar la lengua inglesa (por ejemplo, von Schneidermesser, 2006) o comparar hablantes de inglés entre sí (Byrne y Foulkes, 2004). Además, la trama es atractiva, como se recomienda en la bibliografía para este tipo de textos (Jesus et al., 2015; Martínez-Cifuentes et al., 2020).

Arthur the Rat está traducido al portugués (Guimarães, 2002: 261) y se utiliza habitualmente en investigaciones sobre esta lengua. La versión inglesa más moderna cuenta con 361 palabras (Abercrombie, 1964: 38) y 997 segmentos (Abberton, 2005). La portuguesa con 290 palabras y 26 oraciones, con entre 2 y 24 palabras por oración (duración media de 12.42 palabras por oración) (Guimarães, 2002: 261). Sin embargo, el texto no está traducido al español. Por eso, el objetivo de este trabajo es presentar una propuesta de traducción y adaptación al español de *Arthur the Rat* como el «El ratón Arturo» para construir un texto fonológicamente equilibrado en español, que además es lo suficientemente largo como para pueda ser utilizado tanto en la investigación lingüística como en la práctica clínica, p. ej. logopédica.

2. Metodología

2.1. La adaptación del texto al español

La traducción del texto ha seguido las recomendaciones de traducción de textos para ser comprendidos en la práctica investigadora o clínica: en un registro no demasiado formal, con palabras de uso habitual y fácil de leer (Muñiz et al., 2013). Este diseño reduce la carga cognitiva que puede generar la lectura de un texto (Abberton, 2005). El texto en inglés tiene algunas palabras y construcciones sintácticas poco frecuentes (Powell, 2006), lo que se ha tratado de evitar en la adaptación al español. Asimismo, como el final de la historia en la versión inglesa ha resultado controvertido (Powell, 2006), se ha suavizado para que pueda ser empleado también con niños. Con todo, se ha redactado de forma que sea independiente de la cultura y del momento actual para que pueda ser un texto atemporal.

Además de cumplir con los requisitos de la presencia y distribución de los fonemas del español, se han seguido los consejos de Patel y sus colaboradores (2013) para la confección de textos que sirvan para documentar los problemas motores del habla. Así, el texto se ha construido con oraciones cortas y largas para permitir observar patrones respiratorios y la correcta gestión de los grupos fonatorios. También se han incluido palabras simples y derivadas (más complejas), como *ratón/ratonera*, *techo/tejado*, *lluvia/lluvioso*, lo

que habilita al texto para ser usado como un material más dentro de las tareas típicas que evalúan el control motor del habla; y alguna palabra con acento contrastivo propio del español (*ahí/hay*). Finalmente, se han incluido oraciones con gran cantidad de consonantes nasales (por ejemplo, *media vuelta, en marcha*) para inferir la función velofaríngea al compararla con las oraciones construidas exclusivamente con fonemas orales (*se quedó callado*) (Heurer et al., 2000; Patel et al., 2013).

2.2. Transcripción y análisis fonológico

La transcripción propuesta del texto sigue las preceptivas del Alfabeto Fonético Internacional (IPA 1888[1949]; 1999; 2010) y recoge una transcripción fonológica de la variante del español centronorte peninsular. Se han seguido las convenciones en Martínez-Celdrán y colaboradores (2003) para transcribir el español. Se recogen manifestaciones ortográficas de los 24 fonemas del español: cinco vocales (/a/, /e/, /i/, /o/, /u/) y diecinueve consonantes (/p/, /t/, /k/, /b/, /d/, /g/, /tʃ/, /j/, /f/, /s/, /x/, /m/, /n/, /ɲ/, /r/, /r/, /l/, /θ/ y /ʎ/). No obstante, aunque ortográficamente están representadas palabras con el fonema /ʎ/, como *lluvia*, hay que señalar que en los análisis estadísticos no se ha considerado /ʎ/ como un fonema adicional más, teniendo en cuenta su uso poco habitual actualmente en español como consecuencia del fenómeno del yeísmo (RAE-ASALE, 2011). Se ha hecho así para que el texto pueda ser utilizado, llegado el caso, por hablantes que todavía mantienen el uso distinguidor de /ʎ/ y /j/. La transcripción fonológica ha partido de una versión generada automáticamente por la herramienta TraFo (<https://raco.cat/index.php/EFE/article/view/375474>, Rodríguez, 2020) que se ha revisado manualmente.

Para la comparación entre la distribución de la frecuencia de los fonemas del texto con la frecuencia de los fonemas del habla natural se han tomado los datos de dos corpus de referencia a partir de textos escritos: (i) un corpus formado por textos periodísticos de la agencia EFE con 480 000 palabras y 2 511 856 fonemas (Moreno Sandoval et al., 2006) y (ii) un corpus compuesto por textos literarios formado por 52 661 222 palabras y 236 024 884 fonemas (Arias, 2016). Se ha optado por comparar la propuesta de este nuevo texto con las frecuencias de textos escritos siguiendo la recomendación de Coloma (2015). No obstante, también se han tenido en cuenta los datos de dos corpus orales, compuestos por 348 000 palabras y 1 244 411 fonemas, y 80 367 fonemas, respectivamente (Moreno Sandoval et al., 2006; Machuca y Ríos, 2017, respectivamente).

2.3. Análisis estadístico

Para confeccionar el texto de *El ratón Arturo* de manera fonológicamente equilibrada se ha llevado a cabo un diseño cuantitativo comparativo no experimental descriptivo e inferencial. En primer lugar, se ha confirmado que estuviera presente al menos una grafía de cada uno de los 24 fonemas del español mediante un conteo manual. En segundo lugar, mediante el uso de Python (Casado-Mancebo, 2022a) se ha extraído la cantidad de segmentos totales, cuántos de ellos corresponden a vocales y cuántos a consonantes. En tercer lugar, se ha calculado la frecuencia de distribución de los fonemas en el texto, tanto de manera absoluta como relativa. Adicionalmente, en el análisis descriptivo se ha comparado la distribución de fonemas, sílabas, palabras y oraciones con los seis textos de similares características del español. En la tabla 1 se presentan los valores absolutos y las ratios entre dichos valores.

El análisis estadístico se ha realizado a partir del análisis de normalidad con el estadístico Shapiro-Wilk, utilizado para muestras de datos pequeñas. A continuación, en el análisis inferencial se ha comparado la distribución de los fonemas en los corpus de referencia con la distribución de fonemas en el texto de *El ratón Arturo* mediante el coeficiente de correlación por rangos (r de Spearman) para comprobar si ambas distribuciones frecuenciales se correlacionan. El ajuste es unilateral, ya que se hipotetiza que, si la frecuencia de fonemas de *El ratón Arturo* aumenta, la de los corpus de referencia también lo hará. Finalmente, el coeficiente de correlación ofrece la fuerza de la relación entre dos variables, pero no el acuerdo entre ellas. Para controlar que no exista acuerdo entre las dos muestras o que este sea pequeño, a pesar de la correlación, se ha utilizado una exploración visual de los datos a partir del cálculo normalizado de la media de la diferencia entre ambas distribuciones frecuenciales siguiendo la propuesta de Bland y Altman (1986). En la práctica esto significa comprobar que la presencia de un determinado fonema no esté sobre o infrarrepresentada en el texto en cuestión. Para ello se han normalizado los datos de la frecuencia de cada fonema a partir de la siguiente fórmula:

$$[f(\text{El ratón Arturo}) - f(\text{Corpus de referencia})] \pm (2 \cdot DT)[f(\text{El ratón Arturo}) - f(\text{Corpus de referencia})]$$

donde f significa la frecuencia de cada fonema. Si los valores resultantes están por debajo o por encima de los límites que ofrece la fórmula (± 2 desviaciones típicas, DT , $\alpha = 0.05$), se presume que los fonemas están sobre o infrarrepresentados en el texto. Si los valores frecuenciales se encuentran de acuerdo con los del corpus de referencia, se tienen que localizar dentro de estos umbrales.

Los análisis estadísticos se han realizado con la versión 24 del software de análisis estadístico SPSS (IBM Corporation, Armonk, NY) y el acuerdo entre las muestras y la figura mediante Microsoft Excel.

3. Resultados

3.1. Propuesta de texto fonológicamente equilibrado

Se presenta la propuesta de traducción y adaptación del texto *El ratón Arturo* al español. Dotar al texto de las características arriba mencionadas ha forzado una adaptación que no sigue literalmente el texto en inglés. Para garantizar que todas las palabras pertenecen a un uso relativamente común del español, se utilizó

la herramienta *Explorador de frecuencias* (Casado-Mancebo, 2022b) tomando como base de exploración el CORPES-XXI (Real Academia Española). Se comprobó que las formas del texto estuvieran entre las cien mil más frecuentes según este corpus. Hay que tener en cuenta que el objeto de análisis no es el lema sino la forma, por lo que las frecuencias de aparición son considerablemente más reducidas, lo que incrementa considerablemente su índice en el listado de frecuencia. Como muestra, la forma menos frecuente según los resultados de exploración es *preocupará*, en la posición 126 958. Su lema, sin embargo, ocupa la posición 869.

Seguidamente se incluye la transcripción del texto. No se han transcrito los archifonemas. Se ha optado por transcribir como nasales bilabiales las nasales que están en posición de coda antes de una oclusiva labial a semejanza de otras transcripciones en español con fines similares (Martínez Celadrán et al., 2003; Coloma, 2015). Como no se han separado oclusivas sonoras de espirantes en el conteo de frecuencias se transcriben estos segmentos como oclusivos. Igualmente, las paravocales, aunque se han transcrito como /j,w/, y teniendo en cuenta la discusión sobre el estatus fonológico de las paravocales en español (Roca, 1997), se han considerado vocales en el análisis frecuencial. La consonante palatal se ha transcrito /j/ y analizado conjuntamente en su variante fricativa, tanto para las grafías <y> como <ll>. El discurso directo se transcribe entre comillas para que sea más fácil la comprensión de la transcripción.

Transcripción ortográfica de *El ratón Arturo*

Hubo una vez un joven ratón llamado Arturo que nunca se molestó en tomar decisiones. Cada vez que sus amigos le preguntaban si quería salir con ellos a la calle, él les contestaba que no sabía. No les decía que «sí» pero tampoco les decía que «no». No aprendió a pronunciarse sobre nada de lo que le decían.

Su tía Elena le señaló «Nadie se preocupará por ti jamás si sigues siendo así. Tienes menos cerebro que las briznas de hierba». Arturo la miró con cautela y, una vez más, se quedó callado.

Un día lluvioso los ratones escucharon fuertes ruidos en el almacén donde vivía la familia. ¿Qué estaba pasando? Todas las vigas estaban carcomidas, y finalmente una de las que colgaban se desprendió y cayó de golpe al suelo. Las paredes casi temblaron y a los ratones se les erizó el pelo del miedo. Un ratón viejo, que era el jefe de los ratones, anunció con preocupación «Este sitio ya no servirá. Mandaré a los ratones exploradores a buscar otra casa en la que vivir tranquilos».

Tres horas después, cuatro expedicionarios volvieron con noticias: «Hemos encontrado unas casas de piedra con todo lo que necesitamos. Ahí hay espacio y comida para todas las familias, un caballo afable llamado Nelly, una vaca, una ternera y un jardín con olmos».

Fue entonces cuando el ratón viejo miró al joven Arturo. «Vas a venir, ¿verdad?». «No sé» suspiró Arturo una vez más, «puede que al final el techo aguante». El ratón viejo dijo enfadado «Bien, no podemos esperar todo el día a que te decidas. ¡Media vuelta! ¡En marcha!» y, de esta forma, los ratones se fueron. Arturo permaneció allí viendo a los otros ratones desaparecer bajo la lluvia. La idea de tener que decidir inmediatamente era demasiado para él. «Me iré a la ratonera un rato a pensar», se dijo, «lo justo para tomar la decisión correcta».

Esa misma noche un gran estruendo sacudió la tierra y se vino abajo todo el tejado. Al día siguiente, algunos hombres se acercaron e inspeccionaron las ruinas. Uno de ellos movió un tablero, se agachó y bajo él pudo intuir a un joven ratón tumbado de lado, casi muerto, medio dentro medio fuera de un agujero.

Transcripción fonológica de *El ratón Arturo*

'ubo una 'beθ un 'xoben ra'ton ja'mado ar'turo ke 'nunka se moles'to en to'mar deθi'sjones. 'kada 'beθ ke sus a'migos le pregun'taban si ke'ria sa'lir kon 'ejos a la 'kaje, el les kontes'taba ke no sa'bia. no les de'θia ke «'si» pero tam'poko les de'θia ke «'no». no apren'djo a pronun'θjarse sobre 'nada de lo ke le de'θian.

su 'tia e'lena le seña'lo «'nadje se preokupa'ra por ti xa'mas si 'siges 'sjendo a'si. 'tjenes 'menos θe'rebro ke las 'briθnas de 'jerba». ar'turo la mi'ro kon kaw'tela i, una 'beθ 'mas, se ke'do ka'jado.

un 'dia ju'bjoso los ra'tones esku'tjaron 'fwertes 'rwidos en el alma'θen donde bi'bia la fa'milja. ¿'ke es'taba pa'sando? 'todas las 'bigas es'taban karko'midas, i final'mente una de las ke kol'gaban se despren'djo i ka'jo de 'golpe al 'swelo. las pa'redes 'kasi tem'blaron i a los ra'tones se les eri'θo el 'pelo del 'mjedo. un ra'ton 'bjexo, ke era el 'xefe de los ra'tones, anun'θjo kon preokupa'θjon «este 'sitjo ja no serbi'ra. manda're a los ra'tones eksplora'dores a bus'kar otra 'kasa en la ke bi'bir tran'kilos».

tres 'oras des'pwes, 'kwatro ekspedi'θjo'narjos bol'bjeron kon no'tiθjas: «emos enkon'trado unas 'kasas de 'pjedra kon todo lo ke neθesi'tamos. a'i aj es'paθjo i ko'mida para 'todas las fa'miljas, un ka'bajo a'fable ja'mado 'neji, una 'baka, una ter'nera i un xa'r'din kon 'olmos».

fwe en'tonθes 'kwando el ra'ton 'bjexo mi'ro al 'xoben ar'turo. «bas a be'nir, ¿ber'dad?». «no 'se» suspi'ro ar'turo una 'beθ 'mas, «'pwede ke al fi'nal el 'tetjo a'gwante». el ra'ton 'bjexo 'dixo enfa'dado «'bjen, no po'demos espe'rar todo el 'dia a ke te de'θidas. ¡'medja 'bwelta! ¡en 'martja!» i, de esta 'forma, los ra'tones se 'fweron.

ar'turo permane'θjo a'ji 'bjendo a los 'otros ra'tones desapare'θer 'baxo la 'jubja. la i'dea de te'ner ke deθi'dir inmedjata'mente era dema'sjado para 'el. «me i're a la rato'nera un 'rato a pen'sar», se 'dixo, «lo 'xusto para to'mar la deθi'sjon ko'rekta».

esa 'misma 'notje un gran es'trwendo saku'djo la 'tjera i se 'bino a'baxo 'todo el te'xado. al 'dia si'gjente, al'gunos 'ombres se aθer'karon e inspek'θjo'naron las 'rwinas. uno de ejos mo'bjos un ta'blero, se aga'tjo i baxo 'el 'pudo in'twir a un 'xoben ra'ton tum'bado de 'lado, kasi 'mwerto, medjo 'dentro medjo 'fwera de un agu'xero.

3.2. Estadística descriptiva del texto

La propuesta de texto fonológicamente equilibrado considera los fonemas propios del español en su versión yeísta (23 fonemas). No obstante, se han incluido ejemplos como *lluvia* para poder utilizarlo también en casos de hablantes no yeístas. La comprobación de la presencia de todos los fonemas del español se ha realizado mediante un conteo manual. Asimismo, el texto contiene ejemplos de grupos consonánticos y codas frecuentes en español (Hualde, 2013). Por ejemplo, cuenta con representación de consonante+/r/ como en *brizna* o *piedra* y consonante+/l/ como *exploradores* o *afable*, lo que permitiría evaluar también la producción de grupos consonánticos. Igualmente, aparecen ocurrencias de fonemas en posición de coda en interior y en final de palabra. Por ejemplo, en posición interior aparecen codas simples oclusivas (*correcta*), fricativas (*brizna*, *espacio*), nasales (*entonces*) y líquidas (*almacén*, *Arturo*) y codas complejas compuestas por más de un fonema (*exploradores*, *inspeccionaron*). En posición final aparecen las codas más frecuentes en español en esta posición, /θ, s, n, l, r/, como *vez*, *jamás*, *joven*, *final* y *tomar*.

Como resultado, la propuesta de texto contiene 1612 ocurrencias de fonemas, 772 vocales y 840 consonantes, un texto más largo que los que se utilizan hasta el momento en español que de media presentan 618 fonemas (292 vocálicos y 326 consonánticos), como se comprueba en la tabla 2. La proporción de obstruyentes es de 69 % y la de sonantes es de 31 %. Igualmente, en los textos previos la media de sílabas es 271.83, la de palabras 142.83 y la de oraciones 9.16. En esta propuesta aparecen 724 sílabas, 373 palabras y 28 oraciones. La tabla 1 muestra la estadística descriptiva que compara los valores de los textos fonológicamente equilibrados propuestos para el español hasta el momento con la versión propuesta de *El ratón Arturo*.

Tabla 1. Estadística descriptiva que compara los 6 textos fonológicamente equilibrados en español y el nuevo texto propuesto de *El ratón Arturo*.

Texto	El viento norte	El joyero...	Hay algo ahí...	El abuelo	La familia	El arco iris	El ratón Arturo
Vocales	202	214	345	211	435	345	772
Consonantes	226	238	365	233	498	396	840
Fonemas (ocurrencias)	428	452	710	444	933	741	1612
Fonemas (tipos)	22	23	23	23	23	23	23
Sílabas	184	200	308	195	415	329	724
Palabras (ocurrencias)	97	103	179	101	201	176	373
Palabras (tipos)	60	68	123	80	133	102	218
Oraciones	4	3	12	7	17	12	28
Sílabas/palabras ratio	1.89	1.94	1.72	1.93	2.35	1.86	1.94
Tipos/ocurrencias palabras ratio	0.62	0.66	0.69	0.79	0.66	0.58	0.58
Palabras/oración ratio	24.25	34.33	14.91	14.42	11.82	14.66	13.32

El texto propuesto de *El ratón Arturo* no solamente es un texto más largo que los anteriores (casi 400 palabras, frente a las 142.83 de media de los demás textos), sino que, además, presenta junto con el texto de *El arco iris*, el menor número de palabras repetidas (ratio de 0.58 frente al texto de *El abuelo* que presenta una ratio más alta, 0.79). La complejidad de sílabas por palabra es similar a los textos anteriores, lo que significa que no está constituido por palabras más complejas. También se ha compuesto con oraciones más cortas, con una ratio de palabras/oración, 13.32, muy por debajo de la media (19.06). Asimismo, su lectura en voz alta dura más de dos minutos al superar las 360 palabras. En la tabla 2 se muestran los valores medios y su desviación típica de los valores recogidos en la tabla 1.

Tabla 2. Media y desviación típica (DT) de los textos comparados en la tabla 1 sin incluir el ratón Arturo.

Características analizadas	Media	DT
Vocales	292.00	96.76
Consonantes	326.00	111.71
Fonemas (ocurrencias)	618.00	208.21
Fonemas (tipos)	22.83	0.41
Sílabas	271.83	93.65
Palabras (ocurrencias)	142.83	47.39
Palabras (tipos)	94.33	29.84
Oraciones	9.17	5.42
Sílabas/palabras ratio	1.95	0.21
Tipos/ocurrencias palabras ratio	0.67	0.07
Palabras/oración ratio	19.07	8.61

3.3. Análisis de correlación y de concordancia

Las frecuencias de aparición de los fonemas tanto en el texto de *El ratón Arturo* como en los corpus de referencia no cumplen con el supuesto de normalidad, de acuerdo con el test de Shapiro-Wilk ($p < 0.05$). Se ha calculado la relación entre las distribuciones asumiendo que no son independientes, porque se espera que cuando una aumenta la otra también lo hace. El coeficiente de correlación que se ha empleado es la Rho de Spearman para poder evaluar el acuerdo entre el orden de distribución de fonemas en el texto de *El ratón Arturo* y los corpus de referencia. En todas las correlaciones se encuentra una correlación positiva muy alta ($p < 0.001$) con ajuste unilateral: r de 0.978 y 0.982 con los textos escritos, Moreno Sandoval et al. (2006) y Arias (2016), respectivamente; y r de 0.959 y 0.969 para los textos orales, Moreno Sandoval et al. (2006) y Machuca y Ríos (2017), respectivamente.

En los dos gráficos de la figura 1 se comprueba que no existen frecuencias de fonemas por encima o por debajo de 2 distribuciones típicas con respecto de la diferencia de medias entre las frecuencias de fonemas del texto de *El ratón Arturo* y los dos corpus de referencia de textos escritos. Si existieran fonemas fuera de estos límites no se podría entender que hubiera acuerdo entre ambas distribuciones. Como se puede comprobar visualmente esto no ocurre así, por tanto, se puede afirmar que la distribución de fonemas del texto de *El ratón Arturo* propuesto sigue la distribución típica de fonemas para el español escrito.

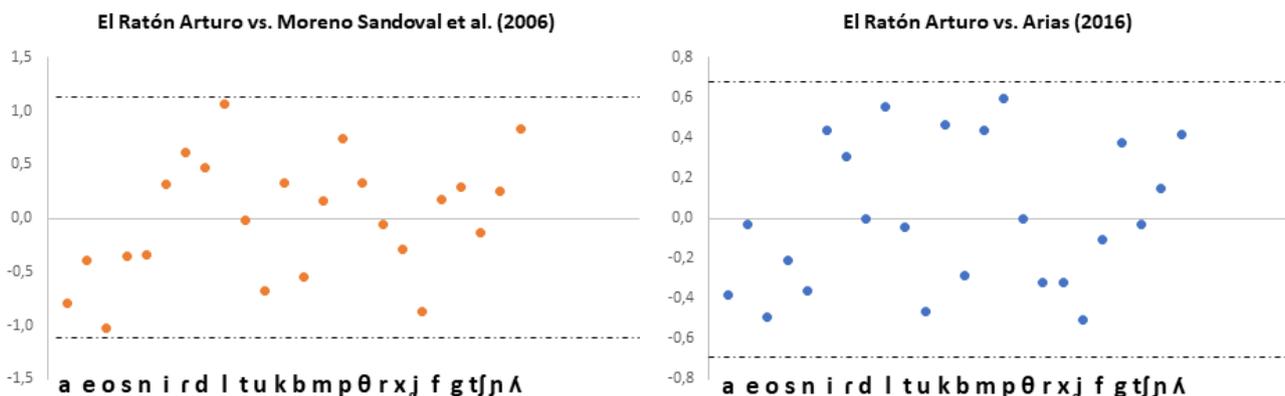


Figura 1. Las dos gráficas muestran la diferencia de medias entre las frecuencias de fonemas del texto de *El ratón Arturo* y los dos corpus de referencia de textos escritos, a la izquierda los textos periodísticos (Moreno Sandoval et al. 2006) y a la derecha los literarios (Arias, 2016). Las líneas punteadas limitan los resultados de las diferencias de medias de las frecuencias de cada fonema 2 desviaciones típicas (DT) por encima y por debajo de cada distribución (-1.11 y 1.12, respectivamente, en la distribución de Moreno Sandoval et al., 2006, y -0.71 y 0.73 por debajo y por encima en la de Arias, 2016).

Adicionalmente, se ha calculado esta diferencia de medias con los dos corpus para textos orales. En este caso las frecuencias propuestas para el fonema /e/ en el texto de *El ratón Arturo* se encuentran por debajo de lo considerado habitual para el español oral, 15.12 % (Moreno Sandoval et al., 2006) o 14.6 % (Machuca y Ríos, 2017). Según las comparaciones con los corpus de referencia, los textos escritos presentan menor número de vocales /e/ que los corpus orales (12.74 %, Moreno Sandoval et al., 2006; 13.1 %, Arias, 2016) y la versión de *El ratón Arturo* presentada se corresponde con las frecuencias de la vocal /e/ para textos escritos (13.13 %) (tabla 3).

Tabla 3. Frecuencias relativas de fonemas (expresadas en porcentajes) de los 4 corpus de referencia más las frecuencias y los valores absolutos para la versión propuesta de *El ratón Arturo*.

	Corpus escritos		Corpus orales		El Ratón Arturo	
	Moreno Sandoval	Arias	Moreno Sandoval	Machuca y Ríos	Frecuencia relativa	Valor absoluto
	Frecuencia relativa	Valor absoluto				
a	12.89	13.3	12.27	13.2	13.68	221
e	12.74	13.1	15.12	14.6	13.13	212
o	9.32	9.85	10.38	9.7	10.34	167
s	7.33	7.47	8.11	7.3	7.68	124
n	7.09	7.07	7.05	6.8	7.43	120
i	7.25	7.37	7.22	6.4	6.93	112
r	6.19	5.88	5.12	5.2	5.57	90
d	5.42	4.95	4.36	4.5	4.95	80
l	5.46	4.95	4.51	4.9	4.40	71
t	4.31	4.29	4.52	4.5	4.33	70
u	3.04	3.25	3.14	3.6	3.72	60
k	3.8	3.93	4.49	3.1	3.47	56
b	2.55	2.81	2.5	2.8	3.10	50
m	2.76	3.04	3.15	3.8	2.60	42
p	2.73	2.58	2.74	2.9	1.98	32
θ	2	1.67	1.5	1.4	1.67	27
r	0.99	0.73	0.42	0.5	1.05	17
x	0.77	0.73	0.62	0.6	1.05	17
j		0.36		0.6	0.87	14
f	0.92	0.64	0.5	0.7	0.74	12
g	1.04	1.12	0.91	2	0.74	12
tʃ	0.18	0.28	0.3	0.3	0.31	5
ɲ	0.31	0.21	0.19	0.3	0.06	1
ʎ	0.83	0.42	0.83	0.4	0	0

4. Discusión

El objetivo del presente trabajo es dotar a la comunidad científica y a la práctica clínica de un texto fonológicamente equilibrado adicional en español. Para ello se ha propuesto una versión adaptada al español de la historia original de *Arthur the Rat*, bajo el título de *El ratón Arturo*. Se ha elegido este texto porque es sencillo de leer, tiene ejemplos de narración y de diálogo, lo que permite documentar el rango de frecuencias de los hablantes y generar un habla más natural, razones por las que ya se había traducido al portugués (Guimarães, 2002: 146-147). Las características con las que se ha construido lo convierten en un texto idóneo para que sea utilizado tanto en la práctica fonética como en la evaluación logopédica, ya que su diseño permite usarlo también ante desórdenes motores como la disartria y la apraxia del habla.

La versión en español se ha construido cumpliendo los dos requisitos que se exige a estos tipos de textos: (i) presentar al menos una ocurrencia de cada uno de los fonemas del español y (ii) distribuir su frecuencia de aparición de acuerdo con la frecuencia de uso de fonemas en español. Por tanto, tiene la ventaja de presentar todos los fonemas del español, requisito que no cumple *El viento norte y el sol*, y presentar una distribución de frecuencias de fonemas de acuerdo con la lengua meta, requisito que no cumplen los textos de *El abuelo*, *La familia* y *El Arco Iris*. También se ha construido de forma que pueda ser utilizado tanto en variedades yeístas como no yeístas. Para que los textos fonológicamente equilibrados puedan recoger todos los fonemas al menos son necesarios 1000 segmentos (Abberton, 2005) y este texto está compuesto por 1612. La proporción de sonantes/obstruyentes cumple la ratio de 3:2 propia de muchas lenguas naturales (Abberton, 2005). Además, se recogen ejemplos de los grupos consonánticos y codas más frecuentes en

español, de palabras simples y derivadas, con acento contrastivo, de oraciones con presencia/ausencia de nasales, para evaluar la función velofaríngea, y con distintas longitudes, lo que permite evaluar el control fono-respiratorio, normalmente rasgos alterados en las apraxias (Patel et al., 2013).

Si se compara con la versión inglesa y la portuguesa, tiene más palabras, 373, frente a la versión inglesa con 361 palabras (Abercrombie, 1964: 38) y la portuguesa con 290 palabras (Guimarães, 2002), pero la complejidad es similar a la portuguesa, 13.32 palabras/oración frente a 12.42 en portugués. A pesar de que no es tanta la diferencia en número de palabras con el inglés, la distribución de fonemas por palabra en español fuerza a que sean necesarios más fonemas en español, 1612 frente a los 997 de la versión inglesa (Abberton, 2005).

A partir de la comparación de los datos entre el texto propuesto y los textos existentes en español se comprueba que el texto de *El ratón Arturo* es más largo (más del doble de largo que la media de 142 palabras de estos textos), especialmente en los casos de *El joyero... y Hay algo ahí...*. La mayor longitud frente a *El joyero... o Hay algo ahí...* garantiza un tiempo de lectura superior a los dos minutos al superar las 360 palabras (Rodero, 2012), lo que lo hace idóneo para ser utilizado en el análisis vocal y la práctica logopédica (Howard, 1998; Abberton, 2005; Lammert et al., 2020). Tiene la longitud suficiente para realizar medidas estadísticamente fiables del rango de frecuencia o de la irregularidad de la voz (Abberton, 2005).

A pesar de la longitud mayor, no es tan largo como para cansar al lector, no es un texto más difícil de leer o de procesar, ya que está compuesto por palabras frecuentes de la lengua española. Además, tiene menos palabras repetidas, una complejidad similar y oraciones más cortas. No obstante, la repetición de palabras también permite evaluar la variación intrahablante, en diferentes contextos o su evolución en el tiempo, por eso también aparecen palabras repetidas a lo largo del texto (Deterding, 2006), aunque en menor medida que en otros textos.

Algunos autores proponen eliminar todas las partes del diálogo para evitar que los hablantes durante la lectura pongan diferentes voces según distintos personajes, lo cual no sería una muestra real de su habla natural (Abberton, 2005). El texto se ha construido manteniendo estas partes precisamente para que al utilizar diversas modalidades oracionales sirva para recoger fenómenos suprasegmentales y evaluar la suficiencia prosódica como el correcto uso de pausas, de velocidad de elocución, duración, entonación, intensidad, etc., normalmente alterada en la disartria (Patel et al., 2013). No obstante, si estos cambios en las voces llegaran a ocurrir, se recomienda su lectura suprimiendo estas partes o convirtiéndolas en estilo indirecto.

El texto reproduce la frecuencia de fonemas del español documentada para la lengua escrita (Moreno Sandoval et al., 2006 y Arias, 2016). No obstante, si se quisiera reproducir la distribución de frecuencias de fonemas de la lengua oral, bastaría con incluir en algunos de los diálogos del texto llamadas como [ehhh], donde aparezca varias veces la vocal [e], haciendo uso de onomatopeyas, interjecciones o préstamos como recomiendan otros autores para ampliar las ocurrencias de fonemas (Hiki et al., 2011).

Finalmente, en la redacción de todo el trabajo se ha optado por la expresión *fonológicamente* equilibrado, en lugar de *fonéticamente* equilibrado, como se suele proponer para este tipo de textos, siguiendo las recomendaciones de Martínez-Cifuentes y sus colaboradores (2020), ya que la construcción del texto se relaciona con conceptos fonológicos, más que fonéticos.

A pesar de que como se ha recogido en la introducción de este trabajo, los textos fonológicamente equilibrados son adecuados para recabar o evaluar datos lingüísticos, la lectura de textos puede ser difícil para personas con déficit de visión, de lectura o de comprensión lectora en voz alta. En estos casos, se recomienda que los participantes en la tarea se familiaricen con el texto leyéndolo previamente antes de la grabación (Podesva y Zsiga, 2016). Igualmente, hay que vigilar que al leer no bajen demasiado la cabeza, en dirección al papel, lo que podría generar distorsiones vocales, por lo que, para evitarlo, el investigador puede colocar la hoja con el texto enfrente del informante en un atril y presentarla con una letra de un tamaño lo suficientemente grande para que sea leída cómodamente.

5. Conclusiones

Tras el análisis de los textos fonológicamente equilibrados existentes en español, se propone un texto nuevo, *El ratón Arturo*, a partir de un texto original en inglés. Este fragmento es más largo que los revisados en español, presenta un menor número de repeticiones de palabras, palabras frecuentes en esta lengua y oraciones sencillas. Su complejidad es similar a la de los otros textos, por lo que, a pesar de la longitud mayor, no significa que sea más difícil de leer o de procesar. Presenta al menos una ocurrencia de cada uno de los fonemas del español y de los grupos consonánticos y codas más frecuentes. Además, los fonemas se encuentran distribuidos según la frecuencia de uso de esta lengua. Asimismo, se ha confeccionado con palabras simples y derivadas, con acento contrastivo, con oraciones con presencia/ausencia de nasales y con oraciones con diversas modalidades oracionales. Por tanto, las características con las que se ha construido lo convierten en un texto idóneo para recabar datos con la validez suficiente requerida tanto para ser utilizado en la investigación lingüística, p.ej. fonética, como en la práctica clínica, p. ej. logopédica.

Agradecimientos

Investigación financiada en el marco del proyecto Generación de Conocimiento del Ministerio de Ciencia e Innovación, convocatoria 2021 «Estereotipos de género en la voz y el habla de las mujeres (ESTEREOFEM)» [PID2021-122607OA-I00].

A Filipa Lã por los comentarios recibidos sobre una primera versión del manuscrito y a los dos evaluadores anónimos por sus valiosos comentarios que han mejorado considerablemente el trabajo.

Contribución de autoría CREDiT

Todos los autores han contribuido por igual. El orden de firma sigue criterios estrictamente alfabéticos.

Referencias bibliográficas

- Abberton Evelyn (2005). Phonetic considerations in the design of voice assessment material. *Logopedics, phoniatrics, vocology*, 30(3-4), 175-180. <https://doi.org/10.1080/14015430500298123>
- Abercrombie, David (1964). *Elements of general phonetics*. Edinburgh: U.P. Edinburgh.
- Arias, Iván (2016). Cálculo de frecuencias de aparición de fonemas y alófonos en español actual utilizando un transcriptor automático. *Loquens*, 3(1), Artículo e029. <https://doi.org/10.3989/loquens.2016.029>
- Baird, Louise, Evans, Nicholas, y Greenhill, Simon J. (2021). Blowing in the wind: Using «North Wind and the Sun» texts to sample phoneme inventories. *Journal of the International Phonetic Association*, 1-42. Disponible en: https://pure.mpg.de/rest/items/item_3339463_2/component/file_3339564/content
- Bland, J. Martin y Altman, Douglas G. (1986) Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet*, 327(8476), 307-10. [https://doi.org/10.1016/S0140-6736\(86\)90837-8](https://doi.org/10.1016/S0140-6736(86)90837-8)
- Boyd, Zac, Elliott, Zuzana, Fruehwald, Josef, Hall-Lew, Lauren y Lawrence, Daniel (2015). «An Evaluation of Sociolinguistic Elicitation Methods», Paper presented at *The 18th International Conference of the Phonetic Sciences*, Glasgow, United Kingdom, 10/08/15 - 14/08/15. Disponible en: https://www.pure.ed.ac.uk/ws/files/21806989/Boyd_etal_2015_ICPhS_final.pdf
- Bruyninckx, Marielle, Harmegnies, Bernard, Llisterri, Joaquim y Poch-Oiivé, Dolores (1994). Language-induced voice quality variability in bilinguals. *Journal of Phonetics*, 22(1), 19-31. [https://doi.org/10.1016/S0095-4470\(19\)30265-7](https://doi.org/10.1016/S0095-4470(19)30265-7)
- Byrne, Catherine y Foulkes, Paul (2004). The «mobile phone effect» on vowel formants. *The International Journal of Speech, Language and the Law*, 11(1), 83-102. <https://doi.org/10.1558/ijssl.v11i1.83>
- Camargo, Mariluz y Marín, José (2013). ¿Es la lectura ««El arco iris»» apropiada para evaluar desórdenes del habla? *Revista Colombiana de Rehabilitación*, 12, 40-45. <http://dx.doi.org/10.30788/RevColReh.v12.n1.2013.42>
- Casado-Mancebo, Mario (2022a). Extraer frecuencias de fonemas en un texto. [Script de Python]. <https://github.com/mario-casado-m/frec-fon>
- Casado-Mancebo, Mario (2022b). Explorador de frecuencias [en línea]. Recuperado el 12/4/2022. <https://frecuencias.mcasado.org/>
- Coloma, Germán (2011). Caracterización fonética de las variedades regionales del español y propuesta de transcripción simplificada. *Revista de Filología Románica*, 28, 11-27. https://doi.org/10.5209/rev_RFRM.2011.v28.37217
- Coloma, Germán (2015). Una versión alternativa de «El viento norte y el sol» en español. *Revista de Investigación Lingüística*, 18, 191-212. Disponible en: <https://revistas.um.es/ril/article/view/248061>
- Deterding, David (2006). The North Wind versus a Wolf: Short texts for the description and measurement of English pronunciation. *Journal of the International Phonetic Association*, 36(2), 187-196. <https://doi.org/10.1017/S0025100306002544>
- Duffy, Joseph (2013). *Motor speech disorders* (3.ª ed.). St. Louis, Missouri: Elsevier Mosby.
- Freed, Donald (2020). *Motor speech disorders: Diagnosis and treatment* (3.ª ed.). San Diego, CA: Plural Publishing.
- Gibbon, Dafydd, Moore, Roger y Winski, Richard (1997). *Handbook of standards and resources for spoken language systems*. Berlín: Mouton de Gruyter.
- Guimarães, Isabel Cristina R. P. (2002). An electrolaryngographic study of dysphonie Portuguese speakers. Tesis de doctorado, University College London.
- González, Rafael A. y Bevilacqua, Jorge A. (2012). Las disartrias. *Revista del Hospital Clínico de la Universidad de Chile*, 23, 299-309.
- Hegde, M.N. y Freed, Donald (2017). *Assessment of Communication Disorders in Adults Resources and Protocols* (2.ª ed.). San Diego, CA: Plural Publishing.
- Heuer, Reinhardt, Towne, Carol, Hockstein, Neil E., Andrade, Deborah F. y Sataloff, Robert. T. (2000). The Towne-Heuer Reading Passage—a reliable aid to the evaluation of voice. *Journal of voice: official journal of the Voice Foundation*, 14(2), 236-239. [https://doi.org/10.1016/s0892-1997\(00\)80031-4](https://doi.org/10.1016/s0892-1997(00)80031-4)
- Hiki, Shizuo, Kakita, Kuniko y Okada, Hideo (2011). A panphonic version of the text of «The North Wind and the Sun» for the illustration of the IPA of Japanese (Tokyo dialect) consonants. *International Congress on Phonetic Sciences XVII*, Hong Kong, 17-21 August, 871-873. Disponible en: <https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2011/OnlineProceedings/RegularSession/Hiki/Hiki.pdf>
- Howard, David M. (1998). Instrumental voice measurement: uses and limitations. En Tom Harris; Sara Harris; John S. Rubin y David M. Howard (Eds). *The voice clinic handbook* (pp. 323-382). London: Whurr Publishers.
- Hualde, José Ignacio (2013). *Los sonidos del español*. Cambridge: Cambridge University Press. (Versión española corregida y aumentada de *The Sounds of Spanish*, 2005).
- IPA [International Phonetic Association] (ed.). (1999). *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet*. Cambridge: Cambridge University Press.

- IPA [International Phonetic Alphabet] (ed.). (2010). The principles of the International Phonetic Association (1988[1949]). *Journal of the International Phonetic Association*, 40, 299–358. doi:10.1017/S0025100311000089.
- Jesus, Luis M.T., Valente, Ana Rita S., y Hall, Andreia (2015). Is the Portuguese version of the passage «The North Wind and the Sun» phonetically balanced? *Journal of the International Phonetic Association*, 45, 1–11. <https://doi.org/10.1017/S0025100314000255>
- Laan, Gitta P. M. (1997). The contribution of intonation, segmental durations, and spectral features to the perception of a spontaneous and a read speaking style. *Speech Communication*, 22, 43–65. [https://doi.org/10.1016/S0167-6393\(97\)00012-5](https://doi.org/10.1016/S0167-6393(97)00012-5)
- Lammert, Adam C., Melot, Jennifer, Sturim, Douglas E., Hannon, Daniel J., DeLaura, Richard, Williamson, James R., Ciccarelli, Gregory y Quatieri, Thomas F. (2020). Analysis of Phonetic Balance in Standard English Passages. *Journal of speech, language, and hearing research: JSLHR*, 63(4), 917–930. https://doi.org/10.1044/2020_JSLHR-19-00001
- Machuca, María Jesús y Ríos, Antonio (2017). Frecuencia de fonemas y alófonos del español en la lengua oral espontánea. En Victoria Marrero y Eva Estebas (Coords.), *Tendencias actuales en fonética experimental: cruce de disciplinas en el centenario del «Manual de Pronunciación Española» (Tomás Navarro Tomás)* (pp. 93–97). Madrid: Universidad Nacional de Educación a Distancia. Disponible en: <http://e-spacio.uned.es/fez/view/bibliuned:DptoLenguaEsp-LingGral-Libros-Vmarrero>
- MacMahon, Michael K.C. (1991). The woman behind «Arthur». *Journal of the International Phonetic Association*, 21, 29–31. <https://doi.org/10.1017/S0025100300005983>
- Martínez-Celdrán, Eugenio, Fernández-Planas, Ana María y Carrera-Sabaté, Josefina (2003). Castilian Spanish. *Journal of the International Phonetic Association*, 33(2), 255–259. <https://doi.org/10.1017/S0025100303001273>.
- Martínez-Cifuentes, Renato, Torres-Bustos, Valeska y Sáez-Carrillo, Katia (2020). Textos utilizados en la evaluación de adultos chilenos con trastornos del habla de origen neurológico. *Revista de Logopedia, Foniatría y Audiología*, 40(2), 77–82. <https://doi.org/10.1016/j.rlfa.2019.11.002>.
- Melle, Natalia (2007). *Guía de intervención logopédica en la disartria*. Madrid: Síntesis.
- Monroy, Rafael y Hernández-Campoy, Juan Manuel (2015). Murcian Spanish. *Journal of the International Phonetic Association*, 45(2), 229–240. <https://doi.org/10.1017/S0025100314000231>
- Moreno Sandoval, Antonio, Torre, Doroteo, Curto, Natalia y de la Torre, Raúl (2006). Inventario de frecuencias fonémicas y silábicas del castellano espontáneo y escrito. En Luis Buera, Eduardo Lleida, Antonio Miguel y Alfonso Ortega (Eds.), *IV Jornadas en Tecnología del Habla* (pp. 77–81). Zaragoza: Universidad de Zaragoza; Red Temática en Tecnologías del Habla. Disponible en: <http://www.llif.uam.es/ESP/Publicaciones/LLI-UAM-4JTH.pdf>
- Moreno Sandoval, Antonio, Torre, Doroteo, de la Torre, Raúl, Garrote, Marta y Guirao, José M. (2008). Developing a phonemic and syllabic frequency inventory for spontaneous spoken Castilian Spanish and their comparison to text-based inventories. En Nicoletta Calzolari, Khalid Choukri, Bente Maegaard, Joseph Mariani, Jan Odijk, Stelios Piperidis y Daniel Tapias (Eds.), *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC 2008)*. Marrakech, Morocco, 28–30 May 2008 (pp. 1097–1100). European Language Resources Association (ELRA). Disponible en: <http://www.lrec-conf.org/proceedings/lrec2008/summaries/283.html>
- Muñiz, José, Elosua, Paula y Hambleton, Ronald K. (2013). Directrices para la traducción y adaptación de los tests: segunda edición. *Psicothema*, 25(2), 151–157. <https://doi.org/10.7334/psicothema2013.24>
- Nolan, Francis, McDougall, Kirsty, de Jong, Gea y Hudson, Toby (2009). The DyViS database: style-controlled recordings of 100 homogeneous speakers for forensic phonetic research. *International Journal of Speech Language and the Law*, 16(1), 31–57. <https://doi.org/10.1558/ijsl.v16i1.31>
- Ortega, Javier, González, Joaquín y Marrero, Victoria (2000). AHUMADA: A large speech corpus in Spanish for speaker characterization and identification. *Speech Communication*, 31(2–3), 255–264. [https://doi.org/10.1016/S0167-6393\(99\)00081-3](https://doi.org/10.1016/S0167-6393(99)00081-3)
- Patel, Rupal, Connaghan, Kathryn, Franco, Diana, Edsall, Erika, Forgit, Dory, Olsen, Laura, Ramage, Lianna, Tyler, Emily y Russell, Scott (2013). «The caterpillar»: a novel reading passage for assessment of motor speech disorders. *American journal of speech-language pathology*, 22(1), 1–9. [https://doi.org/10.1044/1058-0360\(2012/11-0134\)](https://doi.org/10.1044/1058-0360(2012/11-0134))
- Podesva, Robert y Zsiga, Elizabeth (2016). Sound recordings: acoustic and articulatory data. En Robert Podesva y Devyani Sharma (2016, 4.ª ed.). *Research Methods in Linguistics*. (pp.169-194). Cambridge: Cambridge University Press.
- Powell Thomas W. (2006). A comparison of English reading passages for elicitation of speech samples from clinical populations. *Clinical linguistics y phonetics*, 20(2-3), 91–97. <https://doi.org/10.1080/02699200400026488>
- REAL ACADEMIA ESPAÑOLA: Banco de datos (CORPES XXI) [en línea]. Corpus del Español del Siglo XXI (CORPES). <<http://www.rae.es>>
- RAE-ASALE (2011). *Nueva gramática de la lengua española. Fonética y fonología*. Madrid: Espasa Libros.
- Roca, Ignacio (1997). There are no glides, at least in Spanish. *Probus*, 9, 233–65.
- Rodero, Emma (2012). A comparative analysis of speech rate and perception in radio bulletins. *Text & Talk*, 32(3), 391–411. <https://doi.org/10.1515/text-2012-0019>
- Rodriguez, Francesco (2020). TraFo: transcriptor fonológico multidialectal del español. *Estudios de fonética experimental*, 29, 283–00, <https://raco.cat/index.php/EFE/article/view/375474>