

La vocal de relleno en español y en ruso: caracterización acústica e implicaciones teóricas

Vocalic fillers in Spanish and Russian: acoustic characterisation and theoretical implications

JUANA GIL FERNÁNDEZ ¹, JOSÉ MARÍA LAHOZ-BENGOECHEA ²,
JOSÉ VILLA VILLA ³

¹ Instituto Cervantes. Correo electrónico: dirlyo@cervantes.es, juanagilfernandez@gmail.com

² Universidad Complutense de Madrid. Correo electrónico: jmlahoz@ucm.es

³ Universidad Antonio de Nebrija. Correo electrónico: jvilla@nebrija.es, jovillavilla@hotmail.com

En el presente artículo, se revisan el concepto y las funciones de la vocal de relleno y se alude a los estudios que han abordado su estructura acústica en distintas lenguas, particularmente en español y en ruso, trabajos cuyos resultados permiten avanzar algunas hipótesis sobre la auténtica naturaleza de esta vocal. Para tratar de confirmar o falsar dichas hipótesis se ha analizado acústicamente una amplia muestra de vocales de las dos lenguas mencionadas. La discusión de los resultados de este análisis interlingüístico permite extraer algunas conclusiones potencialmente interesantes para la teoría fonético-fonológica. En concreto, se argumenta que la vocal de relleno puede constituir una unidad acústico-articulatoria claramente diferenciada de las vocales que integran los respectivos sistemas fonológicos y estrechamente relacionada con los ajustes articulatorios propios de cada lengua, esto es, con su base de articulación.

Palabras clave: vocal de relleno, fonética contrastiva contrastiva español / ruso, fonética acústica, ajustes articulatorios.

This paper addresses the concept and function of vocalic fillers, reviews other studies that have described their acoustic structure in different languages, especially focusing on those about Spanish and Russian, and outlines some hypotheses about the real nature of these vowels. Next, in order to check those hypotheses, an acoustic analysis is presented, based on large samples of fillers produced by speakers of the two aforementioned languages. The results of this crosslinguistic analysis are then discussed and conclusions are drawn which are potentially interesting for phonetic and phonological theory. Specifically, it is argued that vocalic fillers may constitute an acoustic and articulatory unit clearly differentiated from the vowels that form part of the respective phonological systems, and that their precise characteristics may be highly dependent on language-specific overall articulatory settings (namely, on the base of articulation).

Key words: Fillers, Spanish / Russian contrastive phonetics, acoustic phonetics, articulatory settings.

1. LA ‘VOCAL DE RELLENO’: CONCEPTO, SIGNIFICADOS Y DENOMINACIONES

Una de las características propias del habla espontánea es la presencia en ella de una amplia serie de fenómenos resultantes de la improvisación del discurso a la que el locutor se ve obligado: falsos comienzos, retrocesos, auto-correcciones, repeticiones, alargamientos a comienzo de la emisión, prolongaciones de vocales pre-existentes en final de palabra, pausas, etc. Todos ellos se suelen denominar conjuntamente con el término ‘disfluencias’, en cuanto que suponen una alteración de la fluidez del discurso, y su estudio se ha desarrollado considerablemente en las últimas décadas, como consecuencia del mayor relieve que las nuevas herramientas de análisis y los más abundantes recursos técnicos han permitido conceder al habla natural y no planificada. Las disfluencias, así pues, no se consideran ya elementos carentes de importancia, superfluos, una suerte de ‘ruido’ del mensaje totalmente descartable, sino que se asume que con frecuencia cumplen funciones comunicativas cuya exacta naturaleza está aún por precisar.

De entre todos los fenómenos mencionados, tal vez sean las pausas y sus tipos los aspectos más frecuentemente analizados (v. Goldman-Eisler 1968; Duez 1982, O’Connell y Kowal 1983; y, por lo que se refiere al español, el estado de la cuestión de Machuca, en prensa). Es norma general en los diversos trabajos publicados distinguir entre las pausas llamadas ‘llenas’ (en inglés, *filled pauses*) y las pausas ‘vacías’. Las primeras difieren de las segundas en que son sonoras, es decir, su espacio está ocupado por lo que se denomina en inglés *fillers*: sonidos aislados, prolongaciones de las palabras, vocalizaciones o incluso vocablos íntegros. Dicho de otro modo, en las pausas llenas el flujo del discurso se interrumpe¹ pero no la articulación, al contrario de lo que ocurre en las vacías, en las que sí se produce un silencio. El presente trabajo se centra, concretamente, en las vocales de límites precisos y claramente desgajadas del resto de la secuencia (de modo que no pueden considerarse como un alargamiento de otra vocal anterior ni como parte integrante de las palabras de su entorno) que aparecen en las pausas llenas. No se consideran aquí, por tanto, ni los murmullos nasales (‘mmm’), ni las agrupaciones constituidas por un sonido nasal con una vocal de apoyo (tipo ‘em’, ‘um’)² ni los vocablos completos u otras fórmulas (‘esto’, ‘este’, ‘ya sabes’, ‘o sea’, ‘cómo se dice’, etc.) que a veces se insertan en lugar de una pausa silenciosa.

Todos estos rellenos se conocen también con el nombre de *hesitaciones* (aplicado igualmente a otras disfluencias, como los alargamientos), puesto que parecen responder a la ‘duda’ del hablante acerca de cómo continuar su discurso, pero en realidad esta denominación

¹ Solo si se tiene presente esta idea, que la secuencia lógica del discurso se interrumpe, puede justificarse la denominación *pausa llena*, puesto que es obvio que no se produce ninguna pausa real, sino que el hablante continúa emitiendo sonido. No se trataría, pues, de una “pausa fónica”, sino de una “pausa discursiva”.

² De Leeuw (2007) analiza la preferencia de los hablantes de lenguas diferentes por emplear una vocal acompañada de una resonancia nasal o una vocal sola.

no alcanza a recubrir la variedad de significados que las pausas llenas pueden transmitir³ y que van más allá de la mera expresión de la vacilación del locutor (cf. Rose 1998; Arnold, Fagnano y Tanenhaus 2003; Nicholson, Eberhard y Scheutz 2010). En el caso concreto de los rellenos vocálicos, los trabajos publicados a este respecto se dividen, básicamente, entre a) los que estudian las causas que motivan la aparición de tales *fillers*; y b) los que analizan los efectos que los rellenos producen en los oyentes.

Como resultado de los estudios del primer tipo, se ha señalado que en ocasiones estos recursos reflejan un proceso cognitivo de tipo metalingüístico que el hablante experimenta (necesita tiempo para encontrar el término o la expresión más apropiados para lo que desea transmitir, especialmente si está tratando un tema complejo o en el que se ve implicado un vocabulario amplio o no habitual, v. Beattie y Butterworth 1979; Berger, Karol y Jordan 1989; Schachter et al. 1991; Brennan y Williams 1995; Merlo y Mansur 2004), o bien constituyen una táctica interaccional (cf. Navarretta 2015) con dos posibles objetivos opuestos: evitar la presencia de silencios que puedan hacer perder al hablante el uso de la palabra en beneficio de su interlocutor (Maclay y Osgood 1959; Fox Tree 2002), o sugerir sutilmente a este último que complete el enunciado que al locutor se le hace difícil concluir (Clark y Fox Tree 2002). Las vocales de relleno pueden así entenderse, empleando la terminología adoptada de De Leuw (2007) como ‘síntomas’ de procesos mentales no intencionados, o como ‘señales’ o procedimientos retóricos que el hablante emplea, y controla, para indicar algo a su interlocutor (v., asimismo, Corley y Stewart 2008).

En cuanto a las investigaciones abordadas desde la perspectiva del oyente, los trabajos han tratado de elucidar el modo –automático o, por el contrario, consciente– en el que este filtra los *fillers* durante el procesamiento perceptivo de la secuencia (Christenfeld 1995), o bien han girado en torno a la medida en que el empleo de estos rellenos afecta a la imagen que la audiencia se forma del locutor (Fox Tree 2002; Barr y Seyfeddinipur 2010), tanto en el caso de los hablantes nativos de una lengua, como en el de los aprendices de lenguas extranjeras (Woss 1979).

Finalmente, otros estudios han examinado la preferencia por un tipo u otro de relleno o las reglas que rigen su uso (cf. De Leuw 2007; Clark y Fox Tree 2002), su distribución en la cadena hablada (Shriberg 1994; O’Connell y Kowal 2005) y su frecuencia de aparición, no solo en la lengua materna sino también en la interlengua de los aprendices de L2 (Dewaele 1996; Cenoz 2000; Wu 2008; García Amaya 2015; de Jong 2016), o bien han analizado los *fillers* con vistas a la caracterización idiolectal de los hablantes para fines forenses (Künzel 1997; Brander 2014, Cicres 2014) o para precisar su posible impacto en el rendimiento de los sistemas de reconocimiento automático del habla (Shriberg 2001; Vasilescu, Rosset y Adda-Decker 2010).

³ Aunque habitualmente se piensa que las pausas vacías y llenas sirven a los mismos intereses comunicativos, algunos autores (por ejemplo, Campione y Vèronis 2005) defienden que ambos tipos desempeñan papeles distintos, y que las llenas son las únicas que pueden expresar realmente duda, al igual que los alargamientos, con los que poseen más rasgos funcionales en común que con las pausas vacías.

Se desprende, pues, de esta brevísima síntesis⁴ de la bibliografía que, según apuntábamos antes, denominar ‘vocales de duda’ (o hesitaciones) a los elementos vocálicos que rellenan las pausas, como a veces se hace, conlleva reducir drásticamente el abanico de funciones significativas y de implicaturas conversacionales y pragmáticas que comportan, por lo que en el presente trabajo nos referiremos a ellas siempre con una expresión no alusiva a su significado, como es la de ‘vocales de relleno’.

2. ESTUDIOS DE CARÁCTER ACÚSTICO SOBRE LA VOCAL DE RELLENO: EL CASO DEL ESPAÑOL Y DEL RUSO

Como se ha dicho, las vocales de relleno han generado una considerable cantidad de títulos bibliográficos, que, en su mayor parte, las abordan con el objeto de estudiar los aspectos pragmáticos, discursivos, psicolingüísticos, de adquisición de segundas lenguas, etc., mencionados en el apartado anterior⁵; son escasos, en cambio, los trabajos que se centran específicamente en el análisis de su naturaleza acústica. Excepciones a esta tendencia general son, entre otros, los estudios sobre el alemán (Pätzold y Simpson 1995), el catalán (Cicres 2014), el estonio (Nemoto 2013), el francés (Duez 2001), el inglés (Shriberg 2001), el italiano (Giannini 2003), el portugués (Proença et al. 2013) y algunos otros trabajos de conjunto sobre varias lenguas como los de Candea, Vasilescu y Adda-Decker (2005), Nemoto (2006), Vasilescu y Adda-Decker (2007) Vasilescu, Nemoto y Adda-Decker (2007) y Spreafico (2016). Particularmente interesantes desde nuestro punto de vista son las investigaciones de Silber-Varod, Weiss y Amir (2015) y de Silber-Varod y Amir (2016) sobre los *fillers* en hebreo, puesto que el sistema fonológico de esta lengua comparte con el del español bastantes características: ambos poseen los mismos cinco fonemas vocálicos /a e i o u / integrados en subsistemas en los que ni la duración ni la tensión son distintivas, y que no incorporan una schwa /ə/. Además, los *fillers* en hebreo, como en español, se identifican, al menos perceptivamente, con la vocal anterior media /e/ (véase, más adelante, la nota 10).

⁴ El lector interesado puede encontrar una bibliografía sumamente completa en la página del *Filled Pause Research Center*, <http://filledpause.com>, creada por Ralph Rose y actualizada permanentemente. Véase asimismo Rose (2013).

⁵ A este tipo de trabajos pertenecen, además de los ya citados y entre muchos otros mencionables, los de Riazantseva (2001) sobre el ruso, Watanabe (2003) sobre el japonés, Trofimovich y Baker (2006) sobre el coreano, Tsiaras, Panagiotakis y Stylianou (2010) sobre el griego, Horváth (2011) sobre el húngaro, Urizar y Samuel (2013) sobre el vasco, Wieling et al. (2016) sobre las lenguas germánicas en su conjunto, o Yuan et al. (2016) sobre el chino mandarín. Remitimos de nuevo a <http://filledpause.com> para más referencias sobre distintas lenguas.

En lo que se refiere, concretamente, al ruso y al español, que son las lenguas en las que se centra el presente artículo, no existen muchas investigaciones previas sobre la cuestión. En el caso del ruso, el trabajo de carácter acústico más repetidamente citado es el de Stepanova (2007). Esta autora analizó la duración y los dos primeros formantes de aproximadamente 200 vocales de relleno extraídas de las grabaciones de habla espontánea realizadas a 5 mujeres y 5 hombres de San Petersburgo, y comparó los valores obtenidos con los correspondientes a alrededor de 300 vocales léxicas tónicas /e/ y /a/, puesto que, por lo general, los oyentes asimilan perceptivamente el *filler* empleado en ruso a una de estas dos vocales. Los resultados de su análisis muestran, en cambio, que, desde el punto de vista acústico, las vocales de relleno difieren tanto de la vocal léxica /e/ como de la vocal /a/, o bien en los valores del F1 o bien en los del F2, o bien en los de ambos formantes: se diría que en todos los hablantes registrados (7 en total, porque se eliminaron 3 del estudio) el elemento vocálico que aparece en las pausas sonoras es un sonido más cercano a /a/ o a /e/ dependiendo del locutor, pero distinto de cualquiera de los dos.

Por lo que respecta al español, algunos estudios, como el de Rebollo (1997) o el de Machuca, Llisterri y Ríos (2015), que aportan datos interesantes⁶, no presentan sin embargo análisis acústicos de las vocales, los que sí ofrecen, en cambio, otras publicaciones, como las de Candea, Vasilescu y Adda-Decker (2005), Vasilescu y Adda-Decker (2007), Vasilescu, Nemoto y Adda-Decker (2007) y Machuca y Ríos (2016). De estos, en los tres primeros se subraya la coincidencia o ‘casi coincidencia’ acústica de la vocal de relleno del español con la vocal léxica /e/, si bien se afirma que los corpus utilizados pertenecen a la variedad del español “latinoamericano” (Candea, Vasilescu y Adda-Decker 2005), o a la variante “europea” (Vasilescu y Adda-Decker 2007; Vasilescu, Nemoto y Adda-Decker 2007), etiquetas que resultan excesivamente amplias y que se prestan, por tanto, a generalizaciones incorrectas.

El estudio más reciente, Machuca y Ríos (2016), por su parte, analiza los primeros y segundos formantes acústicos de 68 vocales de relleno, y compara sus valores con los correspondientes tomados de 96 vocales [e] léxicas átonas y 94 tónicas. En esta ocasión, el corpus empleado está constituido por las grabaciones de habla espontánea de cinco informantes masculinos, aunque no se precisa si todos ellos pertenecen a la misma variante del español peninsular. Los resultados obtenidos indican que los valores de frecuencia del F1 de la vocal de relleno tienden a ser los de una vocal media anterior, mientras que los de su F2 (con una media de 1800 Hz aproximadamente) son más elevados que los de las vocales léxicas en habla espontánea, más próximos por tanto a los de estas mismas vocales en habla

⁶ Tanto en uno como en otro la vocal de relleno del español se identifica con /e/, y, en el caso de Machuca, Llisterri y Ríos (2015), se precisa, además, a partir del análisis de 156 pausas sonoras extraídas de los registros de habla espontánea de 16 hablantes, que la realización [e:] se emplea en el 61 % de las ocasiones; el murmullo nasal [m:], en el 32 %, y el 7 % restante corresponde a vocalizaciones variadas. Moreno Sandoval, Campillos y Toledano (2012) realizaron un estudio cuantitativo sobre las disfluencias del español en general, y en él, aunque en realidad no es un trabajo de naturaleza propiamente fonética, se alude también a la vocal de relleno identificándola con <eh>.

de laboratorio. A partir de tales datos, los autores concluyen que “la estructura formántica de las pausas sonoras no es comparable con la de las vocales extraídas de un corpus de habla espontánea, ya que estas últimas sufren un fenómeno de centralización” (74). Es decir, la diferencia observada entre las vocales de relleno y las léxicas, tanto tónicas como átonas, con las que se las compara se debería a que estas últimas están extraídas de registros en los que los hablantes se expresan de modo natural, hipoarticulado, lo que ocasiona su centralización y reducción temporal. Se desprende de todo ello que las vocales de relleno [e:] serían más bien acústicamente similares a las vocales [e] en habla hiperarticulada de laboratorio.

Considerados todos estos precedentes, en el presente trabajo nos propusimos analizar la vocal de relleno del ruso y del español comparándolas con sus vocales léxicas de referencia a partir de a) un número más elevado de informantes, de manera que los resultados alcanzaran una mayor significatividad, y b) registros no solo de habla espontánea (para conseguir la cantidad adecuada de *fillers*), sino también de lectura de fragmentos de textos especialmente seleccionados para este objetivo (en los que las vocales léxicas tónicas se pronuncian de manera hiperarticulada y sin acortamientos ni centralización marcados).

3. EL TIMBRE DE LAS VOCALES DE RELLENO: ALGUNAS HIPÓTESIS

Con base en las conclusiones de las investigaciones llevadas a cabo hasta el momento, es posible enumerar algunos aspectos en los que las vocales de relleno de las diversas lenguas parecen coincidir. En primer lugar, resultan ser más largas que la(s) vocal(es) léxica(s) similar(es), como señalan, entre muchos otros, Clark y Fox Tree (2002) o Shriberg (2001); además, el contorno que la f_0 presenta en ellas es bastante estable, o, en todo caso, propende a descender (véanse Quimbo, Kawahara y Doshita 1998; Goto, Itou y Hayamizu. 1999; Campione y Véronis 2005; Belz y Reichel 2015), en ocasiones con valores tan bajos que pueden desembocar en un registro *creak* (cf. Shriberg 1999; Candeia, Vasilescu y Adda-Decker 2005, Horne 2009, Maekawa y Mori 2015); finalmente, en numerosas lenguas –aunque no en todas las analizadas⁷– se documenta una tendencia a la centralización de su timbre, en relación inversa con su duración, es decir, cuanto más larga es la vocal de relleno, menos centralizada se pronuncia.

No obstante, al lado de estas características compartidas, se han señalado otros rasgos que distinguen a las vocales de relleno interlingüísticamente: el más obvio es que su timbre es diferente en las diversas lenguas, y se percibe por los oyentes de modo también claramente diferenciado, como lo prueba la propia representación ortográfica que se les suele dar (*er/uh* en inglés, *euh* en francés, *äh* en alemán, *øh* en danés, etc., v. Vasilescu y Adda-Decker 2013). Se ha comprobado asimismo que los oyentes pueden identificar una

⁷ Esta tendencia, documentada para el árabe, el chino mandarín, el francés, el alemán, el portugués europeo y el inglés americano por Candeia, Vasilescu y Adda-Decker (2005), no caracteriza según estas mismas autoras, al español latinoamericano (*sic*).

lengua a partir tan solo de la vocal de relleno que escuchan (Vasilescu, Candea y Adda-Decker 2005), de forma que no puede negarse que son sonidos perceptivamente disímiles.

Tanto los rasgos que se han demostrado comunes como los específicos de cada lengua subyacen en los argumentos que se han ido avanzando respecto a dos cuestiones fundamentales: a) si la vocal de relleno basa su cualidad en el grado de esfuerzo requerido para pronunciarla, y b) si el timbre de la vocal de relleno de una lengua coincide siempre con el de una de las vocales que integran su sistema fonológico.

Relacionar los *fillers* con el esfuerzo articulatorio presuponiendo que, por su carácter de relleno, serán sonidos que impliquen escasos movimientos articulatorios y demanden poca energía muscular, como hicieron Pätfold y Simpson (1995) en su estudio sobre el alemán, sustentaría el hecho de que en muchas lenguas se pronuncie una vocal centralizada tipo schwa (O'Saughnessy 1993), cuya producción naturalmente implica el menor costo en este sentido. La ya citada Stepanova (2007), en referencia al ruso, incide en esta línea de razonamiento al apuntar que el sonido resultante puede deberse a que los hablantes “produce sounds that are as neutral as possible, demanding minimal movements of tongue and lips” (1327). Esta hipótesis, sin embargo, no encaja con el hecho de que la cualidad de la vocal varía de unas lenguas a otras y no siempre tiende a la centralización como sería esperable si la economía fuera la razón última y universal de su aparición.

Habida cuenta de estas contradicciones, muchos otros autores se han inclinado por justificar la cualidad de la vocal de relleno en función del inventario fonológico de la lengua en cuestión, a partir del presupuesto de que los hablantes emplean una de sus unidades vocálicas como *filler* (Levelt 1989; Shriberg 1994: 24-25; Giannini 2003: 2656). No obstante, esta hipótesis también se enfrenta con problemas, porque ya se ha comprobado que en algunas lenguas (o dialectos) la vocal de relleno no se corresponde con ningún sonido de su inventario (por ejemplo, en el caso mencionado por Giannini del italiano del sur, cuyo sistema no contiene una schwa, siendo esta, en cambio, la realización del *filler*; o el estudio sobre el francés de Canadá de Seguin 2008), y porque, aun cuando exista tal correspondencia, el grado de similitud entre ambos elementos varía de una lengua a otra (Vasilescu, Nemoto y Decker 2007).

La hipótesis de la que se parte en el presente trabajo se aparta de los dos planteamientos que se acaban de exponer, y supone que la vocal de relleno refleja en realidad la ‘postura’ o el ajuste articulatorio preferido por cada lengua, esto es, la configuración del tracto vocal que constituye el anclaje en torno al cual gira todo su entramado fónico. No se trataría, por tanto, de que la vocal de relleno haya de coincidir con una posición de reposo universal que entrañe poco esfuerzo articulatorio, sino de que tal posición en cada lengua es diferente porque se aviene de manera natural con la postura más ‘cómoda’ para sus hablantes, la cual es el ajuste articulatorio característico y propio del idioma en cuestión. De igual forma, la vocal de relleno no tendría por qué coincidir con una vocal del inventario fonológico y, de hacerlo, más bien podría hablarse de que es la vocal la que se acerca a la postura por defecto, en lugar de identificar a esta última previamente con la vocal.

Con el objetivo de comprobar lo que sucede en dos lenguas tipológicamente muy dispares y en qué medida esos hechos contribuyen a falsar o confirmar esta hipótesis, hemos

procedido a estudiar, en fases sucesivas, la vocal de relleno en castellano centro-peninsular y en ruso⁸.

4. LAS VOCALES DE RELLENO EN CASTELLANO Y EN RUSO: ESTUDIO ACÚSTICO CONTRASTIVO

4.1. *La vocal de relleno en castellano*

En un trabajo previo sobre la vocal de relleno en la variedad castellana centro-peninsular del español (cf. Villa, Gil y Lahoz-Bengoechea, 2017) analizamos las vocales de este tipo extraídas de las grabaciones de habla espontánea de 50 informantes masculinos, cuyos registros constituyen una parte del corpus *CIVIL*, elaborado durante los años 2010-2013 en el Laboratorio de Fonética del Consejo Superior de Investigaciones Científicas, en Madrid⁹. Nuestra finalidad fue comprobar, en primer lugar, cuáles eran las características acústicas de la vocal articulada como relleno por los distintos locutores y, en segundo lugar, comprobar en qué medida tales rasgos coincidían con los de la vocal /e/ léxica del español, partiendo de las múltiples observaciones documentadas en la bibliografía acerca de la evidente similitud entre ambos elementos (§ 3). Para soslayar la influencia del estilo de habla espontáneo en las características de las vocales léxicas, estas se extrajeron de los fragmentos de lectura también incluidos en el corpus mencionado. Analizamos en este primer estudio todos los parámetros (duración, frecuencia fundamental, los tres primeros formantes y las distancias existentes entre los formantes y entre el primer formante y la frecuencia fundamental) que se enumeran en la Tabla 1, en la cual se reproducen los resultados obtenidos con respecto a cada uno de ellos para los dos tipos de vocal. Los parámetros con subíndices diferentes son aquellos que presentan diferencias significativas entre la vocal /e/ léxica y la de relleno, lo cual se refleja igualmente, pero con mayor grado de detalle, en la Tabla 2. En el gráfico de dispersión que se ofrece en la Figura 1 se pueden apreciar también con gran claridad las similitudes y las diferencias acústicas encontradas entre los dos tipos de vocal.

⁸ La elección del ruso y no de otra lengua también alejada del español se debe a que la presente investigación se enmarca en el proyecto de fonética forense denominado “EMULANDO: Estudios Multilingües del Acento No nativo Disimulado, financiado por el Ministerio de Economía español” (Ref. FFI2014-59848-C2-1-P). El foco de interés de este proyecto se centra en la imitación con propósitos delictivos del acento ruso, inglés o francés por hispanohablantes, y, de las tres lenguas que se ven implicadas, el ruso es, sin duda, la que menos se ha estudiado en relación con el español y, en este sentido, la que más investigación requiere.

⁹ El corpus *CIVIL* (*Cualidad Individual de la Voz e Identificación del Locutor*) está integrado por las grabaciones de 100 locutores (50 masculinos y 50 femeninos) de entre 20 y 35 años, todos ellos hablantes de la variante castellana del español. Cada locutor efectuó una tarea de lectura y otra de habla espontánea en modo de fonación modal y en registros *falsetto* y *creak*. A todos se los grabó en 2 sesiones idénticas de aproximadamente 10 minutos realizadas en días diferentes en la cabina insonorizada del Laboratorio siempre con el mismo equipo. Más detalles sobre *CIVIL* en Alves et al. (2014).

A partir de todos estos datos, en el estudio mencionado (Villa, Gil y Lahoz-Bengoechea, 2017) concluimos que el timbre de la vocal de relleno, en castellano centro-peninsular, se asemeja sin duda alguna al de la vocal léxica /e/, como era esperable, aunque ambas son *significativamente* distintas. No solo las de relleno son claramente más largas –una media de 463 ms frente a los 56 ms de las léxicas tónicas–, sino que su F2 se mueve en torno a unos valores frecuenciales sistemáticamente más elevados que el F2 de las vocales léxicas

	Tipo de vocal			
	Léxica		Relleno	
	Media	Desviación estándar	Media	Desviación estándar
Duración	56 _a	19 _c	463 _b	219 _d
f_0	117 _a	19 _c	113 _a	44 _d
F1	467 _a	54 _c	473 _a	49 _c
F2	1707 _a	164 _c	1903 _b	188 _c
F3	2536 _a	190 _c	2578 _b	160 _c
F1 - f_0	349 _a	55 _c	360 _a	65 _d
F2 - F1	1240 _a	152 _c	1431 _b	202 _c
F3 - F2	829 _a	187 _c	674 _b	192 _c

Tabla 1: Descripción acústica comparada de la vocal de relleno y la vocal léxica /e/ en español.

Efecto del tipo de vocal sobre el valor medio de las distintas variables	
Variación significativa	Variación no significativa
Duración t (208.14) = 26.60, p < 0.001	f_0 t (276.99) = -1.36, ns
F2 t (402.41) = 11.33, p < 0.001	F1 t (407.04) = 1.19, ns
F3 t (398.03) = 2.42, p < 0.05	F1 - f_0 t (397.61) = 1.78, ns
F2 - F1 t (380.87) = 10.82, p < 0.001	
F3 - F2 t (409.73) = -8.29, p < 0.001	

Tabla 2: Variaciones acústicas entre la vocal de relleno y la vocal léxica /e/ en español.

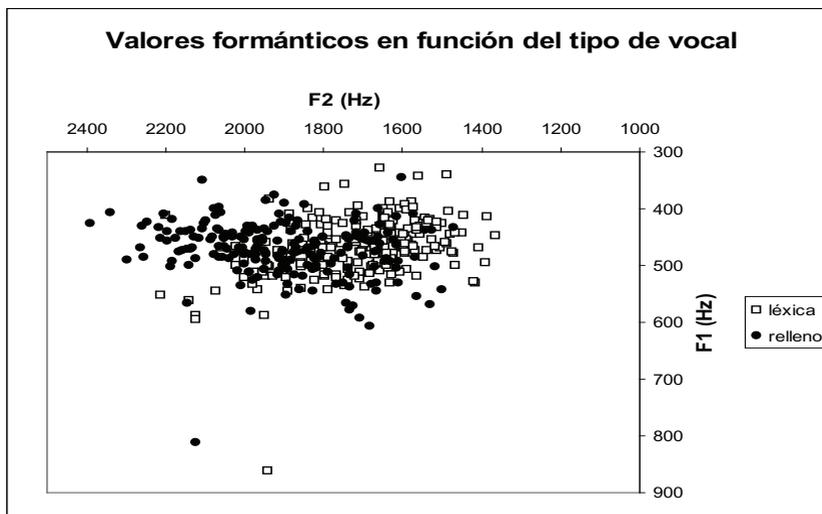


Figura 1: Dispersión de los valores de F1 y de F2 de la vocal de relleno y de la vocal léxica /e/ en español.

(con una media de 1903 Hz en los *fillers* frente a un valor medio de 1713 Hz en las léxicas), lo cual significa que las vocales de relleno son, en general, más adelantadas que las léxicas. Su grado de abertura es, no obstante, coincidente con el de la vocal léxica /e/. La vocal de relleno está sometida, aun así, a un grado mayor de varianza que la vocal léxica, como reflejan los parámetros de la Tabla 1, especialmente el referido a la duración. En el presente trabajo, teniendo en cuenta estos datos previos alcanzados sobre el español, abordamos el análisis de las vocales del ruso con el procedimiento y los resultados que se exponen a continuación.

4.2. La vocal de relleno en ruso: metodología del estudio

Se grabó en el Laboratorio del Departamento de Fonética de la Universidad de San Petersburgo a 30 locutores nativos, todos ellos masculinos, nacidos y residentes en la ciudad, hablantes de la misma variante dialectal del ruso, con edades comprendidas entre los 18 y los 45 años y sin ninguna patología articulatoria ni auditiva. En las sesiones de grabación todos ellos tuvieron que realizar dos tipos de tareas.

La primera consistió en la lectura de 13 frases en ruso, preparadas especialmente para este propósito de modo que incorporaran vocales /a/ y /e/ léxicas en posición tónica; en cada oración se anotó un ejemplar de /e/ tónica y otro de /a/ tónica, de forma que se obtuvieron 13 casos de cada una de ellas. La segunda tarea, concebida para conseguir que

los informantes articularan un número suficiente de vocales de relleno, se subdividía en tres sub-tareas: en primer lugar, se les presentaba en la pantalla del ordenador una serie abundante de dibujos correspondientes a objetos cotidianos, de los que debían enumerar la mayor cantidad posible en un tiempo limitado, lo que inevitablemente provocó que surgiera un buen número de pausas llenas; en segundo lugar, se les proponía que ofrecieran argumentos convincentes a favor de ciertas afirmaciones ‘heterodoxas’, como por ejemplo “fumar es bueno” o “el deporte perjudica”, para provocar que el esfuerzo cognitivo que deberían realizar fuera mayor que si se tratara de aseveraciones de amplio consenso social; finalmente, se les proponía que hablaran sobre un tema libre de conversación relacionado con sus aficiones favoritas, sus hábitos, sus estudios, etc.

En total, se midieron y etiquetaron 367 vocales /e/ tónicas léxicas y 390 vocales /a/ tónicas léxicas. La segmentación y el etiquetado se realizaron de forma manual, mientras que la extracción de los valores acústicos fue automática, en el punto central de la duración de cada vocal, mediante un *script* de Praat (Boersma y Weenink 2017), si bien las cifras obtenidas se revisaron después a mano para corregir los datos que se salían de los rangos esperados. En el curso de esta revisión posterior hubo que descartar algunos casos de /e/ porque eran extremadamente breves, por tanto difícilmente segmentables y analizables. Ello explica que, en el caso de esta vocal, no se alcanzara el número de ejemplares previsto en el diseño experimental (390).

Por lo que respecta a las vocales de relleno, se optó –al igual que se había hecho en el estudio sobre el español– por considerar como tales solo aquellos sonidos claramente separados del resto de la secuencia que presentaran una duración igual o superior a 200 ms. Esta decisión se tomó por dos razones fundamentales: porque hacerlo así es una práctica relativamente habitual en la bibliografía (cf. por ejemplo Candea, Vasilescu y Adda-Decker 2005) y porque, como comprobaron Machuca, Llisterrí y Ríos (2015) en los experimentos que realizaron, pausas sonoras de hasta 156 ms pueden pasar inadvertidas por el oyente, de manera que el umbral de percepción para este tipo de elementos parece ser bastante alto. Siguiendo estos criterios, se obtuvieron, se etiquetaron y se midieron 660 vocales de relleno.

Los parámetros considerados en cada vocal fueron también los mismos que se analizaron en Villa, Gil y Lahoz-Bengoechea, 2017 para el castellano: duración, frecuencia fundamental, los tres primeros formantes y las distancias existentes entre los formantes y entre el primer formante y la frecuencia fundamental.

4.3. *La vocal de relleno en ruso: Resultados*

En la Tabla 3 se presentan los promedios de todas las mediciones llevadas a cabo en el estudio de las vocales de relleno del ruso, comparadas con las realizaciones tónicas de las vocales /e/ y /a/ léxicas de esa lengua. Como ocurría en la Tabla 1 con respecto al castellano, los parámetros con subíndices distintos son aquellos que presentan diferencias significativas entre los distintos tipos de vocal. En la Figura 2, por otra parte, se reproduce la carta de formantes con los valores de F1 y F2 de los tres tipos de vocal.

Tabla 3: Descripción acústica comparada de la vocal de relleno y de las vocales léxicas /a/ y /e/ en ruso.

	/e/		/a/		Relleno	
	Media	Desviación estándar	Media	Desviación estándar	Media	Desviación estándar
Duración	70 _a	23	78 _a	17	468 _b	210
f_0	132 _a	26	142 _b	35	105 _c	20
F1	451 _a	59	571 _b	62	519 _c	104
F2	1760 _a	200	1367 _b	233	1222 _c	172
F3	2478 _a	208	2419 _b	218	2366 _c	225
$F1 - f_0$	319 _a	66	430 _b	65	416 _c	110
F2 - F1	1309 _a	227	796 _b	258	703 _c	179
F3 - F2	718 _a	201	1052 _b	273	1144 _c	238

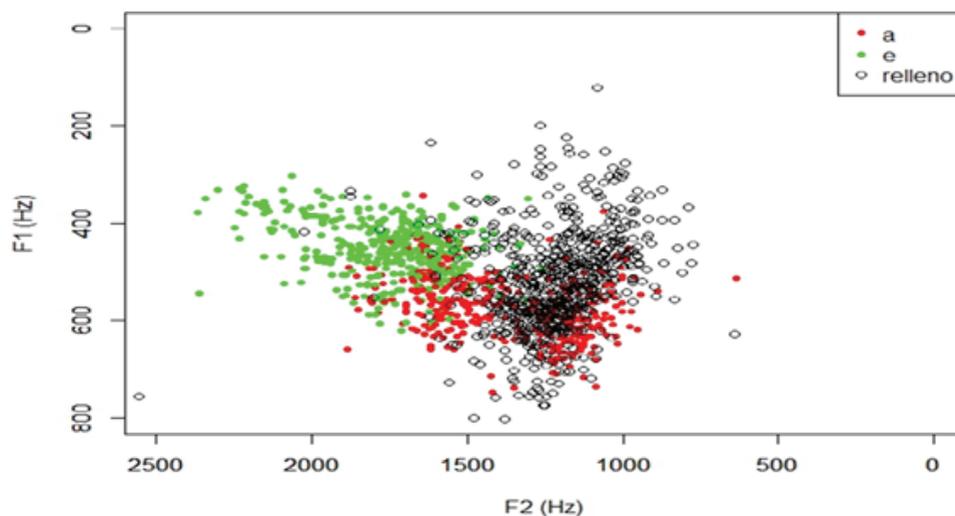


Figura 2: Dispersión de los valores de F1 y de F2 de la vocal de relleno y de las vocales léxicas /a/ y /e/ en ruso.

Exactamente igual que sucedía en castellano, la vocal de relleno presenta una duración mucho mayor que las dos vocales léxicas (que, sin embargo, arrojan valores similares entre sí a este respecto), pero, a diferencia de lo que se comprobó para el español, su frecuencia fundamental es considerablemente más baja. También difieren entre sí las tres vocales en cuanto a los valores del F1 y del F2 (y, consecuentemente, en todas las otras mediciones que los implican), según puede apreciarse con más claridad en los gráficos de las Figuras 3 y 4, respectivamente.

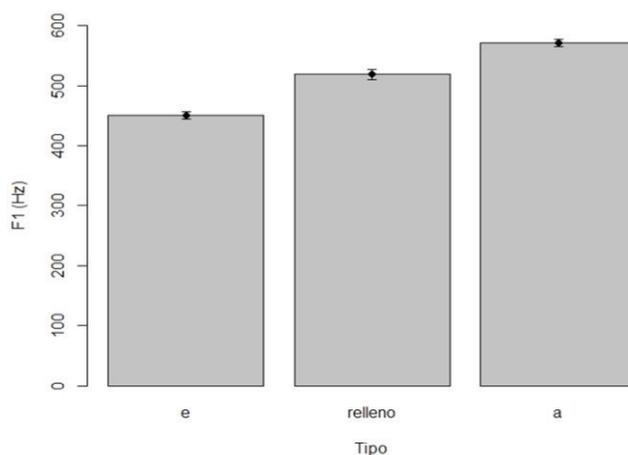


Figura 3: Comparación de los valores de F1 de la vocal de relleno y de las vocales /a/ y /e/ en ruso

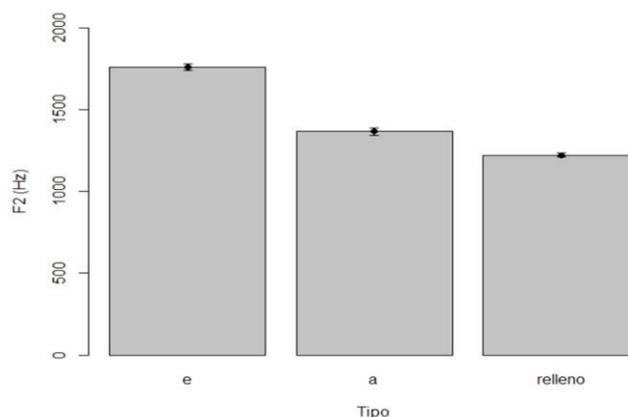


Figura 4: Comparación de los valores de F2 de la vocal de relleno y de las vocales léxicas /a/ y /e/ en ruso

Se aprecia, pues, a partir de los valores medios de F1, que la vocal de relleno es un poco más cerrada que la vocal léxica [a], pero un poco más abierta que la [e], mientras que de los valores de F2 se desprende que su zona de articulación es más retrasada que la de las otras dos vocales de referencia. La variabilidad que presentan las cifras en el caso de la vocal de relleno es mayor en relación con el F1. En definitiva, lo que resulta sumamente interesante es que los tres tipos de sonidos vocálicos difieren significativamente entre sí en lo que se refiere tanto a la duración como al timbre.

4.4. Comparación de los datos sobre la vocal de relleno en castellano y en ruso

Una vez comprobado que ni en la variante castellana del español ni en la variante del ruso hablada en San Petersburgo las vocales de duda coinciden acústicamente con las vocales léxicas con las cuales se las suele identificar, llevamos a cabo la comparación interlingüística entre las realizaciones de la vocal de relleno en ambas lenguas. En la Figura 5 se reproduce el gráfico de dispersión que muestra cómo es la distribución de estos elementos y la distancia clara que media entre los *fillers* de ambos idiomas.

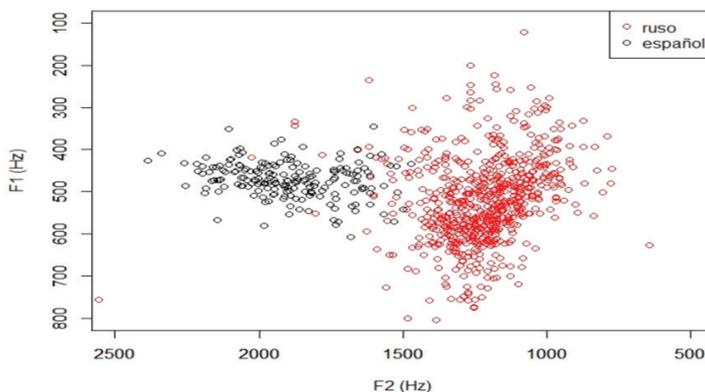


Figura 5: Gráfico de dispersión de los valores formánticos de la vocal de relleno en español y en ruso

Aunque básicamente coinciden por lo que respecta a su duración y a su frecuencia fundamental, es evidente que se trata de dos sonidos distintos: en el caso del español castellano, tenemos un sonido más anterior que [e] y con un grado de abertura similar al de esta última; en el caso de la variante del ruso estudiada, un sonido más posterior que la [e] e incluso que la [a] y, desde luego, mucho más retrasado que el *filler* español, también con un mayor grado de abertura que este último (véase la Tabla 4 para cotejar los valores medios concretos de los parámetros más relevantes).

	Vocal de relleno (ruso S. P.)		Vocal de relleno (castellano)	
	<i>Media</i>	<i>Desviación estándar</i>	<i>Media</i>	<i>Desviación estándar</i>
Duración	468	210	463	219
f_0	105	20	113	44
F1	519	104	473	49
F2	1222	172	1903	188
F3	2366	225	2578	160

Tabla 4: Comparación de los valores acústicos de la vocal de relleno en ruso (variedad de San Petersburgo) y en español (variedad castellana).

5. DISCUSIÓN DE LOS RESULTADOS

A la espera de nuevas y más amplias indagaciones sobre otras lenguas, los datos obtenidos en el presente trabajo con respecto al castellano y al ruso confirman que la vocal de relleno no se corresponde con una posición de reposo pretendidamente universal, puesto que difiere claramente en las dos lenguas (cf. lo apuntado en el § 3 a propósito de otros idiomas).

Es posible que, en efecto, exista una posición universal que pudiéramos llamar también ‘de reposo’ o ‘de descanso’, pero se tratará en todo caso de la posición de inactividad completa de los órganos articulatorios en un estado de inercia fisiológica, cuando no hay habla, y no de la que aquí nos interesa, la *Sprachliche oder active Normalstand der Organe*, o, en palabras de uno de los primeros fonetistas que clarificaron esta distinción, Wilhelm Scherer (1868: 23, cit. en Kedrova y Borisoff 2013: 3), “el estado normal y activo de los órganos al cual estos regresan más fácil y cómodamente durante su actividad” (traducción nuestra). Amado Alonso (1974) explicó perfectamente esta diferencia hace ya muchos años:

En las inscripciones quimográficas, ya preparado el sujeto de la experiencia con la bocina colocada, pero sin haber iniciado la pronunciación, la plumilla marca sobre el cilindro una línea recta (posición de descanso). Pero unas 40 centésimas de segundo antes de iniciar la pronunciación, la plumilla da un salto de un par de milímetros, aunque la línea sigue siendo recta en su nueva altura hasta la aparición del primer fonema. La membrana del tamborcillo ha sufrido un minúsculo desplazamiento (no debido a salida alguna del soplo) que se mantiene invariable hasta la aparición de las articulaciones, y éstas se marcan sobre el papel ahumado partiendo de esta nueva altura y no de la de total descanso. ¿Qué significa ese desplazamiento? Hasta ese instante el sujeto tenía los órganos articuladores en reposo; en ese instante

los pone en disposición de hablar. Esa disposición especial de los órganos articuladores coincide con la bien conocida ‘posición de trabajo’ de los órganos respiratorios y de los fonadores. (...) La base fonética comienza no en la posición de reposo, sino en la disposición de trabajo que adoptan en cada comunidad lingüística los órganos articuladores. Desde esa específica disposición de los órganos se desarrolla la cadena de articulaciones según un modo también específico de memoria motriz. (Alonso 1974: 263-4).

También hace tiempo, Hartmann y Stork (1972) en su diccionario ya diferenciaron ambas nociones y Öhman (1967) proporcionó, más o menos por la misma época, las primeras pruebas cuantitativas, con respecto a algunos músculos articuladores, de que tal distinción tenía sentido.¹⁰ Mucho más recientemente, Gick et al. (2005), Wilson (2006) y Schaeffler, Scobbie y Mennen (2008) han estudiado las ‘posturas de trabajo’ de varias lenguas y han conseguido establecer mediante técnicas objetivas (sistemas de captura de movimientos en 3D por ultrasonidos, cineradiografía, etc.) que las –así llamadas por todos ellos– *inter-speech postures* divergen, en efecto y significativamente, al menos entre la variedad canadiense del inglés y del francés.

A partir de estos precedentes, podemos pensar que los *fillers* coinciden, no con la posición de reposo, sino con la posición de trabajo, lo que explicaría no solo el hecho de que no sean universales, sino asimismo que no se hayan visto tampoco aquí corroboradas las otras hipótesis mencionadas en el § 3.

Es el caso de la avanzada por Pätzold y Simpson (1995), según la cual los *fillers* implican un menor esfuerzo articulatorio, que demanda mínimos movimientos de los órganos articuladores y produce la realización más cómoda posible, ya sea para todos los locutores, ya sea –si nos atenemos a la reflexión suscrita por Stepanova (2007)– para cada hablante. Hemos comprobado, por el contrario, que la vocal de relleno del español es una vocal anterior más periférica que la vocal /e/, por lo que los gestos articulatorios a los que responde demandan en principio un mayor desarrollo en el espacio y en el tiempo, es decir que, para generar el gesto de constricción que la produce, la lengua se tiene que desplazar más y durante más tiempo de su posición de reposo, lo cual lógicamente comporta un incremento en el esfuerzo articulatorio global asociado con el segmento. En consecuencia, cabe pensar que el empleo de una determinada vocal de relleno no deja traslucir una tendencia inconsciente del hablante a esforzarse menos, sino una preferencia natural de ese mismo hablante por una determinada posición articulatoria, que constituye el reflejo de un ajuste articulatorio básico de su lengua en cuestión, esto es, de la posición de trabajo

¹⁰ Brian Annan (1972), por su parte, llamó la atención sobre el hecho de que la postura de reposo durante el habla (la de trabajo, en los términos de Alonso, que podemos adoptar aquí) no es común a todas las lenguas y criticó con contundencia el que Chomsky y Halle (1968) basaran su clasificación fonológica de los sonidos en dicha posición supuestamente compartida por todas ellas.

de ese idioma o dialecto concreto¹¹. Algo similar parecen pensar Gick et al. (2005) cuando escriben, refiriéndose al inglés:

Además del caso de la schwa, otro posible contexto en el que aparece una ‘vocal neutra’ relacionada con los A[justes] A[rticulatorios básicos] puede ser el de las denominadas ‘pausas llenas’ como *uh* o *um* en inglés. La hipótesis sería que al producirse la fonación durante la producción de un AA subyacente ello daría lugar a un resultado del tipo de estas pausas llenas” (221, traducción nuestra).

Igualmente, van Buuren (1995) había afirmado con anterioridad que “gran parte del habla (o quizá toda) ‘pivota’ alrededor de la postura de una vocal particular (= lengua más labios), y quizá esto resulta más evidente si esa vocal subyacente o ‘básica’ está en la periferia del espacio vocálico más que en algún lugar de la zona central” (133, traducción nuestra).

Todo esto explicaría que la vocal de relleno no tenga por qué coincidir necesariamente con una de las vocales que integran el inventario fonológico de la lengua en cuestión, porque la posición de trabajo no es identificable con un segmento concreto, sino que es la postura preparatoria para producir cualquiera de ellos. Por decirlo de otro modo, los resultados de este estudio apuntan a que algunas de las propiedades de la /e/ del español puedan estar, en efecto, determinadas por las de la posición de trabajo, y acercarse más a ellas que las de otros segmentos, pero ponen de manifiesto también que la /e/ no se corresponde con una simple versión sonora de tal posición. Ni en ruso ni en español el *filler* habitualmente empleado coincide plenamente con ningún fonema integrante de sus respectivos sistemas fonológicos, aunque pueda asociarse perceptivamente con alguno de ellos.

En el caso del castellano, podría contra-argumentarse, al hilo de las observaciones realizadas en Machuca y Ríos (2016), que las diferencias apreciadas entre el timbre más periférico de la vocal de relleno y el de la vocal léxica /e/ no se deben a que se trate de dos elementos realmente distintos, sino al hecho de que la léxica es más breve y está sujeta a la centralización típica del habla espontánea. En otras palabras, la divergencia entre una y otra vendría dada por el mayor o menor tiempo del que disponen los órganos para conformar, respectivamente, la vocal de relleno y la léxica, y por la hipoarticulación de esta última. En habla de laboratorio, pues, los valores de las dos deberían aproximarse o directamente coincidir.

Sin embargo, en el presente estudio, a diferencia de lo que ocurre en el de Machuca y Ríos, las vocales léxicas analizadas eran todas ellas tónicas y fueron extraídas de lecturas realizadas en laboratorio, de modo que no se vieron sujetas a los mismos condicionamientos que las procedentes de habla espontánea. Además, el caso del ruso es revelador: las vocales de

¹¹ Al ser esa la posición favorecida por la lengua, resultan tal vez más fáciles de entender algunos fenómenos particulares. Por ejemplo, tanto en hebreo como en español, la vocal que suele aparecer en casos de epéntesis para deshacer grupos consonánticos no permitidos es la vocal /e/, que se considera fonéticamente no marcada en las dos lenguas. En ambas también, como se dijo en el § 2, la vocal de relleno es bastante similar a la /e/, lo que indica que puede existir coincidencia también en la posición de trabajo subyacente que actúa en los dos idiomas y que podría justificar la aparición ‘espontánea’ de un sonido similar en determinadas posiciones.

relleno no se identifican más con las vocales /a/ y /e/ de referencia dependiendo del tiempo que se invierta en su articulación; muy al contrario, las disimilitudes entre ellas se agudizan a medida que se incrementa la duración de los *fillers*. En concreto, por lo que respecta a la vocal léxica rusa /a/, cuanto mayor es su duración, más desciende su F2, esto es, menos adelantada se pronuncia, $r = -0.21$, y más asciende su F1, es decir, más abierta se articula, $r = 0.28$. Por otro lado, si la vocal léxica rusa /e/ se prolonga, su F2 aumenta, lo que la hace más anterior, $r = 0.16$, y más se incrementa asimismo su F1, de modo que también se abre más, $r = 0.18$. En cambio, cuanto más se alarga la vocal de relleno rusa, más descienden su F2, $r = -0.16$, y su F1, $r = -0.09$, es decir más se atrasa y más se cierra. Aunque las magnitudes de las correlaciones son más bien limitadas, las tendencias muestran que el comportamiento de la vocal de relleno no coincide con el de las vocales léxicas y es distinto del de las categorías fonológicas de referencia. Si la vocal de relleno fuera una /a/, tendría que hacerse más abierta con la duración. Si la vocal de relleno fuera una /e/, tendría que hacerse más abierta y más adelantada. Sin embargo, experimenta justo el cambio contrario¹².

Cabría aún suponer que la variabilidad inter- e intrahablante que se constata en la realización de la vocal de relleno tanto en español castellano como en ruso, y que no siempre es explicable por la duración del sonido, va en contra de la idea de que esta vocal sea el reflejo sonoro de la postura de trabajo característica de cualquiera de las dos lenguas. Sin embargo, Wilson (2006) comprobó no solo la divergencia entre las –en sus términos– *inter-speech postures* del inglés y del francés canadienses, como ya se ha mencionado, sino también que ambas posturas pueden variar en función del contexto, y que no lo hacen de la misma manera en las dos lenguas. Por ejemplo, en el caso del francés, la disposición de los labios característica de la posición de trabajo se vio afectada de manera clara por el contexto fonético inmediatamente anterior. Si esto es así, cabe pensar también que las propiedades del sonido que refleja la posición de trabajo, esto es, de la vocal de relleno, pueden verse alteradas debido a los efectos coarticulatorios ejercidos por el entorno precedente y siguiente. El hecho de que las vocales de relleno analizadas en nuestro estudio vayan situadas entre pausas no es óbice en principio para una explicación de este tipo, ya que durante esas pausas los articuladores adoptan ya la posición de trabajo, susceptible de verse influida por los sonidos vecinos, que solo adquirirá forma sonora al pronunciarse la vocal. En esta ocasión, no hemos controlado el entorno fonético en el que las vocales de relleno aparecían, y el consecuente modo en el que se podía ver modificada su naturaleza, pero esta es, sin duda, una línea de investigación que merece abordarse en próximos trabajos.

¹² De hecho, en el trabajo ya citado con el que los autores iniciamos el análisis de la vocal de relleno en español (Villa, Gil y Lahoz-Bengochea, 2017), concluimos que probablemente la diferencia entre las vocales léxicas y la de relleno podría deberse, no tanto a una hipoarticulación de la primera (puesto que, en nuestro estudio, habíamos seleccionado ya vocales plenas por ser tónicas y en habla leída y de laboratorio), sino a una hiperarticulación de la de relleno, favorecida por su gran duración, lo que podría haber llevado a que esta sobrepasara el supuesto objetivo articulatorio propio de una /e/. El hecho de ampliar el alcance de ese trabajo incorporando todos estos datos del ruso y revisando los aportados para otras lenguas, nos aboca a otra interpretación diferente que, creemos, resulta mucho más convincente y acorde con los resultados obtenidos.

También relacionado con la variabilidad de las vocales de relleno, se plantea un interrogante interesante que suscitan los resultados hallados, a saber, cuál es la trascendencia fonológica que poseen tales elementos, es decir, si se pueden considerar un objetivo articulatorio (en inglés, *articulatory target*) de pleno derecho o no, o lo que es lo mismo, si requieren igual grado de planificación, coordinación temporal y precisión de movimientos que el que pueda exigir la pronunciación de las vocales fonológicas¹³, lo cual tendrá que ver con el margen de variación que les es permitido a unas y a otras. Gick et al. (2005) compararon las mediciones del tracto vocal durante la pronunciación de la vocal /i/ con las obtenidas durante la articulación de la vocal de relleno en inglés y francés, y estimaron la variabilidad que presentaban los dos tipos de realizaciones en cinco hablantes de cada lengua. Posteriormente, se calcularon para cada sujeto las desviaciones estándar en los dos tipos de sonidos y no se observó una diferencia significativa entre los resultados relativos al fonema vocálico y a la vocal de relleno. No se ha realizado un estudio similar sobre el español, pero se debería emprender cuanto antes.

6. CONCLUSIONES

Son muchas las cuestiones a propósito de las vocales de relleno de las lenguas que restan por aclarar. En el caso de la variante castellana del español y de la variante del ruso aquí estudiadas, una vez comprobadas sus características acústicas, lo que resulta imprescindible es contar con análisis articulatorios de las posturas de trabajo con las que parecen corresponderse, similares a los realizados para el inglés o el francés por los autores más arriba citados, con numerosos informantes, normalizando las mediciones de ellos obtenidas, controlando el contenido fonético y el tipo discursivo de los estímulos, y con la necesaria aplicación de los medios técnicos más actualizados. Por lo que se refiere concretamente al español, si, habiéndose alcanzado un conocimiento más profundo de esta cuestión, se corrobora la idea de que la vocal de relleno es efectivamente el reflejo sonoro de la postura por defecto de la lengua, tan determinante para el idioma desde cualquier punto de vista, entonces tal vocal debería ocupar un puesto muy destacado entre los objetos de estudio de investigadores y de docentes.

AGRADECIMIENTOS

Los autores desean hacer constar su agradecimiento a la Dra. Anna Dmítrieva por su ayuda con los textos en ruso, y a Ennoia Torre por su colaboración en el proceso de grabación de los informantes en San Petersburgo.

¹³ El hecho de que las pausas llenas tengan por lo general una frecuencia fundamental más baja que la de los sonidos próximos en la secuencia no deja de ser un indicio de que constituyen un objetivo articulatorio con sus propias peculiaridades especificadas.

OBRAS CITADAS

- Adda-Decker, Martine, Benoît Habert, Claude Barras, Gilles Adda, Philippe Boula de Mareuil y Patrick Paroubek. 2003. "A disfluency study for cleaning spontaneous automatic transcripts and improving speech language models". *Papers in Theoretical Linguistics* 90: 67-70.
- Alonso, Amado. 1974. *Estudios lingüísticos. Temas españoles*. Madrid: Gredos.
- Alves, Helena, Juana Gil, Carolina Pérez y Eugenia San Segundo. 2014. "La cualidad individual de la voz y la identificación del locutor: El proyecto CIVIL". *Fonética Experimental, Espacio Europeo Educación Superior e Investigación*. Eds. Yolanda Congosto, M^a Luisa Montero y Antonio Salvador. Sevilla: Universidad de Cádiz / Universidad de Sevilla. 591-612.
- Annan, Brian. 1972. "The 'articulation base' and Chomsky's neutral position". *Proceedings of the 7th International Congress of Phonetic Sciences*. Eds. André Rigault y René Charbonneau. La Haya: Mouton. 1080-1082.
- Arnold, Jennifer E., Maria Fagnano y Michael K. Tanenhaus. 2003. "Disfluencies signal thee, um, new information". *Journal of Psycholinguistic Research* 32: 25-36.
- Barr, Dale J. y Mandana Seyfeddinipur. 2010. "The role of fillers in listener attributions for speaker disfluency". *Language and Cognitive Processes* 25: 441-455.
- Beattie, Geoffrey W. y Brian L. Butterworth. 1979. "Contextual probability and word frequency as determinants of pauses and errors in spontaneous speech". *Language and Speech* 22: 201-211.
- Belz, Malte y Uwe Reichel. 2015. "Pitch characteristics of filled pauses". *7th Workshop on Disfluency in Spontaneous Speech (DiSS-15)*, Edimburgo: Universidad de Edimburgo.
- Berger, Charles R., Susan H. Karol y Jerry M. Jordan. 1989. "When a lot of knowledge is a dangerous thing: the debilitating effects of plan complexity on verbal fluency". *Human Communication Research* 16: 91-119.
- Boersma, Paul y David Weenink. 2017. "Praat: doing phonetics by computer" [Programa informático, version 6.0.31]. Amsterdam: Universidad de Amsterdam.
- Brander, Dario. 2014. "Phonetic characteristics of hesitation vowels in Swiss German and their use for forensic speaker identification". *IAFPA 2014. 23rd Annual Conference of the International Association for Forensic Phonetics and Acoustics*. Zürich.
- Brennan, Susan E. y Maurice Williams. 1995. "The feeling of another's knowing: Prosody and filled pauses as cues to listeners about the metacognitive states of speakers". *Journal of Memory and Language* 34: 383-398.
- Campione, Estelle y Jean Véronis. 2005. "Pauses and hesitations in French spontaneous speech". *5th Disfluency in Spontaneous Speech Workshop (DiSS-05)*. http://www.isca-speech.org/archive_open/archive_papers/diss_05/dis5_043.pdf. 21 de julio de 2017.
- Candea, Maria, Ioana Vasilescu y Martine Adda-Decker. 2005. "Inter- and intra-language acoustic analysis of autonomous fillers". *5th Disfluency in Spontaneous Speech Workshop (DiSS-05)*. Aix-en-Provence: Universidad de Aix-en-Provence. <https://halshs>.

- archives-ouvertes.fr/halshs-00321914. 21 de julio de 2017.
- Cenoz, Miren Jasone. 2000. "Pauses and hesitation phenomena in second language production". *ITL: Review of Applied Linguistics* 127-8: 53-69.
- Chomsky, Noam y Morris Halle. 1968. *The Sound Pattern of English*. Cambridge, MA.: The MIT Press.
- Christenfeld, Nicholas. 1995. "Does it hurt to say um?" *Journal of Nonverbal Behavior* 19: 171-186.
- Cicres, Jordi. 2014. "Comparación forense de voces mediante el análisis multidimensional de las pausas llenas". *Revista Signos. Estudios de Lingüística* 47: 365-384.
- Clark, Herbert H. y Jean E. Fox Tree. 2002. "Using uh and um in spontaneous speaking". *Cognition* 84: 73-111.
- Corley, Martin y Oliver W. Stewart. 2008. "Hesitation disfluencies in spontaneous speech: The meaning of um". *Language and Linguistic Compass* 2/4: 589-602.
- De Jong, Nivja. 2016. "Predicting pauses in L1 and L2 speech: the effects of utterance boundaries and word frequency". *International Review of Applied Linguistics in Language Teaching* 54: 113-132.
- De Leeuw, Esther. 2007. "Hesitation markers in English, German, and Dutch". *Journal of Germanic Linguistics* 19: 85-114.
- Dewaele, Jean Marc. 1996. "Les phénomènes d'hésitation dans l'interlangue française: Analyse de la variation interstylistique et interindividuelle". *Rassegna Italiana da Linguistica Applicata* 28: 87-103.
- Duez, Danielle. 1982. "Silent and non-silent pauses in three speech styles". *Language and Speech* 25: 11-28.
- . 2001. "Caractéristiques acoustiques et phonétiques des pauses remplies dans la conversation en français". *Travaux Interdisciplinaires du Laboratoire Parole et Langage* 20: 31-48.
- Fox Tree, Jean E. 2002. "Interpreting pauses and ums at turn exchanges". *Discourse Processes* 24: 37-55.
- García Amaya, Lorenzo. 2015. "A longitudinal study of filled pauses and silent pauses in second language speech". *7th Workshop on Disfluency in Spontaneous Speech (DiSS-15)*. Edimburgo: Universidad de Edimburgo. http://www.disfluency.org/DiSS_2015/Programme_files/Garcia-Amaya-DISS2015.pdf. 29 de agosto de 2017.
- Giannini, Antonella. 2003. "Hesitation phenomena in spontaneous Italian". *Proceedings of the 15th International Congress of Phonetic Sciences*. Eds. María Josep Solé, Daniel Recasens y Joaquín Romero. Adelaida: Casual Productions. 2653-2656.
- Gick, Bryan, Ian Wilson, Karsten Koch y Clare Cook. 2005. "Language-specific articulatory settings: Evidence from interutterance rest position". *Phonetica* 61: 1-14.
- Goldman-Eisler, Frieda. 1968. *Psycholinguistics: Experiments in Spontaneous Speech*. Londres / Nueva York: Academic Press.
- Goto, Masataka, Katunobu Itou y Satoru Hayamizu. 1999. "A real-time filled pause detection system for spontaneous speech recognition". *Eurospeech'99*: 227-230. <https://>

- staff.aist.go.jp/m.goto/PROJ/fp.html. 15 de junio de 2017.
- Hartmann, Reinhard Rudolf Karl y Francis C. Stork. 1972. *Dictionary of Language and Linguistics*. Lincoln, G.B.: Applied Science Publishers.
- Horne, Merle. 2009. "Creaky fillers and speaker attitude: data from Swedish". *Where Prosody Meets Pragmatics: Research at the Interface*. Eds. Dagmar Barth-Weingarten, Nicole Déhé y Anne Wichmann. Lund: Emerald Group Publishing Limited. 277-288.
- Horváth, Viktória. 2010. "Filled pauses in Hungarian: Their phonetic form and function". *Acta Linguistica Hungarica* 57: 288-306.
- Kedrova, Galina E. y Constantine Leo Borisoff. 2013. "The concept of 'basis of articulation' in Russia in the first half of the 20th century". *Historiographia Linguistica* 40: 151-197.
- Künzel, Hermann J. 1997. "Some general phonetic and forensic aspects of speaking tempo". *Forensic Linguistics* 4: 48-83.
- Levelt, Willem J. M. 1989. *Speaking: From Intention to Articulation*. Cambridge, MA.: The MIT Press.
- Machuca, María Jesús, Joaquim Llisterri y Antonio Ríos. 2015. "Las pausas sonoras y los alargamientos en español: un estudio preliminar". *Normas. Revista de Estudios Lingüísticos Hispánicos* 5: 81-96.
- Machuca, María Jesús y Antonio Ríos. 2016. "Estructura formántica de las pausas sonoras en español". *53 reflexiones sobre aspectos de la fonética y otros temas de lingüística*. Ed. Ana M^a. Fernández Planas. Barcelona: Laboratorio de Fonética de la Universidad de Barcelona. 67-76.
- Machuca, María Jesús. En prensa. "Las pausas". *Fonética y fonología descriptivas de la lengua española*. Eds. Juana Gil y Joaquim Llisterri. Washington/Madrid: Georgetown University Press / Consejo Superior de Investigaciones Científicas (CSIC).
- Maclay, Howard y Charles Osgood. 1959. "Hesitation phenomena in spontaneous speech". *Word* 15: 19-44.
- Maekawa, Kikuo y Hiroki Mori. 2015. "Voice quality analysis of Japanese filled pauses: a preliminary report". *7th Workshop on Disfluency in Spontaneous Speech (DiSS-15)*. Edimburgo: Universidad de Edimburgo. <http://www2.ninjal.ac.jp/kikuo/AllPublicationKM.html>. 7 de agosto de 2017.
- Merlo, Sandra y Leticia Mansur. 2004. "Descriptive discourse: topic familiarity and disfluencies". *Journal of Communication Disorders* 37: 489-503.
- Moreno Sandoval, Antonio, Leonardo Campillos y Doroteo Toledano. 2012. "A quantitative study of disfluencies in formal, informal and media spontaneous speech in Spanish". *IberSpeech 2012. Jornadas en Tecnologías del Habla y III Iberian SL Tech Workshop*. http://www.lllf.uam.es/leonardo/pdf/IberSPEECH2012_Dysfluencies.pdf. 27 de julio de 2017.
- Navarretta, Costanza. 2015. "The functions of fillers, filled pauses and co-occurring gestures in Danish dyadic conversations". *Proceedings from the 3rd European Symposium on Multimodal Communication*. Dublín: Universidad de Dublín. 55-61.

- Nemoto, Rena. 2006. *Hésitation vocalique autonome vs. système phonétique de la langue: étude acoustique en plusieurs langues*. Mémoire de DESS (Tesis de Máster). París: LIMSI-CNRS.
- - -. 2013. "Which kind of hesitations can be found in Estonian spontaneous speech?" *13th Workshop on Disfluency in Spontaneous Speech (DiSS-13)*: 53-54.
- Nicholson, Hannele, Kathleen Eberhard y Matthias Scheutz. 2010. "Um... I don't see any': The function of filled pauses and repairs". *5th Workshop on Disfluency in Spontaneous Speech – 2nd International Symposium on Linguistic Patterns in Spontaneous Speech*, Tokio: Universidad de Tokio. 89-92.
- O'Connell, Daniel C. y Sabine Kowal. 1983. "Pausology". *Computers in Language Research*. Eds. Walter A. Sedelow y Sally Yeates Sedelow. Berlín: De Gruyter. 221-301.
- - -. 2005. "Uh and um revisited: Are they interjections for signaling delay?" *Journal of Psycholinguistic Research* 34: 555-576.
- O'Shaughnessy, Douglas. 1992. "Recognition of hesitations in spontaneous speech". *ICASSP'92. Proceedings of the 1992 IEEE International Conference on Acoustics, Speech and Signal Processing*. San Francisco: IEEE Computer Society. 521-524.
- Öhman, S. E. G. 1967. "Peripheral motor commands in labial articulation". *Quarterly Progress Status Report of the Speech Transmission Laboratory, Royal Institute of Technology –KTH– of Stockholm (QPSR)* 4: 30-63.
- Pätzold, Matthias y Adrian P. Simpson. 1995. "An acoustic analysis of hesitation particles in German". *Proceedings of the 13th International Congress of Phonetic Sciences*. Estocolmo. 512-515.
- Proença, Jorge, Dirce Celorico, Arlindo Veiga, Sara Candeias y Fernando Perdigão. 2013. "Acoustical characterization of vocalic fillers in European Portuguese". *6th Workshop on Disfluency in Spontaneous Speech (DiSS-13)*. 63-66.
- Quimbo, Félix C. M., Tatsuya Kawahara y Shuji Doshita. 1998. "Prosodic analysis of fillers and self-repair in Japanese speech". *Proceedings of the International Conference on Spoken Language Processing (ICSLP)*: 3313-3316.
- Rose, Ralph Leon. 1998. *The Communicative Value of Filled Pauses in Spontaneous Speech* (Tesis de Máster). Birmingham: Universidad de Birmingham. www.roselab.sci.waseda.ac.jp/resources/file/madissertation.pdf. 20 de Agosto de 2017.
- - -. 2013. "Crosslinguistic Corpus of Hesitation Phenomena: A corpus for investigating first and second language speech performance". *Interspeech 2013. Proceedings of the 14th Annual Conference of the International Speech Communication Association*: 992-996.
- Rebollo, Leticia. 1997. "Pausas y ritmo en la lengua oral. Didáctica de la pronunciación". *El español como lengua extranjera: del pasado al futuro*. Eds. Francisco Moreno, María Gil y Kira Alonso. Alcalá de Henares: Universidad de Alcalá. 667-676.
- Riazantseva, Anastasia. 2001. "Second language proficiency and pausing: a study of Russian speakers of English". *Studies in Second Language Acquisition* 23: 497-526.
- Schachter, Stanley, Nicholas Christenfeld, Bernard Ravina y Frances Bilous. 1991. "Speech

- disfluency and the structure of knowledge”. *Journal of Personality and Social Psychology* 60: 362-367.
- Schaeffler, Sonja, James Scobbie e Ineke Mennen. 2008. “An evaluation of inter-speech-postures for the study of language-specific articulatory settings”. *Proceedings of the International Seminar on Speech Production (ISSP)*. Eds. Susanne Fuchs, Rudolph Sock e Yves Laprie. 121-124. <http://issp2008.loria.fr/Proceedings/PDF/issp2008-24.pdf>. 20 de junio de 2017.
- Scherer, Wilhelm. 1868. *Zur Geschichte der deutschen Sprache*. Berlín: Franz Duncker.
- Séguin, Marie Claude. 2008. “Analyse formantique des pauses remplies chez les adolescents unilingues et outaouais québécois: une étude pilote”. http://homes.chass.utoronto.ca/~cla-acl/actes2008/CLA2008_Seguin.pdf. 15 de agosto de 2017.
- Shriberg, Elisabeth. 1994. *Preliminaries to a Theory of Speech Disfluencies*. (Tesis de Doctorado). Berkeley: Universidad de California, Berkeley.
- - -. 1999. “Phonetic consequences of speech disfluency”. *Proceedings of the 14th International Congress of Phonetic Sciences*. Eds. John Ohala, Yoko Hasegawa, Manjari Ohala, Daniel Granville y Ashlee C. Bailey. San Francisco: Universidad de California. 619-622.
- - -. 2001. “To ‘errrr’ is human: ecology and acoustics of speech disfluencies”. *Journal of the International Phonetic Association* 31: 153-169.
- Silber-Varod, Vered y Noam Amir. 2016. “Formant analysis of the mid-front vowel as realized in hesitation disfluencies in Hebrew”. *Phonetician* 113: 49-60.
- - -. Adva Weiss y Noam Amir. 2015. “Can you hear these mid-front vowels? Formants analysis of hesitation disfluencies in spontaneous Hebrew”. *7th Workshop on Disfluency in Spontaneous Speech (DiSS-15)*. Póster. Edimburgo: Universidad de Edimburgo. http://www.disfluency.org/DiSS_2015/Programme_files/Silber-Varod-DISS2015.pdf. 23 de julio de 2017.
- Spreafico, Lorenzo. 2016. “Filled pauses in multilingual speech: an acoustic analysis”. *Linguistica e Filologia* 36: 99-116.
- Stepanova, Svetlana. 2007. “Some features of filled hesitation pauses in spontaneous Russian”. *Proceedings of the 16th International Congress of Phonetic Sciences*. Eds. Jürgen Trouvain y William J. Barry. Saarbrücken: Universidad de Saarbrücken. 1325-1328.
- Trofimovich, Pavel y Wendy Baker. 2006. “Learning second language suprasegmentals: Effect of L2 experience on prosody and fluency characteristics of L2 speech”. *Studies in Second Language Acquisition* 28: 1-30.
- Tsiaras, Vassilis, Costas Panagiotakis y Yannis Stylianou. 2009. “Video and audio based detection of filled hesitation pauses in classroom lectures”. *Proceedings of the 17th European Signal Processing Conference (EUSIPCO 2009)*. Glasgow: IEEE. 834-838.
- Urizar, Xabier y Arthur Samuel. 2013. “A corpus-based study of fillers among native Basque speakers and the role of *zera*”. *Language and Speech* 0: 1-29.
- Van Buuren, Laurens. 1995. “Postura: clear and dark consonants, etcetera”. *Studies in Gen-*

- eral and English Phonetics: Essays in Honour of Professor J. D. O'Connor*. Ed. John Windsor-Lewis. Londres: Routledge. 130-142.
- Vasilescu, Ioana, Maria Candea y Martine Adda-Decker. 2004. « Hésitations autonomes dans 8 langues: une étude acoustique et perceptive ». *Colloque Modélisation pour l'identification des langues et des variétés dialectales (MIDL)*. <https://perso.limsi.fr/madda/publications/PDF/midl04filledp.pdf>. 25 de julio de 2017.
- Vasilescu, Iona, Martine Adda-Decker y Rena Nemoto. 2008. "Acoustic and prosodic characteristics of vocalic hesitations in three languages". https://rs2007.limsi.fr/TLP_Page_14.html. 29 de agosto de 2017.
- Vasilescu, Ioana y Martine Adda-Decker. 2007. "A cross-language study of acoustic and prosodic characteristics of vocalic hesitations". *Fundamentals of Verbal and Non-verbal Communication and the Biometric Issue*. Eds. Anna Esposito, Maja Bratanić, Eric Keller y María Marinaro. Ámsterdam: IOS Press. 140-148.
- - -. Maria Candea y Martine Adda-Decker. 2005. "Perceptual salience of language-specific acoustic differences in autonomous fillers across eight languages". *Interspeech 2005*. <https://halshs.archives-ouvertes.fr/hal-00875151/document>. 22 de julio de 2017.
- - -. Rena Nemoto y Martine Adda-Decker. 2007. Vocalic hesitations vs. vocalic systems: A cross-language comparison. <http://www.icphs2007.de/conference/Papers/1504/index.html>. 24 de julio de 2017.
- - -. Sophie Rosset y Martine Adda-Decker. 2010. "On the functions of the vocalic hesitation euh in interactive man-machine question answering dialogs in French". <https://www.semanticscholar.org/paper/On-the-functions-of-the-vocalic-hesitation-euh-in-Vasilescu-Rosset/9f75bf25cc7b1da13c48ec781fc9f62247f17f0a>. 25 de julio de 2017.
- Villa, José, Juana Gil y José M^a Lahoz-Bengoechea. 2017. "Las vocales de relleno en español: nuevos datos y algunas reflexiones". *Nuevos estudios sobre comunicación social*. Eds. Leonel Ruiz Mirayes, María R. Álvarez Silva y Alex Muñoz Alvarado. La Habana: Centro de Lingüística Aplicada. 165-169.
- Watanabe, Michiko. 2003. "The constituent complexity and types of fillers in Japanese". *Proceedings of the 15th International Congress of Phonetic Sciences*. Eds. Maria Josep Solé, Daniel Recasens y Joaquín Romero. Adelaida: Casual Productions. 2473-2476.
- Wieling, Martijn, Jack Grieve, Gosse Bouma, Josef Fruehwald, John Coleman y Mark Liberman. 2016. "Variation and change in the use of hesitation markers in Germanic Languages". *Language Dynamics and Change* 6: 199-234.
- Wilson, Ian Lewis. 2006. *Articulatory Settings of French and English Monolingual and Bilingual Speakers*. (Tesis de Doctorado). Universidad de la Columbia Británica.
- Voss, Bernd. 1979. "Hesitation phenomena as sources of perceptual errors for non-native speakers". *Language and Speech* 22: 129-144.
- Wu, Chen Huei. 2008. "Filled pauses in L2 Chinese: A comparison of native and non-native speakers". *Proceedings of the 20th North American Conference on Chinese Linguistics*. Eds. Marjorie K. M. Chan y Hana Kang. Columbus: Universidad de Ohio.

213-227.

Yuan, Jiahong, Xu Xiaoying, Wei Lai y Mark Liberman. 2016. "Pauses and pause fillers in Mandarin monologue speech: The effects of sex and proficiency". *Speech Prosody*: 1167-1170.