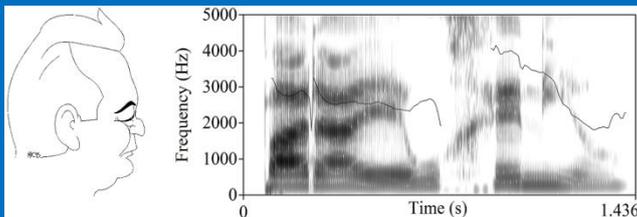


Estructura formántica de las pausas sonoras en español

María J. Machuca Ayuso

Antonio Ríos Mestre



Fernández Planas, A. Ma. (ed.) (2016): *53 reflexiones sobre aspectos de la fonética y otros temas de lingüística*, Barcelona, págs. 67-76.

ISBN: 978-84-608-9830-6.

Estructura formántica de las pausas sonoras en español

María J. Machuca Ayuso
Universitat Autònoma de Barcelona
MariaJesus.Machuca@uab.cat

Antonio Ríos Mestre
Universitat Autònoma de Barcelona
Antonio.Rios@uab.cat

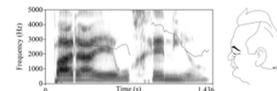
Para Eugenio, un fonetista infatigable.

1. INTRODUCCIÓN

Entendemos por pausas sonoras aquellos fenómenos de hesitación que no van asociados a ningún elemento léxico; son elementos sonoros que produce el hablante sin ningún significado concreto y que aparecen generalmente en un estilo de habla espontáneo (Machuca, Llisterri y Ríos, 2015). Se afirma que su aparición en el discurso oral se considera como una de las estrategias lingüísticas más empleadas (véase, por ejemplo, Nicholson, Eberhard y Scheutz, 2010; Wennerstrom, 1994), ya que no solo sirven para señalar cuál es el momento adecuado para efectuar un cambio de turno de palabra en la conversación (Cestero, 2000), sino también para darle tiempo al hablante a planificar y a reorganizar el discurso oral (Stepanova, 2007).

Algunos autores han señalado que la cualidad del sonido que se emplea para emitir una pausa sonora depende del sistema fónico de cada lengua (Levelt, 1989; O'Shaughnessy, 1993; Stepanova, 2007; Giannini, 2003) o recurren a una vocal no marcada (Gil, 2007); en todo caso, los hablantes realizan sonidos que no requieren amplios movimientos articulatorios (Pätzold y Simpson, 1995). En español, según Rebollo (1997), los fonemas /e/ y /m/ son los únicos que pueden constituir verdaderas pausas sonoras, ya que pueden insertarse en el discurso sin que necesariamente deriven del alargamiento de una palabra. En Machuca, Llisterri y Ríos (2015) se observa que, de 156 pausas sonoras extraídas de un corpus de habla espontánea, un 61,5% de los casos corresponde a un sonido que se puede transcribir como [e:]; un 32%, a un murmullo nasal y el 7% restante a vocalizaciones que se podrían transcribir como [a:] y a vocalizaciones que no se pueden categorizar en ninguno de los casos anteriores (p. 86).

El objetivo de este estudio es analizar, en español, la estructura formántica de las pausas sonoras que se pueden transcribir como un sonido [e:] y compararla con la de la vocal /e/ producida en interior de palabras, tanto en sílabas tónicas como átonas, para determinar si ambos elementos poseen características acústicas diferentes, como sucede en otras lenguas, por ejemplo, en alemán (Pätzold y Simpson, 1995) y en inglés (Ward, 2006). En la misma línea, los resultados obtenidos en Stepanova (2007) para el ruso indican que los valores de frecuencia del primer formante de la pausa sonora se sitúan entre los valores de



una /a/ y los de una /e/, pero los del segundo formante presentan los mismos valores que la vocal /a/. Por otro lado, los resultados de Giannini (2003) para el italiano muestran que la estructura formántica de la pausa sonora varía en función de las variedades analizadas: en el italiano meridional la pausa sonora muestra unos valores semejantes a los de una vocal central, aun teniendo en cuenta que es una vocal que no existe en el sistema fonológico de esta variante, mientras que en el norte las características acústicas de este elemento son las propias de una vocal anterior. En catalán, Cicres (2014) proporciona valores de los 3 primeros formantes de las pausas sonoras cuyo timbre se encuentra en un espacio comprendido entre los sonidos [e:] y [ə:]; el valor medio del primer formante, considerando los 10 hablantes analizados, es de 666 Hz y el del segundo, de 1611 Hz, lo que indica que es una vocal central, denominada también *schwa*.

2. PROCEDIMIENTO EXPERIMENTAL

El corpus de este trabajo está constituido por las grabaciones de habla espontánea de cinco informantes masculinos. Dos de ellas forman parte de las entrevistas grabadas para el proyecto *El español en contacto con el catalán: variación diatópica y bilingüismo* (Ministerio de Economía y Competitividad: FFI2012-33499). Cada una de las grabaciones dura unos 30 minutos, de los que se han analizado los 15, aproximadamente, que corresponden a las intervenciones del hablante entrevistado. El resto de las grabaciones forman parte del corpus de habla espontánea de AHUMADA (Ortega, González y Marrero, 2000), que fue analizado en el proyecto VILE, *Estudio acústico de la variación inter e intralocutor en español* (BFF2001-2551, 2001-2004 y HUM2005-06980, 2005-2008). Cada uno de los informantes fue grabado en tres sesiones separadas en el tiempo, por lo que en este trabajo se ha analizado un total de 9 minutos de grabación, 3 minutos por hablante.

En este corpus se han seleccionado todas las pausas sonoras realizadas como una vocal semejante a la [e] y un total, como máximo, de 30 casos de vocales [e] tónica y 30 de vocales [e] átona situadas en una sílaba CV, seguidas y precedidas de consonantes dentales y alveolares, ya que estos son los puntos de articulación que menos influyen en los valores formánticos de las vocales (Iskarous et al, 2013). Por otro lado, también se ha considerado que en el entorno consonántico no aparecieran nasales, pues estas pueden crear un formante nasal en la estructura formántica de las vocales analizadas (Chen, 1997). Mediante el programa PRAAT (Boersma y Weenink, 2011) se han medido la duración de las pausas sonoras y de las vocales, así como los valores de frecuencia de los dos primeros formantes. Para obtener los datos de frecuencia se ha tomado el valor medio de las vocales analizadas, debido a que poseen una duración muy breve en habla espontánea. Por otra parte, también se tomó el valor medio de los dos primeros formantes de las pausas sonoras, pues la frecuencia se mantenía estable a lo largo de todo el segmento.

En total se han analizado 68 casos de pausas sonoras, 94 de vocales tónicas y 96 de vocales átonas. Se ha de tener en cuenta que el número de pausas sonoras depende de cada informante; por ejemplo, el informante 5 presenta solo 8 en 3 minutos de grabación, mientras que el informante 1 muestra 19 en 15 minutos. En cuanto al número de casos de



las vocales, no solo depende del informante, sino de la longitud total de las grabaciones analizadas. Así, en los informantes cuyas grabaciones duraban 15 minutos, se han llegado a analizar 30 casos de cada tipo de vocal; en cambio, en los informantes de los que solo se dispone de 3 minutos para cada uno de ellos, el número de vocales tónicas oscila entre 9 y 16, y el de vocales átonas entre 9 y 17, por tanto, debemos considerar los resultados con cautela, ya que el número de casos reducido hace que las conclusiones no sean definitivas.

3. RESULTADOS

La figura 1 muestra los valores de frecuencia de los dos primeros formantes (F1 y F2) de los tres segmentos analizados ([e], [e] y pausas sonoras), comparando los datos en el mismo informante. La finalidad de adoptar este procedimiento era observar si la relación que se podía establecer entre las características acústicas de los elementos analizados era la misma en todos los hablantes. Como se puede observar en la figura 1, los valores de F1 son similares para todos ellos y en los tres tipos de segmentos. Los valores de significación de un test ANOVA, en el que la variable independiente son los segmentos y la dependiente los valores de F1 de cada uno de los tres tipos de segmentos, normalizados por los valores medios correspondientes, así lo indican ($F(2, 256) = 0.1, p > .05$). No obstante, aunque no existan diferencias significativas entre los valores de las vocales átonas y tónicas, se mantiene la tendencia general en español de que los valores de frecuencias son más altos en las tónicas que en las átonas (Albalá et al, 2008), excepto para los informantes 1 y 4. En cambio, los valores de frecuencia del F2 no muestran diferencias significativas entre la vocal tónica y la vocal átona en ninguno de los cinco informantes analizados ($F(4, 186) = 0.5, p > .05$), pero si se comparan estos valores con los de las pausas sonoras, sí se obtienen diferencias ($F(2, 256) = 81.1, p < .05$). Como se puede observar en la figura 1, los valores de frecuencia de F2 de las pausas sonoras se aproximan a la línea de los 1800 Hz o la sobrepasan, mientras que no ocurre lo mismo en los valores de frecuencia del segundo formante de las vocales.

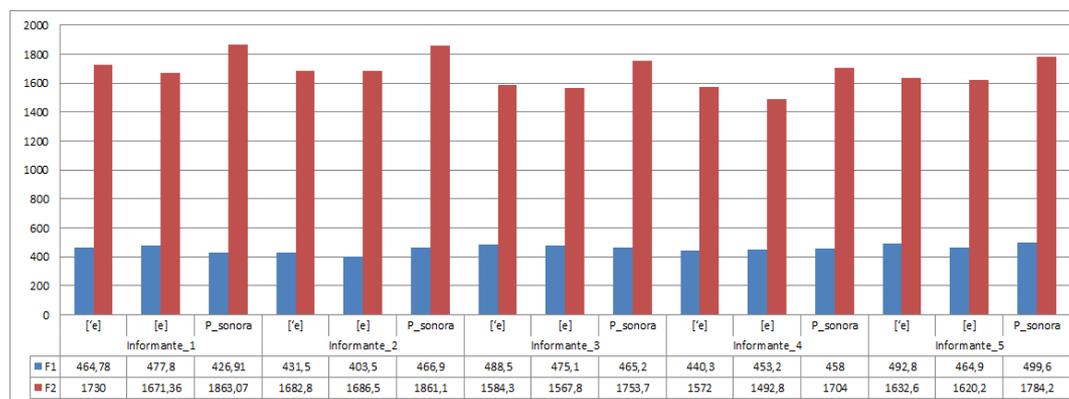
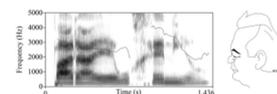


Figura 1. Valores de frecuencia en Hz del primer (F1) y del segundo formante (F2) de los tres segmentos analizados: [e], [e] y pausa sonora.



En la figura 2, donde se muestran las áreas de dispersión de los valores de frecuencia de F1 y F2 de los segmentos analizados, puede observarse cómo los valores que se agrupan en la parte superior del gráfico corresponden a las pausas sonoras; esto corrobora lo que se había comentado en el párrafo anterior: las pausas sonoras presentan valores de frecuencia del segundo formante más altos que los valores de las vocales.

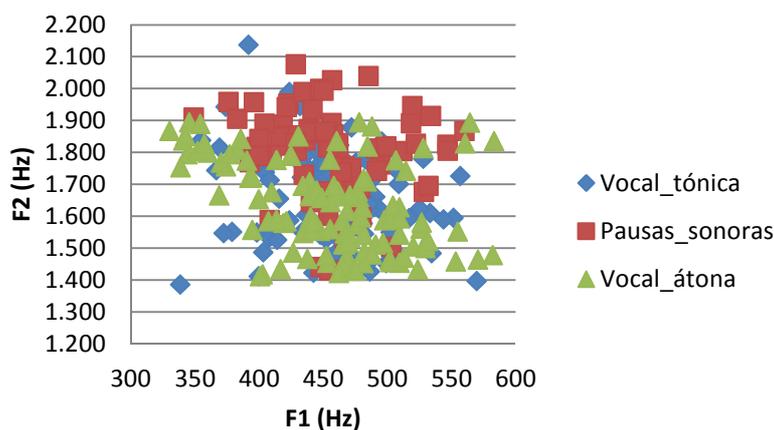
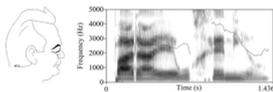


Figura 2. Áreas de dispersión de los valores de F1 y F2 de los tres segmentos analizados: [e], [e] y pausa sonora.

En cuanto a los valores de duración, excepto para el informante 4, al igual que ocurría con los valores de frecuencia del primer formante, los datos muestran unos valores algo más elevados en las vocales tónicas que en las átonas, pero las diferencias no son significativas, como lo muestran los valores de significación de un test ANOVA en la que la variable dependiente son las duraciones en cada uno de los informantes y la dependiente, la tonicidad ($F(4, 186) = 1.3, p > .05$). Obviamente, el parámetro de duración resulta significativo cuando consideramos el valor temporal de la pausa sonora ($F(8, 250) = 8.6, p < .05$). En la tabla 1 pueden observarse los valores medios y la desviación estándar de cada tipo de segmento analizado.

	[e]		[e]		Pausa sonora	
	Media	Desv. Est.	Media	Desv. Est.	Media	Desv. Est.
Informante_1	60,38	14,2	56,89	10,2	729,5	285,9
Informante_2	55,3	14,2	43,9	8,3	367	196,5
Informante_3	63	16,7	57	18	565,3	278,1
Informante_4	63,3	16,3	67	15,5	547	226,7
Informante_5	76	17,8	64	12,9	349	166,8

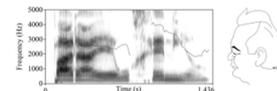
Tabla 1. Valores de duración en ms de los tres segmentos analizados: [e], [e] y pausa sonora.



En cuanto a la duración de la pausa sonora, puede observarse que tanto los valores medios como los de desviación estándar son muy altos y dispares entre los informantes, ya que es el hablante el que selecciona el tiempo que emplea en una pausa sonora como una estrategia discursiva individual.

4. DISCUSIÓN

Aunque el número de casos de pausas sonoras analizadas en este estudio es reducido, pues su aparición en el discurso depende, como ya se ha mencionado, del hablante y de la duración de las grabaciones analizadas, se puede observar a partir de los valores de los parámetros analizados un comportamiento similar: los valores de frecuencia del primer formante de la pausa sonora tienden a los de una vocal palatal de grado medio de abertura en español, mientras que los del segundo formante presentan un valor de frecuencia más alto que el de las vocales [e] y [e] extraídas de una situación comunicativa correspondiente a un estilo de habla espontánea. Los datos de este experimento muestran que las vocales en este tipo de habla sufren una centralización, es decir, los valores de los formantes se acercan a una vocal realizada en el espacio que se encuentra entre las vocales de grado medio de abertura y la vocal abierta. Este fenómeno ya se ha puesto de relieve en numerosos trabajos en los que se comparan las características acústicas de las vocales en diferentes estilos de habla (Lindblom et al, 1996, por ejemplo). En español, las áreas de dispersión trazadas con los valores de frecuencia de los dos primeros formantes están perfectamente definidas y separadas en las realizaciones de habla de laboratorio, en cada una de las cinco vocales, mientras que con los datos obtenidos en habla espontánea las áreas de dispersión son más extensas y los valores formánticos se superponen (Harmegnies y Poch, 1992; Poch y Harmegnies, 2010). Por otro lado, los valores del segundo formante de la pausa sonora correspondiente a un sonido que podría transcribirse como [e:] se mantienen estables durante su duración, por esta razón, la pausa sonora es uno de los elementos que se analiza en la identificación de un hablante en trabajos relacionados con la fonética judicial (Künzel, 1997; Rose, 2002, entre otros). A partir de nuestros datos, podría afirmarse que los valores de ese segundo formante se acercan más a lo que es una vocal [e] no centralizada, propia del habla de laboratorio. En Machuca y Poch (2015) el valor medio de frecuencia del segundo formante de la vocal [e] en español, mezclando tanto vocales átonas como tónicas, en un corpus de lectura de palabras aisladas producidas por hablantes bilingües de catalán y de español con un porcentaje de uso mayor del español, es de 1758 Hz. Los datos de frecuencia del segundo formante obtenidos en este trabajo oscilan entre 1704 Hz, el valor medio más bajo, y 1863 Hz, el valor medio más alto. Si comparamos estos datos con los obtenidos en Cicres (2014), 666 Hz para el primer formante y 1611 Hz para el segundo, se puede observar una clara diferencia respecto a los nuestros. Los valores que muestra en su estudio corresponden a lo que realmente es una vocal central, ya que los hablantes analizados son todos catalanohablantes. No debemos olvidar que el catalán presenta dos subsistemas fonológicos en la variedad central: el subsistema tónico que comprende los segmentos /i e ε a o u / y el subsistema átono que está compuesto de tres unidades /i u a/ y de una vocal centralizada /ə/, cuyo estatus fonológico es puesto en tela de juicio por muchos autores (Julià, 2002). Por tanto, los hablantes bilingües en español y en catalán utilizan dos



lenguas que presentan notables diferencias fonológicas, pues se distinguen en que el español presenta tres grados de abertura y el catalán posee cuatro, además del fenómeno de reducción que existe en catalán. Cuando el hablante es de dominancia catalana, como se observa en el estudio de Cicres, la vocal resultante para la realización de una pausa sonora muestra la estructura correspondiente a una vocal central, mientras que cuando los hablantes son de dominancia española, la vocal resultante es una vocal [e].

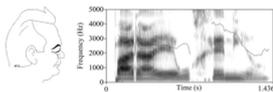
Como ya se ha mencionado en la introducción, existen autores que indican que las pausas sonoras no pueden identificarse con un sonido de la propia lengua y que todas las lenguas tienden a una vocal central (Pätzold y Simpson, 1995), mientras que otros relacionan la estructura formántica de la pausa sonora con algún sonido de la lengua en cuestión (Giannini, 2003, para el italiano; Stepanova, 2007, para el ruso). Los valores obtenidos en este trabajo indican que la estructura formántica del sonido con el que se realiza una pausa sonora puede relacionarse con la de una vocal palatal de grado medio de abertura, pero correspondiente a los valores que se muestran en un estilo de habla hiperarticulado, como puede ser el habla de laboratorio. En cuanto a la duración, los valores de las vocales muestran una duración similar a pesar de que unas son tónicas y otras, átonas, mientras que la duración de las pausas sonoras muestra una gran desviación en todos los informantes analizados.

5. CONCLUSIONES

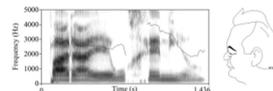
Los resultados de este trabajo muestran que la estructura formántica de las pausas sonoras no es comparable con la de las vocales extraídas de un corpus de habla espontánea, ya que estas últimas sufren un fenómeno de centralización. Por tanto, puede darse una relación entre la duración y la frecuencia de los formantes. Las vocales en habla espontánea, independientemente de que se realicen como tónicas o como átonas, poseen una duración menor que las del habla de laboratorio y sufren una centralización. Por el contrario, la mayor duración de las pausas sonoras permite al hablante realizar un elemento vocálico con una estructura formántica estable, por esta razón, los valores de frecuencia obtenidos poseen un timbre similar al de una vocal palatal de grado medio de abertura en habla de laboratorio. Estos resultados ponen de manifiesto que, aunque existen trabajos en los que se ha analizado este fenómeno que defienden que la vocalización de una pausa sonora se realiza mediante una vocal central (O'Shaughnessy, 1993), en español, según muestran los resultados de este experimento, el hablante tiende a una vocal palatal de grado medio de abertura. Debería considerarse si se debe a que no existen vocales neutras en el sistema fonológico de esta lengua o a que la frecuencia de aparición de esta vocal [e] en la lengua es muy alta. Finalmente, sería interesante considerar en próximos estudios cómo la vocalización de esa pausa sonora repercute en el grado de acento extranjero cuando los hablantes aprenden otra lengua.

6. REFERENCIAS BIBLIOGRÁFICAS

- ALBALÁ, M. J.; E. BATTANER; M. CARRANZA; J. GIL; J. LLISTERRI; M. J. MACHUCA; N. MADRIGAL; M. MARQUINA; V. MARRERO; C. DE LA MOTA; M. RIERA y A. RÍOS (2008): «VILE: nuevos datos acústicos sobre vocales del español» en A. Pamies, M. C. Amorós y J. M. Pazos (eds):



- Language Design. Journal of Theoretical and Experimental Linguistics. Special Issue 1: New Trends in Experimental Phonetics*, Granada, Método Ediciones, pp.1-14.
- BOERSMA, P. y D. WEENINK (2011): *Praat: Doing phonetics by computer* (version 5.3.02). <http://www.praat.org/> [15/12/2015].
- CESTERO, A. M. (2000): *El intercambio de turnos de habla en la conversación. Análisis socio-lingüístico*, Alcalá de Henares, Universidad de Alcalá de Henares, Servicio de Publicaciones.
- CICRES, J. (2014): «Comparación forense de voces mediante el análisis multidimensional de las pausas llenas», *Revista Signos. Estudios de Lingüística*, 47(86), pp. 365-384. <http://www.scielo.cl/pdf/signos/v47n86/a02.pdf> [15/12/2015].
- CHEN, M Y. (1997): «Acoustic correlates of English and French nasalized vowels», *The Journal of the Acoustical Society of America*, 102(4), pp. 2360-2370.
- GIANNINI, A. (2003): «Hesitation Phenomena in Spontaneous Italian», en M. J. Solé, D. Recasens y J. Romero (eds): *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona, pp. 2653-2656.
- HARMEGNIES, B. y D. POCH-OLIVÉ (1992): «A study of style-induced vowel variability. Laboratory versus spontaneous speech in Spanish», *Speech Communication*, 11, pp. 429-437. [http://dx.doi.org/10.1016/s0167-6393\(99\)00081-3](http://dx.doi.org/10.1016/s0167-6393(99)00081-3) [15/12/2015].
- ISKAROUS, K.; C. MOOSHAMMER; P. HOOLE; D. RECASENS; C. H. SHADLE; E. SALTZMAN y WHALEN, D. H. (2013): «The coarticulation/invariance scale: Mutual information as a measure of coarticulation resistance, motor synergy, and articulatory invariance», *The Journal of the Acoustical Society of America*, 134(2), pp. 1271-1282. <http://doi.org/10.1121/1.4812855> [15/12/2015].
- JULIÀ, J. (2002): «Els sons del català», en Solà, J.; M. R Lloret; J. Mascaró y M. Pérez Saldanya (dirs): *Gramàtica del català contemporani*, vol. I, Barcelona, Empúries, pp. 37-87.
- KÜNZEL, H. J. (1997). «Some general phonetic and forensic aspects of speaking tempo», *The International Journal of Speech, Language and the Law*, 4(1), pp. 48-83.
- LEVELT, W. J. M. (1989): *Speaking: from intention to articulation*, Cambridge, MA: MIT Press.
- LINDBLOM, B.; S. BROWNLEE y R. LINDGREN (1996): «Formant undershoot and speaking styles: An attempt to resolve some controversial issues» en A. Simpson y M. Pätzold (eds): *Sound patterns of connected speech, models and explanation*, Kiel, Institut für Phonetik und Digitalsprachverarbeitung, Universität Kiel, pp. 119-130.
- MACHUCA, M. J. y D. POCH-OLIVÉ (en prensa): «Dinámica de las vocales del español en contacto con el catalán», *Oraia. Análisis del discurso oral*, 19.
- MACHUCA, M. J.; J. LLISTERRI y A. RÍOS (2015): «Las pausas sonoras y los alargamientos en español: un estudio preliminar», *Normas. Revista de Estudios Lingüísticos Hispánicos*, 5, pp. 81-96. <https://ojs.uv.es/index.php/normas/article/view/6823> [15/12/2015].
- NICHOLSON, H.; K. EBERHARD y M. SCHEUTZ (2010): «“Um... I don't see any”: The function of filled pauses and repairs», *Proceedings of DiSS-LPSS Joint Workshop 2010. 5th Workshop on Disfluency in Spontaneous Speech and 2nd International Symposium on Linguistic Patterns in Spontaneous Speech*, Tokio, pp. 89-92.
- ORTEGA, J.; J. GONZÁLEZ y V. MARRERO (2000): «AHUMADA: A large speech corpus in Spanish for speaker characterization and identification», *Speech Communication*, 31(2-3), pp. 255-264.
- O'SHAUGHNESSY, D. (1993): «Locating disfluencies in spontaneous speech: an acoustical analysis», *Proceedings of European Conference on Acoustics of Speech Communication and Technology*, 3, pp. 2187-2190.
- PÄTZOLD, M. y A. SIMPSON (1995): «An acoustic analysis of hesitation particles in German», en K. Elenius y P. Branderud (eds.): *Proceedings of the 13th International Congress of Phonetic Sciences*, Estocolmo, pp. 512-515.
- POCH-OLIVÉ, D. y B. HARMEGNIES (2010): «Centralización y reducción en las lenguas románicas», en M. Iliescu, H. Siller-Rungaldier y P. Danler (eds): *Actes du XXVe Congrès Internationale de Linguistique et de Philologie Romanes*, Innsbruck, De Gruyter, pp. 137-146.



-
- REBOLLO, L. (1997): «Pausas y ritmo en la lengua oral. Didáctica de la pronunciación», en F. Moreno; M. Gil y K. Alonso (eds.): *El español como lengua extranjera: del pasado al futuro. Actas del VIII Congreso Internacional de la Asociación para la Enseñanza del Español como Lengua Extranjera*, Alcalá de Henares, Servicio de Publicaciones de la Universidad de Alcalá, pp. 667-676.
- ROSE, P. (2002): *Forensic speaker identification*, Londres, Taylor & Francis.
- STEPANOVA, S. (2007): «Some features of filled hesitation pauses in spontaneous Russian» en J. Trouvain y W. J. Barry (eds.): *Proceedings of ICPHS 2007. 16th International Congress of Phonetic Sciences*, Saarbrücken, pp. 1325-1328.
- WARD, N. (2006): «Non-Lexical conversational sounds in American English», *Pragmatics and Cognition*, 14(1), pp. 113-184.
- WENNERSTROM, A. (1994): «Intonational meaning in English discourse: A study of non-native speakers», *Applied Linguistics*, 15(4), pp. 399-420.
<http://dx.doi.org/10.1093/applin/15.4.399> [15/12/2015].