CARACTERIZACIÓN DEL HABLANTE CON FINES JUDICIALES: FENÓMENOS FÓNICOS PROPIOS DEL HABLA ESPONTÁNEA

Joaquim Llisterri
Universitat Autònoma de Barcelona
María J. Machuca
Universitat Autònoma de Barcelona
Antonio Ríos
Universitat Autònoma de Barcelona

RESUMEN

El análisis de 4232 hesitaciones (pausas y alargamientos) extraídas de un corpus de habla espontánea grabado por 21 hablantes de español pone de manifiesto que la frecuencia de aparición de los diferentes tipos de hesitación, la duración global de los silencios y la duración de los alargamientos vocálicos presentan una elevada variación interlocutor y una escasa variabilidad intralocutor. Además, los alargamientos vocálicos y las pausas silenciosas son los dos tipos de hesitación más frecuentes en el corpus. Estos resultados, aunque preliminares, permiten concluir que el estudio de las hesitaciones puede ofrecer datos relevantes para la comparación entre muestras de habla en el contexto de la fonética judicial.

Palabras clave: hesitaciones, pausas, alargamientos, habla espontánea, fonética judicial, fonética acústica, español

ABSTRACT

An analysis of 4232 hesitations (pauses and prolongations) extracted from a corpus of spontaneous speech recorded by 21 speakers of Spanish shows that the frequency of occurrence of the different types of hesitations, the global duration of silent pauses and the duration of vowel lengthening exhibit high inter-speaker variability and low intra-speaker variation. Moreover, prolongations in vowels and silent pauses are the most frequent types of hesitations found in the corpus. The results, although preliminary, support the conclusion that the study of disfluencies may offer relevant insights for the comparison of speech samples in forensic work.

Keywords: hesitations, disfluencies, pauses, prolongations, spontaneous speech, forensic phonetics, acoustic phonetics, Spanish

1. INTRODUCCIÓN Y OBJETIVOS

El análisis del habla espontánea ha puesto de manifiesto la existencia de una serie de características específicas que se explican, fundamentalmente, por la necesidad de elaborar el discurso en tiempo real; ello da lugar a un conjunto de fenómenos que se han descrito como *hesitaciones* o *disfluencias* (Shriberg, 2001). El presente trabajo[1] se centra en los tipos de hesitación que atañen al nivel fonético: los alargamientos segmentales, tanto vocálicos como consonánticos, la conjunción *y* con valor expletivo, las pausas silenciosas —también conocidas como *vacías*— y las pausas sonoras —denominadas, en ocasiones, *llenas*—.

Aunque el estudio de las hesitaciones en el plano fónico se ha llevado a cabo desde diversos ámbitos, la perspectiva que se adopta en esta investigación es la de la fonética judicial, con el fin de obtener datos sobre las alteraciones en la fluidez del discurso que resulten pertinentes para las tareas que se llevan a cabo cuando se comparan muestras de habla a fin de determinar si corresponden a la misma persona o si proceden de locutores diferentes. Se pretende, pues, establecer las características de los alargamientos y de las pausas que presentan una mayor variabilidad entre hablantes diferentes y, a la vez, un menor grado de variación en un mismo locutor, una propiedad considerada esencial en los parámetros utilizados para comparar muestras de habla en contextos judiciales (Foulkes y French, 2012; Jessen, 2008).

2. METODOLOGÍA

El material analizado procede de las grabaciones de habla espontánea del corpus CIVIL (San Segundo, Alves y Fernández Trinidad, 2013). En ellas, una entrevistadora y los participantes conversaban durante tres o cuatro minutos sobre temas de trabajo, aficiones, etc., en dos sesiones separadas en el tiempo. La muestra seleccionada está formada por 21 hablantes femeninas.

Los archivos sonoros se segmentaron y etiquetaron mediante el programa Praat (Boersma y Weenink, 2018), siguiendo los criterios definidos en el proyecto VILE (Llisterri, Machuca y Ríos, 2017).

[1]El estudio se ha realizado en el marco del proyecto ESTIVOZ, *Grado de espontaneidad y factores temporales en el disimulo de la voz* (FFI2014-59848-C2-2-P) financiado por el Ministerio de Economía y Competitividad del Gobierno de España.

En las pausas silenciosas (figura 1) se incluyen los silencios propiamente dichos que se encuentran en el interior de los turnos de palabra de los participantes. Si después del silencio aparece una oclusiva sorda, se considera que la pausa incluye la oclusión cuando la ausencia de energía sonora supera los 90 ms.

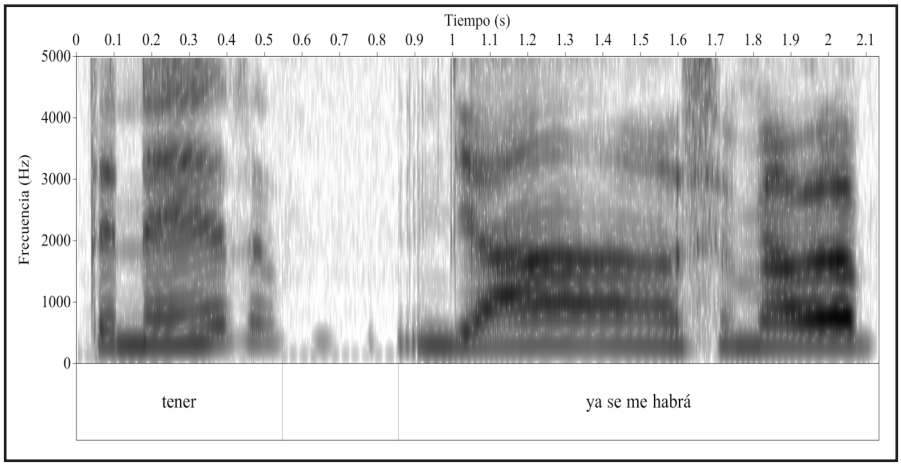


Figura 1. Pausa silenciosa en el enunciado: "claro, el acento que pudiera tener ya se me habrá perdido"

Se han considerado como pausas sonoras (figura 2) las vocalizaciones y los murmullos sonoros que no forman parte de una palabra.

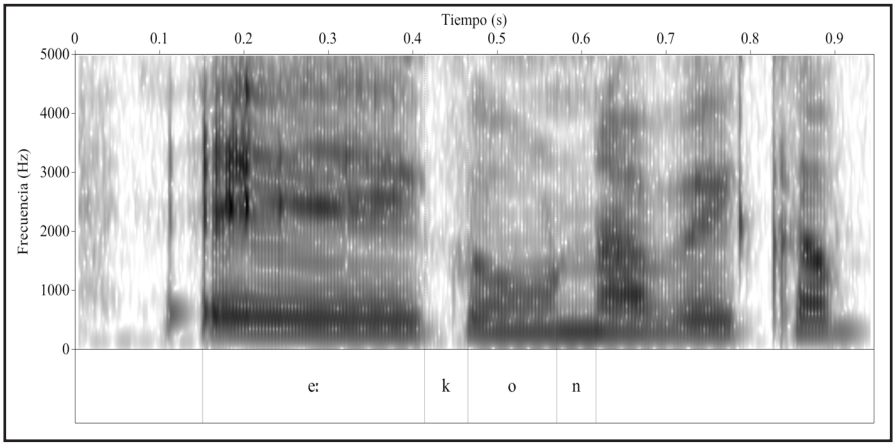


Figura 2. Pausa sonora realizada como [e:] en el enunciado: "eh, con una beca"

En el caso de los alargamientos vocálicos (figura 3) y consonánticos (figura 4), se han analizado separadamente los que se realizan en palabras funcionales y los que aparecen en palabras con contenido y, para ambos tipos de hesitaciones, se ha diferenciado entre los situados en posición prepausal y los que se encuentran en otros contextos.

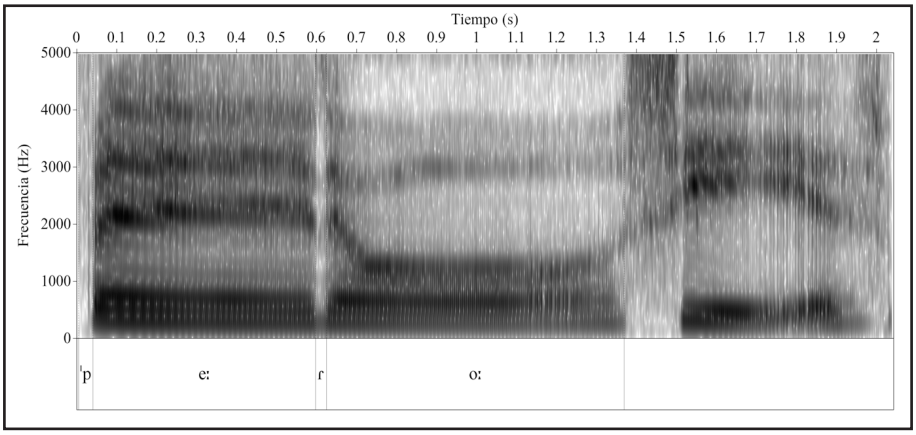


Figura 3. Alargamientos vocálicos en el enunciado: "pero, sí, es muy bonito de estudiar"

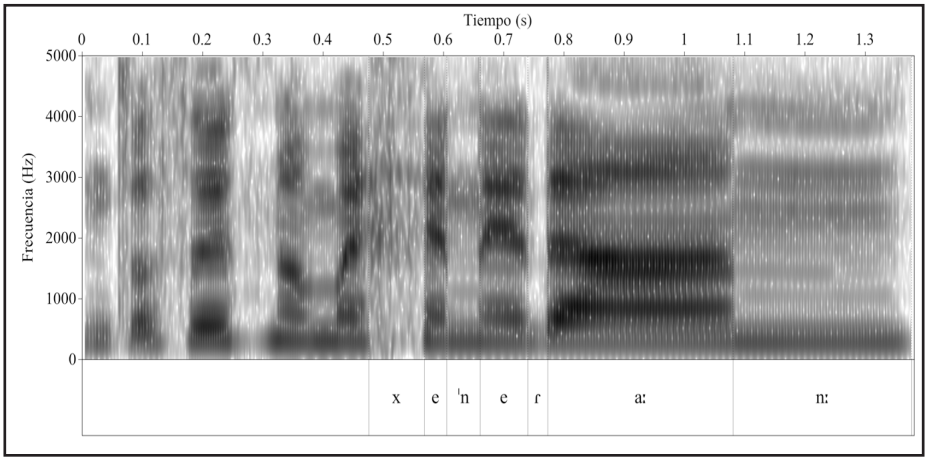


Figura 4. Alargamiento consonántico en el enunciado: "entonces se me generan"

Los casos en los que la conjunción *y* pierde su función de categoría léxica que une sintagmas y oraciones se han clasificado como expletivos (Cortés, 1991); aun consistiendo en el alargamiento de una vocal, se asemejan a una pausa sonora, como puede comprobarse en los ejemplos de las figuras 2 y 5.

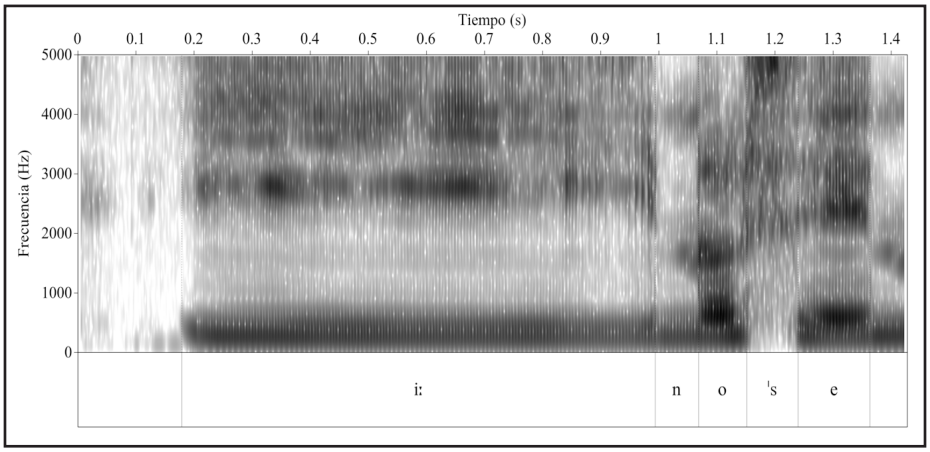


Figura 5. Alargamiento del elemento expletivo en el enunciado: "y, no sé, no he pedido más cosas"

En total, se han analizado 4 232 hesitaciones, distribuidas como se muestra en la tabla 1.

Fenómeno	Sesión	Número de casos
Pausas silenciosas	1	829
	2	880
	Total	*1 709*
Pausas sonoras	1	103
	2	80
	Total	*183*
Alargamientos consonánticos	1	111
	2	148
	Total	*259*
Alargamientos vocálicos	1	988
	2	1 012
	Total	*2 000*
Elementos expletivos	1	44
	2	37
	Total	*81*

Tabla 1. Número de casos para cada uno de los tipos de hesitaciones analizadas en las dos sesiones de grabación

Los datos se han analizado estadísticamente considerando, por un lado, la frecuencia de aparición de las hesitaciones en cada hablante y, por otro, la duración de cada una de ellas. En el primer caso, se han empleado modelos multinomiales, ya que la variable de respuesta (es decir, la dependiente) es la frecuencia de aparición de cada tipo de hesitación; en el segundo, se han utilizado modelos mixtos en los que la duración es la variable de respuesta.

3. RESULTADOS

Se presentan, en primer lugar, los resultados más generales, correspondientes a la frecuencia de aparición de las hesitaciones y, a continuación, los datos relativos a los valores de duración de los alargamientos y de las pausas.

3.1 Frecuencia de aparición de las hesitaciones

En la figura 6 se recoge la frecuencia de aparición de los distintos tipos de hesitaciones. Como puede comprobarse, en consonancia con los datos de la tabla 1, los alargamientos vocálicos y las pausas silenciosas son los más frecuentes; a continuación se sitúan los alargamientos consonánticos y las pausas sonoras, siendo el elemento expletivo el menos frecuente.

Globalmente, se constata que en las pausas silenciosas y en los alargamientos consonánticos aumenta el número de apariciones en la segunda sesión de entrevistas, al contrario de lo que ocurre con los alargamientos vocálicos, las pausas sonoras y con el elemento expletivo; sin embargo, estas diferencias no resultan estadísticamente significativas ($\chi^2(4, N = 4232) = 10{,}55$, $p > 0{,}05$).

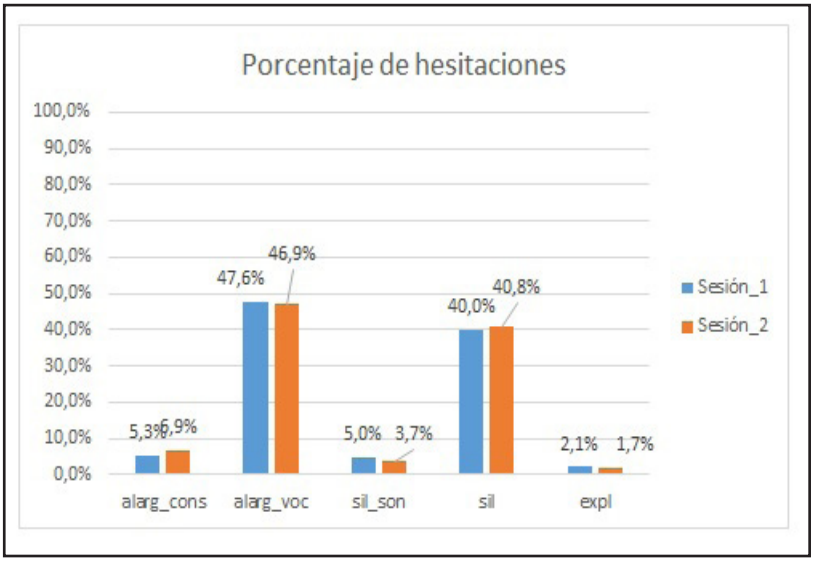
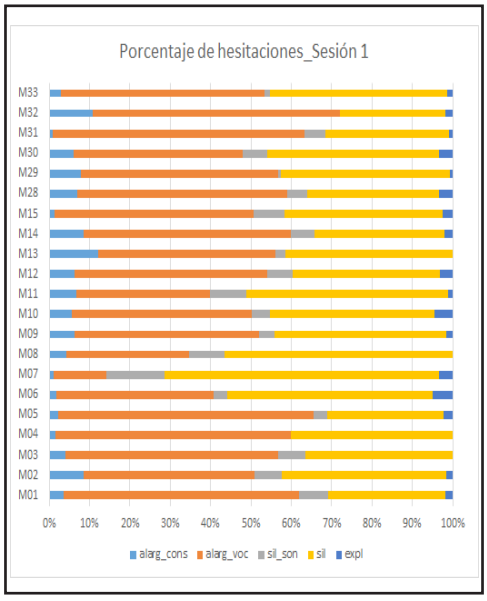
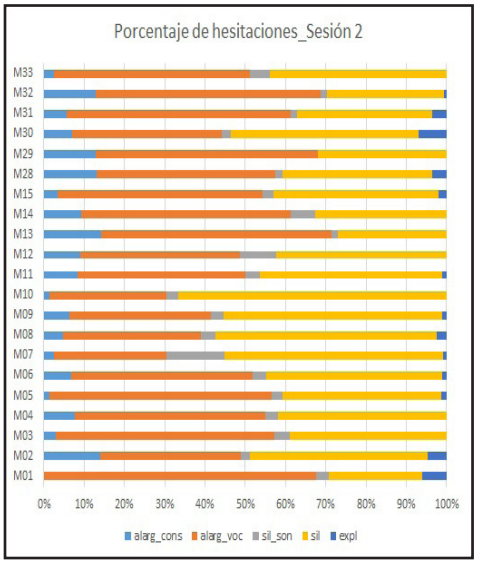


Figura 6. Frecuencia de aparición cada tipo de hesitación en función de la sesión de grabación. Tipos de hesitación: alarg_cons = alargamientos consonánticos; alarg_voc = alargamientos vocálicos; sil_son = pausas sonoras; sil = pausas silenciosas; expl = elementos expletivos

La frecuencia de aparición de las distintas clases de hesitaciones muestra diferencias significativas entre los hablantes ($\chi^2(80, N = 4232) = 354{,}7$, $p < 0{,}05$), mientras que el porcentaje con el que aparecen los diferentes tipos de hesitación deja de resultar significativo si se comparan las dos sesiones de grabación de un mismo locutor ($\chi^2(80, N = 4232) = 101{,}5$, $p = 0{,}05$). Al contrastar las distribuciones que se visualizan en la figura 7 se aprecia que, aunque los participantes utilizan todas las hesitaciones, los alargamientos vocálicos y las pausas silenciosas son las que se emplean con mayor frecuencia. De hecho, un análisis discriminante llevado a cabo solamente con estas dos clases de hesitación permite clasificar correctamente el 56 % de los 21 locutores analizados.



Sesión 1



Sesión 2

Figura 7. Frecuencia de aparición de las hesitaciones empleadas por cada hablante en cada sesión de grabación. Tipos de hesitación: alarg_cons = alargamientos consonánticos; alarg_voc = alargamientos vocálicos; sil_son = pausas sonoras; sil = pausas silenciosas; expl = elementos expletivos

3.2 Duración de las hesitaciones

Los resultados del análisis de la duración de cada clase de hesitación estudiada se resumen en la figura 8. Como puede apreciarse, los valores dependen del tipo de hesitación (F (4, 4205) = 187,5, p < 0,05). Los alargamientos presentan un valor medio de 286,5 ms; los expletivos, de 511,38 ms; los silencios, de 521,3 ms y las pausas sonoras, de 375,2 ms Si se comparan los valores temporales del conjunto de hesitaciones en cada una de las dos sesiones analizadas para cada hablante, se observa que apenas se encuentran diferencias (F (1, 4205) = 0,1, p > 0,05). La interacción entre sesión y hesitación, tomando al hablante como variable aleatoria, tampoco proporciona resultados significativos (F (4, 4149) = 0,9, p > 0,05).

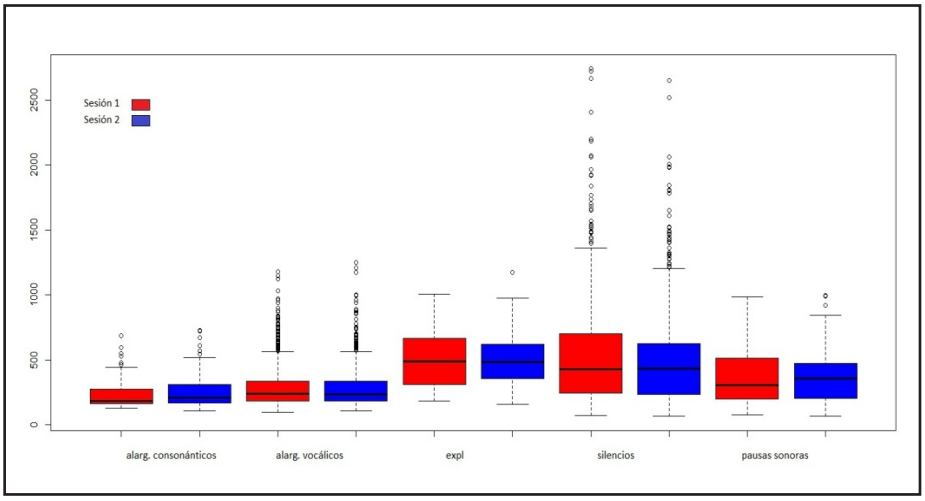


Figura 8. Diagramas de cajas de los valores de duración (en ms) de las hesitaciones en función del tipo de hesitación y de la sesión de grabación. Tipos de hesitación: alarg = alargamientos; expl = expletivos; silencios = pausas silenciosas

En lo que respecta a los alargamientos (figura 9), los resultados ponen de manifiesto que en los consonánticos no se dan diferencias de duración entre los que aparecen en posición prepausal y los que no se encuentran antes de pausa, tanto en palabras funcionales (F (1, 107) = 0,3, p = 0,5) como en palabras con contenido (F (1, 151) = 0,022, p = 0,8); en cambio, en los vocálicos pueden encontrarse diferencias temporales —en palabras con contenido: F (1, 1324) = 15,89, p = 0,000; en palabras funcionales: F (1, 674) = 26,59, p = 0,000— debidas a la clase de palabra, como se recoge en la tabla 2.

Clase de palabra	Posición	Duración media	Desviación típica
Palabras con contenido	Prepausal	284,74	147,7
	No prepausal	225,29	113,3
Palabras funcionales	Prepausal	396,3	204
	No prepausal	311,39	166

Tabla 2. Duración (en ms) de los distintos tipos de alargamiento vocálicos analizados

Cabe añadir, que la duración de los alargamientos, tanto vocálicos como consonánticos, es mayor en las palabras funcionales que en las palabras con contenido.

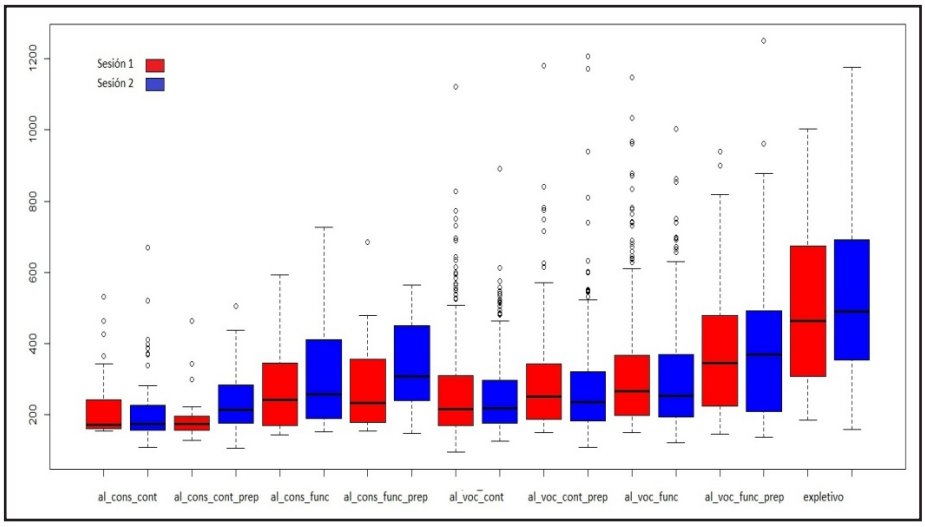


Figura 9. Diagramas de cajas de los valores de duración (en ms) de los alargamientos en función del tipo de alargamiento y de la sesión de grabación. Tipos de hesitación: al = alargamiento; cons = consonante; cont = palabra con contenido; prep = posición prepausal; func = palabra funcional

En cuanto a la relevancia de los alargamientos para diferenciar los locutores, se observa que la duración de los alargamientos consonánticos no contribuye a la discriminación entre hablantes (F (20, 239) = 1,3 p > 0,05), mientras que la duración de los alargamientos vocálicos (F (20, 1979) = 9,4, p < 0,05) y la de los expletivos (F (17, 81) = 1,04, p = 0,4) constituyen dos parámetros que resultan de utilidad para esta tarea.

En el caso de las pausas silenciosas, los valores temporales no muestran diferencias significativas entre las que preceden a una oclusiva sorda y las que no aparecen en esta posición (F (1, 1708) = 0,3, p > 0,05). Las pausas silenciosas ante oclusiva sorda presentan una duración media de 527,8 ms (DT = 404 ms) y, las situadas en el resto de los contextos, de 514,77 ms (DT = 374 ms). Los valores de duración de los silencios, considerados globalmente, contribuyen a la diferenciación entre hablantes (F (20, 1708) = 4,9, p < 0,05).

Por último, las pausas sonoras se perciben con tres realizaciones fonéticas diferentes, con un número distinto de casos, como se puede observar en la tabla

3. Sin embargo, las diferencias temporales entre las tres realizaciones no son significativas (F (2, 182) = 0,1, p = 0,8) y la duración de la pausa sonora no resulta de utilidad para discriminar entre hablantes (F (20, 182) = 1,1, p > 0,05).

Realización	Número de casos	Duración media	Desviación típica
[mː]	92	373	209,38
[eː]	86	391	233,8
[aː]	5	361,5	62

Tabla 3. Número de casos y duración (en ms) de las realizaciones fonéticas de las pausas sonoras

4. DISCUSIÓN Y CONCLUSIONES

Al analizar las 4 232 hesitaciones detectadas en dos sesiones de grabación de entrevistas realizadas a 21 hispanohablantes se ha observado, en primer lugar, que la frecuencia con la que aparece cada tipo de hesitación difiere entre locutores, pero no lo hace entre las dos sesiones de un mismo participante, por lo que, como también encontraron Ishihara y Kinoshita (2010), Braun y Rosin (2015) o McDougall y Duckworth (2017) para otras lenguas, se trata de un fenómeno que ofrece una de las características que son relevantes en el ámbito judicial.

En segundo lugar, se ha constatado una elevada frecuencia de aparición de alargamientos vocálicos y de pausas silenciosas, así como una escasa presencia de elementos expletivos; esta jerarquía coincide, en buena parte, con la que se obtuvo en el análisis de dos corpus de habla espontánea en español (Machuca, Llisterri y Ríos, 2015).

Finalmente, en la duración total de las hesitaciones se aprecia una elevada variación entre hablantes, mientras que no se detectan diferencias al comparar las dos sesiones de grabación de un mismo participante, al igual que sucede con la duración de los alargamientos vocálicos. También el tiempo global de silencio en la producción de cada hablante permite discriminar entre locutores, como se pone igualmente de manifiesto en los datos recogidos en Ríos y Machuca (2017).

En conjunto, pues, puede pensarse que las hesitaciones propias del habla espontánea contienen datos valiosos para la fonética judicial. Aun así, el caso de las pausas llenas —cuya duración no resulta efectiva para distinguir entre

locutores en ausencia de información sobre los valores de los formantes— hace patente que, como también señalan Braun y Rosin (2015) o Ishihara y Kinoshita (2010), es la combinación de diversos elementos lo que contribuye, en última instancia, a establecer los rasgos que permiten a un experto determinar la semejanza entre muestras de habla.

REFERENCIAS BIBLIOGRÁFICAS

Boersma, P. y Weenink, D. 2018. *Praat: Doing Phonetics by Computer* [Programa informático]. URL: http://www.praat.org.

Braun, A. y Rosin, A. 2015. "On the speaker-specificity of hesitation markers". En The Scottish Consortium for ICPhS 2015 (ed.), *Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow: The University of Glasgow. URL: https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0731.pdf.

Cortés, L. 1991. *Sobre conectores, expletivos y muletillas en el español hablado*. Málaga: Ágora.

Foulkes, P. y French, P. 2012. "Forensic speaker comparison: A linguistic–acoustic perspective". En L. Solan y P. Tiersma (eds.), *The Oxford Handbook of Language and Law*. Oxford: Oxford University Press, 557-572.

Ishihara, S. y Kinoshita, Y. 2010. "Filler words as a speaker classification feature". En M. Tabain, J. Fletcher, D. Grayden, J. Hajek y A. Butcher (eds.), *Proceedings of the 13th Australasian International Conference on Speech Science and Technology*. Melbourne, Australia, 34-37.

Jessen, M. 2008. "Forensic phonetics". *Language and Linguistics Compass*, 2: 671-711.

Llisterri, J., Machuca, M. J. y Ríos, A. 2017. *VILE-P: un corpus para el estudio prosódico de la variación inter e intralocutor*. Comunicación presentada en Subsidia: herramientas y recursos para las ciencias del habla. Málaga, junio 2917, España.

Machuca, M. J., Llisterri, J. y Ríos, A. 2015. "Las pausas sonoras y los alargamientos en español: un estudio preliminar". *Normas. Revista de estudios lingüísticos hispánicos*, 5: 81-96.

McDougall, K. y Duckworth, M. 2017. "Profiling fluency: An analysis of individual variation in disfluencies in adult males". *Speech Communication*, 95: 16-27.

Ríos, A. y Machuca, M. J. 2017. "Parámetros acústicos en la variación inter e intralocutor en español (VILE-I)". En V. Marrero (ed.), *Introducción a la fonética judicial*. València: Tirant Lo Blanch, 111-129.

San Segundo, E., Alves, H. y Fernández Trinidad, M. 2013. "CIVIL Corpus: Voice quality for speaker forensic comparison". *Procedia – Social and Behavioral Sciences*, 95: 587-593.

Shriberg, E. 2001. "To 'errrr' is human: Ecology and acoustics of speech disfluencies". *Journal of the International Phonetic Association*, 31: 153-169.