

El ritmo del español y el del inglés.

Repercusiones metodológicas sobre su medición acústica*

Speech rhythm in Spanish and English.

Methodological implications about its acoustic measurement

Gonzalo Eduardo Espinosa

Universidad Nacional del Comahue/IPEHCS-CONICET

gonzalo.espinosa@fadel.uncoma.edu.ar

Recibido: 14 de abril de 2018

Aceptado: 23 de julio de 2018

RESUMEN

Desde un abordaje fonológico, el ritmo de habla se considera como la suma de diversas propiedades fonético-fonológicas, como la estructura silábica, el rol del acento y la reducción vocálica. El español y el inglés representan dos tipologías rítmicas marcadamente diferenciadas. Para dar cuenta de las diferencias rítmicas de las lenguas se han instrumentado diversas mediciones acústicas. En este trabajo adoptamos dos de las principales métricas rítmicas de duración: %V y VarcoV, empleadas en dos tratamientos distintos según el cómputo de las pausas en el habla (métodos A y B del programa Correlatore) y según los elementos finales antes de las pausas (inclusión y exclusión del intervalo final). Por medio de un corpus obtenido de forma experimental, comparamos las distintas mediciones para dar cuenta de los cómputos más productivos a la hora de reflejar las diferencias del ritmo del español de la Patagonia argentina y el del inglés estándar de Inglaterra y Estados Unidos. Nuestros resultados indican que la manera más óptima de diferenciar estos dos ritmos corresponde al empleo de la métrica VarcoV por medio del método B y con exclusión del intervalo final.

Palabras clave: ritmo de habla, métricas, español, inglés

ABSTRACT

From a phonological approach, speech rhythm has been conceived of as the product of different phonetic and phonological properties, such as syllabic structure, the role of stress and vowel reduction. Spanish and English represent two clearly different rhythmic typologies. Many acoustic measurements have been developed in order to account for the rhythmic differences among languages. In this study two rhythm metrics are adopted: %V and VarcoV. These metrics are computed considering two treatments for pauses (method A and B from the software Correlatore) and the treatment for intervals before pauses (with inclusion or exclusion of final intervals). By means of a corpus obtained experimentally, different measurements are compared so as to get the most productive way of distinguishing the rhythm of Spanish in Patagonia, Argentina, and the rhythm from standard British English and standard American English. The results indicate that the most effective technique for reflecting rhythmical differences is by means of VarcoV, method B and exclusion of final intervals.

Keywords: speech rhythm, metrics, Spanish, English

* Quiero agradecer a los informantes que participaron de las distintas tareas de elicitación y a Faith Youth y Gabriela Portal, por su colaboración en la recolección del corpus en Inglaterra y Estados Unidos, respectivamente. Agradezco también a los evaluadores anónimos por sus comentarios y sugerencias.

1. Introducción

Según un enfoque tradicional (Pike 1945, Abercrombie 1967), los dos principales tipos de ritmos en las lenguas comprenden un ritmo acentual (como el inglés, el alemán y el holandés) y un ritmo silábico (como el español, el francés y el italiano). En este tipo de clasificación, el acento y la sílaba son los elementos estructuradores del ritmo acentual y silábico, respectivamente. Estos elementos son considerados como eventos que ocurren de manera periódica y, por lo tanto, se habla de la isocronía de alguno de ellos en el ritmo de las lenguas. Sin embargo, los trabajos que han intentado demostrar esta isocronía en lenguas como el español y el inglés no obtuvieron resultados esperados. Para el inglés, los aportes de Lehiste (1973) demuestran de manera experimental que hay una ausencia de isocronía en los acentos, por lo que este autor concluye que la isocronía se vincula con la percepción de los hablantes y no con la manifestación acústica del ritmo. Esta idea de percepción rítmica se ve reforzada por Allen (1975), quien sugiere que los intervalos regulares de tiempo son efecto de la percepción del ritmo. Para el español, los trabajos de Borzone de Manrique & Signorini (1983) y de Pamies Bertrán (1999) dan cuenta de la falta de duración constante en la sílaba, en contra de cómo se considera un ritmo silábico desde una mirada tradicional.

Por su parte, Roach (1982) realiza estudios experimentales en las lenguas que han sido consideradas con ritmos prototípicamente distintos. Es decir, este autor analiza si una lengua con ritmo silábico como el español presenta una duración silábica más constante que los intervalos entre acentos y, asimismo, si una lengua con ritmo acentual como el inglés presenta duraciones silábicas más variadas y duraciones de intervalos interacentuales más uniformes. Sus resultados indican que tanto las duraciones de las sílabas como las de los intervalos interacentuales no son uniformes en las lenguas y que no es posible hacer una distinción categórica de los ritmos. En suma, según Roach, no existen valores absolutos en las características rítmicas de las lenguas y éstas presentan una superposición de tipologías rítmicas. Del mismo modo, este autor sugiere estudiar características fonético-fonológicas como la estructura silábica y la reducción vocálica con el fin dar cuenta de las diferencias rítmicas en las lenguas.

Si bien adoptamos una terminología tradicional, en este trabajo tenemos en cuenta el ritmo de habla como el producto de diversas propiedades fonético-fonológicas (Bertinetto 1981, 1989, Dasher & Bolinger 1982, Dauer 1983, 1987). Las principales propiedades que se tienen en cuenta son el acento léxico principal, la estructura silábica y la reducción vocálica. Un ritmo acentual como el del inglés presenta mayores efectos del acento, una estructura silábica variada y compleja y, por último, procesos de reducción vocálica. Por otra parte, un ritmo silábico como el del español se caracteriza por tener menores efectos del acento, una estructura silábica menos variada y simple y no existen casos de reducción vocálica.

Con el objetivo de dar cuenta del ritmo del habla por medio de un análisis acústico, se han diseñado las métricas rítmicas (cfr. Ramus *et al.* 1999, White & Mattys 2007, Dellwo & Wagner 2003, Grabe & Low 2002). Estas métricas computan los

parámetros acústicos del habla como la duración, la intensidad y la sonoridad, reflejando las características fonológicas que dan forma al ritmo. Para el empleo de las métricas se debe, en primer lugar, segmentar los fragmentos de habla en intervalos vocálicos (v) y consonánticos (c). Cada intervalo corresponde a un segmento o grupo de segmentos. Por ejemplo, en una frase como *el habla* [el.'a.βla], la segmentación es la siguiente |v|c|v|cc|v|, con tres intervalos vocálicos y dos intervalos consonánticos.

Basándose en la duración, Ramus *et al.* (1999) introducen tres métricas: %V tiene en cuenta la proporción de los intervalos vocálicos, ΔV mide la desviación estándar de las duraciones de los intervalos vocálicos y, por último, ΔC mide la desviación estándar de las duraciones de los intervalos consonánticos. Estos autores sostienen que los índices métricos logran discriminar las lenguas según la tipología rítmica a la que pertenecen, especialmente cuando los valores se computan por medio de dos métricas en conjunto. En concreto, las lenguas como el español y el francés se asocian a un ritmo silábico, reflejado en valores bajos para ΔV y ΔC y altos para %V; mientras que las lenguas como el inglés y el alemán corresponden a un ritmo acentual, que se manifiesta con valores altos para ΔV y ΔC y bajos para %V.

Debido a la influencia de la velocidad del habla en los cálculos métricos, las métricas se han modificado a modo de atenuar esta influencia, Dellwo (2006) propone la métrica VarcoC, con la que se divide la desviación estándar de los intervalos consonánticos por la duración promedio de todos los intervalos. White & Mattys (2007) implementan la métrica VarcoV, que, al igual que el aporte de Dellwo (2006), logra atenuar los efectos de la velocidad del habla a través del cómputo con los intervalos vocálicos. Según White & Mattys (2007), el cálculo con vocales, como lo hace VarcoV, permite reflejar mejor las distinciones rítmicas de las lenguas. Estos autores analizan el holandés, el inglés, el español y el francés y sugieren utilizar la métrica VarcoV en conjunto con la métrica %V, porque este cómputo en conjunto permite una mayor diferenciación entre un ritmo acentual (holandés e inglés) y un ritmo silábico (español y francés).

Para facilitar el cómputo de las métricas, Mairano & Romano (2010) diseñaron el programa Correlatore, que trabaja en conjunto con los *textgrids* de Praat (Boersma & Weenink 2015). Esta herramienta agiliza el análisis acústico del habla, porque presenta las fórmulas de las principales métricas de duración. Es decir, una vez que se obtienen los *textgrids* de la segmentación de los fragmentos de habla, Correlatore arroja los índices métricos que se le soliciten de forma automática. Los valores obtenidos se pueden procesar en el mismo Correlatore para el diseño de gráficos o se pueden pasar a una planilla de cálculo para otros tipos de análisis estadísticos.

Las métricas rítmicas han recibido críticas en cuanto a su precisión a la hora de reflejar la estructura del ritmo. Los aportes de Arvaniti (2009, 2012) ponen en duda la medición del ritmo a través de los correlatos acústicos como lo hacen las métricas. Parte de sus argumentos están relacionados con la idea de que medir la duración es parte del componente rítmico, pero la duración no representa el ritmo en su totalidad. Es decir, el

componente fonológico no estaría representado en los índices métricos. Además, según la autora, estos índices no logran captar la naturaleza rítmica de las lenguas, porque existe mucha variación entre los hablantes, los tipos de tareas de elicitación y las cuestiones fonotácticas. De todos modos, hemos decidido adoptar los índices métricos debido a que han tenido resultados exitosos en la literatura, consignados en los distintos antecedentes que citamos a lo largo de nuestro trabajo. En otras palabras, seguimos los lineamientos de Langus, Mehler, & Nespor (2017), quienes plantean que la estructura rítmica queda plasmada en distintos niveles: el segmental, el de los pies métricos y el de las frases fonológicas. Por lo tanto, al analizar los índices métricos basados en la duración de los segmentos podemos acceder a la información rítmica de las lenguas.

Con respecto a las maneras de calcular los índices de las métricas, existen dos puntos que han dado lugar a resultados diversos en la literatura y que motivan, en parte, el desarrollo del presente trabajo: el tratamiento de las pausas y el de los intervalos ubicados antes de una pausa.

En relación con las pausas, éstas se pueden incluir o excluir en el cómputo de las métricas. Correlatore permite obtener valores para estas dos formas de considerar las pausas que se hallan en un mismo *textgrid*. Cabe aclarar que la ausencia de cualquier tipo de texto en un *textgrid* es interpretada por Correlatore como una pausa, es decir, una pausa corresponde a los casos en los que no se transcribe un intervalo vocálico |v| o consonántico |c|. Por medio del método A o global, este programa arroja los valores de las métricas midiendo todos los intervalos sin tener en cuenta las pausas. Por otra parte, a través del método B o local, Correlatore considera cada secuencia de segmentos por separado, obtiene un promedio de cada secuencia y, por último, suma todos los promedios.

En lo que concierne a los intervalos pre-pausa, se registran mayores diferencias entre los antecedentes sobre el ritmo, porque se tiene en cuenta que los segmentos finales presentan un alargamiento cuando se encuentran al final de una frase entonativa¹. Una postura sostiene que la exclusión de este intervalo permite obtener mejores distinciones entre las tipologías rítmicas (Loukina *et al.* 2009). Otra visión sobre este intervalo sugiere que es necesario incluir el intervalo final, porque este alargamiento que sufre este intervalo –mayor en algunas lenguas y menor en otras– es parte de la estructura rítmica (Grabe & Low 2002, Kireva & Gabriel 2015, White & Mattys 2007).

De todas las métricas propuestas en la literatura (véanse Fuchs 2016, Loukina *et al.* 2011, Wiget *et al.* 2010, para un resumen), en el presente trabajo adoptamos las métricas %V y VarcoV, porque han sido algunas de las más empleadas y han sido consideradas como las más sensibles para distinguir tipologías rítmicas (Fuchs 2016, White & Mattys 2007).

¹ Entendemos por frase entonativa la unidad prosódica mayor que presenta un patrón entonativo y que está delimitada por una pausa o por un tono de frontera (Beckman & Pierrehumbert 1986).

A modo de ejemplo para la métrica %V, el ritmo del español ha alcanzado un índice de 42,61 para el español peninsular y americano (Mairano 2011), de 48 para el español peninsular (White & Mattys 2007), de 46,94 para el español de Buenos Aires (Gabriel & Kireva 2014) y de 54 en Lima (O'Rourke 2008). Para el inglés, con la métrica %V se ha caracterizado el inglés estadounidense con un índice de 41,37 (Mairano 2011) y el inglés británico con un índice de 38 (White & Mattys 2007), de 40,1 (Ramus *et al.* 1999) y de 41,5 (Fuchs 2016). Por su parte, la métrica VarcoV ha permitido describir el ritmo del español con valores de 41 en la variedad peninsular (White & Mattys 2007), de 36 para el español de Lima (O'Rourke 2008) y de 49,68 para el español de Buenos Aires (Gabriel & Kireva 2014). El inglés estadounidense ha alcanzado valores de 51,33 para VarcoV (Mariano 2011) y el inglés británico valores de 64 (White & Mattys 2007) y de 51,7 (Fuchs 2016). Cabe destacar que los distintos índices arrojados por las métricas dependen tanto del tipo de dialecto o variedad lingüística como de la metodología empleada en cada investigación. Por lo tanto, la explicitación metodológica resulta fundamental a la hora de comparar resultados.

Por límites de espacio, en el presente trabajo no discutimos la cercanía o lejanía de nuestros índices métricos con los valores de otros antecedentes dentro de una misma lengua. Teniendo en cuenta las métricas %V y VarcoV y los tratamientos de las pausas y los intervalos finales antes de pausa, en este trabajo pretendemos demostrar qué métrica y qué tipo de tratamiento resulta ser más productivo para discriminar el ritmo silábico del español de la Patagonia argentina y el ritmo acentual del inglés proveniente de Inglaterra y de Estados Unidos. En la sección 2, a continuación, presentamos la metodología de nuestra investigación. Posteriormente, incluimos los resultados (sección 3) y el análisis de los datos en conjunto con algunas discusiones referidas a nuestros objetivos (sección 4). Culminamos con una breve conclusión (sección 5).

2. Metodología

2.1. Participantes

Se entrevistó a hablantes nativos del español patagónico (n=7) y a nativos de inglés de las variedades del sudeste de Inglaterra (n=7) y del noreste de Estados Unidos (n=7). Los hablantes nativos de español (3 mujeres y 4 hombres) tenían una edad promedio de 29 años (DE 6,7, máx.=39, mín=21) al momento de ser entrevistados, nacieron en Argentina y se criaron en la zona norte de la Patagonia. Sus segundas lenguas abarcaban inglés, francés, mapudungun, italiano y ruso, todas aprendidas en contexto de instrucción formal años después de haber adquirido la primera lengua. Solamente 2 informantes contaban con destrezas lingüísticas orales avanzadas en inglés, mientras que el resto reportó habilidades orales con un nivel inicial. Sus padres y tutores hablaban español argentino como lengua materna. En cuanto al plano socioeducativo, un informante contaba con estudios universitarios completos, mientras que los demás estaban en la mitad o culminando sus carreras universitarias.

El grupo de informantes de Inglaterra (4 mujeres, 3 hombres), por su parte, corresponde a hablantes del inglés estándar del sudeste de Inglaterra, con una edad promedio de 25 años (DE 3,8, máx.=32, mín.=21). En relación con sus lugares de nacimiento, 5 nacieron en la zona del Gran Londres, uno en Sudáfrica (con padres escoceses) y uno en Irlanda. Todos se criaron en los alrededores de Londres, con padres que hablaban el inglés como lengua materna². Los informantes reportaron las siguientes segundas lenguas con habilidades orales buenas: gaélico, portugués y tibetano; mientras que presentaban un nivel regular o pobre para español, francés, árabe y danés. Las segundas lenguas fueron adquiridas tanto en contextos naturales como en contexto de instrucción formal. Con respecto al nivel de educación alcanzado, 4 informantes presentaban estudios universitarios completos, mientras que 3 de ellos se encontraban terminando sus carreras de grado.

Por último, el grupo de informantes de Estados Unidos (4 mujeres, 3 hombres) presentó una edad promedio de 25 años (DE 5,1, máx.=34, mín.=18). Todos ellos nacieron y se criaron en el noreste de Estados Unidos, en los estados de Virginia, Nueva York, Massachusetts y el Distrito de Columbia. Sus padres nacieron en el mismo país y hablaban inglés como lengua materna. Con respecto a las segundas lenguas, algunos reportaron un nivel bueno y muy bueno en todas las destrezas lingüísticas de español y portugués y solo un informante declaró un nivel avanzado de francés. En la mayoría de los casos la adquisición de la L2 ocurrió en contexto de instrucción formal. Con respecto al nivel de educación, 5 informantes se hallaban en la mitad o culminando sus estudios universitarios, mientras que 2 de ellos presentaban estudios superiores completos.

2.2. Tareas de elicitación

Para la presente investigación computamos los datos acústicos de cuatro tareas de elicitación lingüística, dos tareas son de habla controlada y dos de habla espontánea. La tarea 1-LEC corresponde a la lectura de cinco oraciones³ en las que se controló la cantidad y tipos de sílabas entre las dos lenguas. En la tarea 2-REP los informantes debían escuchar fragmentos de habla sintetizada y repetirlos de manera más natural. La síntesis del habla se llevó a cabo a través de Praat con una velocidad de 115 palabras por minuto y con una pausa de 0,03 segundos entre palabras. Para estas dos tareas controladas, mostramos el total de estímulo de 1-LEC en la tabla 1 y ejemplos de 2-REP en la tabla 2.

² Solamente uno de los tutores hablaba portugués como primera lengua, pero la lengua de comunicación en el hogar era el inglés y el informante había aprendido portugués como segunda lengua.

³ Las oraciones se adoptaron de Ramus *et al.* (1999) y Prieto *et al.* (2012).

Tabla 1: Oraciones de la tarea de elicitación 1-LEC

Español	<ol style="list-style-type: none"> 1. Se enteraron de la noticia en este diario. 2. Tuve que ir a buscarlo lo más rápido posible 3. Los vecinos de mis abuelos son gente muy agradable. 4. Las madres salen cada vez más rápido de la maternidad. 5. El director declaró que la situación estaba bajo control.
Inglés	<ol style="list-style-type: none"> 1. A hurricane was announced this afternoon on the TV. 2. The committee will meet this afternoon for a special debate. 3. This rugby season promises to be a very exciting one. 4. Having a big car is not something I would recommend in this city. 5. The government is planning a reform of the educational program.

Tabla 2: Ejemplos de la tarea de elicitación 2-REP

Español	<ol style="list-style-type: none"> 1. hola soy yo de nuevo 2. quiero que sepas que estoy bien 3. y que ya no hago turnos durante el día en el trabajo
Inglés	<ol style="list-style-type: none"> 1. hi it is me again 2. first I want you to know that I am doing fine 3. and I am not working day shifts any more

Con la tarea 3-REL se obtuvo habla espontánea a través del relato de un cuento. Los informantes debían contar un cuento breve que se les presentó a través de una historia representada en una secuencia de 8 cuadros. Por último, la tarea 4-INS corresponde a habla espontánea obtenida por medio de instrucciones en las direcciones del mapa de una ciudad. Los informantes debían decir cómo llegar de un punto a otro del mapa, completando un total de 5 itinerarios. De las cuatro tareas de elicitación llevadas a cabo por los 21 informantes se obtuvo un total de 9315 sílabas en 1128 frases entonativas. Este corpus permitió la medición de aproximadamente 18000 intervalos en total. La tabla 3 muestra el detalle de producción lingüística obtenida por tarea y por grupo de informante⁴.

⁴ Para las tareas 3-REL y 4-INS se seleccionaron entre 90 y 100 sílabas por informante agrupadas entre 10 y 15 frases entonativas.

Tabla 3: Cantidad de frases entonativas y sílabas en el corpus analizado

Tarea		ES	II	IE	Subtotal
1-LEC	frases	63	48	45	156
	sílabas	564	551	551	1666
2-REP	frases	183	142	133	458
	sílabas	1463	1071	1071	3605
3-REL	frases	77	87	82	246
	sílabas	669	697	664	2030
4-INS	frases	80	92	96	268
	sílabas	684	663	667	2014
Total	frases	403	369	356	1128
	sílabas	3380	2982	2953	9315

2.3. Segmentación y medición

La segmentación entre intervalos vocálicos y consonánticos se llevó a cabo manualmente siguiendo los criterios de segmentación en la bibliografía consultada (Ramus *et al.* 1999, White & Mattys 2007, Fuchs 2016, Mairano 2011, Wiget *et al.* 2010, Mairano & Romano 2011, Prieto *et al.* 2012). Para identificar los límites entre vocales y consonantes nos basamos en el análisis perceptual y en la información espectrográfica. Dentro de las principales decisiones, podemos mencionar las siguientes: las deslizantes en posición prevocálica se consideraron como intervalos consonánticos; las deslizantes ubicadas después de la vocal se consideraron intervalos vocálicos; las consonantes silábicas del inglés se consideraron intervalos consonánticos; se tuvieron en cuenta todas las fases de las oclusivas (cierre, contención, explosión y aspiración); las consonantes elididas no se segmentaron; para los sonidos róticos se tuvo en cuenta el formante 3; por último, para delimitar las vocales se prestó atención a los formantes 1 y 2.

La medición de los intervalos obtenidos se llevó a cabo teniendo en cuenta las duraciones de cada intervalo. Como ya se adelantó, para este cómputo se emplearon dos métricas de duración. Por un lado, la métrica %V (Ramus *et al.* 1999) mide la proporción de los intervalos vocálicos, es decir, la suma de los intervalos vocálicos dividida por la duración total del enunciado (vocales + consonantes) multiplicado por 100. Por otro lado, la métrica VarcoV (Dellwo 2006, Dellwo & Wagner 2003, White & Mattys 2007) divide la desviación estándar de los intervalos vocálicos por la duración vocálica promedio y se multiplica por 100, controlando así la velocidad del habla.

2.4. Procedimiento

Las grabaciones se llevaron a cabo en tres universidades distintas. Las correspondientes al grupo de español se grabaron en una sala especial con aislación acústica de la Universidad Nacional del Comahue, en General Roca, Argentina. Los registros de los informantes de Inglaterra se tomaron en una sala de grabación de *University College* de Londres, mientras que los de los hablantes de Estados Unidos se realizaron en una sala silenciosa en *Hamilton College*, Clinton, Nueva York. Los micrófonos empleados fueron Røde NT1-A, Røde NT2-A y CAD GXL2200. Los informantes contaban con las tareas de elicitación en una presentación de PowerPoint y tenían la posibilidad de controlar la transición de las diapositivas. Las tareas se reordenaron al azar a modo de obtener distintas secuencias de actividades y evitar algún efecto de fatiga. Antes de comenzar a leer o hablar, los participantes tenían el tiempo suficiente para familiarizarse con las tareas. De cada informante se obtuvo un archivo de audio WAV sin comprimir, con una frecuencia de muestreo de 44.100 Hz y una cuantificación de 16 bits mono. Los archivos se editaron con Audacity (Equipo de Audacity 2015), quitando ruidos extralingüísticos. Posteriormente se procedió a etiquetar la segmentación de los fragmentos de habla utilizando Praat y se computaron los índices métricos de %V y VarcoV por medio de Correlatore. Los valores métricos se tabularon con el programa Infostat (Di Rienzo *et al.* 2008) a fin de llevar a cabo el tratamiento estadístico descriptivo, el diseño de gráficos y el análisis estadístico inferencial por medio de ANOVA y la prueba *post hoc* de Tukey.

3. Resultados

A continuación presentamos los resultados sobre cómo se diferencian los distintos grupos de informantes en las distintas tareas de elicitación. Mostramos primero los resultados del método A y luego los del método B, incluyendo y excluyendo el intervalo final para %V y VarcoV. Para medir el efecto de grupo de informante –y por lo tanto el efecto de lengua– se fijó *tipo de tarea* (1-LEC, 2-REP, 3-REL, 4-INS) como factor y *grupo de informantes* (ES, II e IE) como variable dependiente. Para el análisis estadístico inferencial se corrieron análisis ANOVA⁵ y para observar semejanzas significativas entre grupos se aplicó la prueba *post hoc* de Tukey⁶. Por último, concluimos cada subsección de los resultados con gráficos que muestran el ritmo en un espacio bidimensional por medio de las métricas en conjunto.

⁵Con el análisis ANOVA se logra observar cuán significativas son las diferencias de los grupos de valores según las variables. El nivel de significación es mayor cuando el estadístico *F* es alto y cuando el *p*-valor es bajo.

⁶Una prueba *post hoc*, como la de Tukey, permite observar qué grupos de valores presentan semejanzas o diferencias significativas. En nuestro análisis, las semejanzas se representan con una misma letra; por ejemplo cuando dos grupos mantienen la letra ‘a’. Las diferencias, en cambio, se expresan por medio de letras distintas; por ejemplo cuando un grupo presenta la letra ‘a’ y otro la letra ‘b’.

3.1. Resultados según el método A

La tabla 4 muestra los resultados de la métrica %V. Cuando se computó el intervalo final, esta métrica logró discriminar las lenguas en las tareas 2-REP ($F=18,97$; $p<,0001$) y 4-INS ($F=10,83$; $p=,0008$), mientras que cuando no se computó este intervalo, solamente diferenció las lenguas en la tarea 3-REL ($F=8,79$; $p=,0022$). La prueba *post hoc* diferenció el grupo de español (letra ‘b’), por un lado, y las dos variedades del inglés (letra ‘a’), por el otro.

Tabla 4: %V y efecto del tipo de lengua con el método A

Tareas	ES	II	IE	F	p-valor	Post hoc ^a		
						ES	II	IE
con intervalo final								
1-LEC	44,88 (2,07)	43,07 (3,52)	43,30 (1,14)	1,14	0,3431	a	a	a
2-REP	48,22 (3,46)	39,89 (2,54)	40,74 (2,20)	18,97	<0,0001**	b	a	a
3-REL	50,41 (1,16)	49,80 (1,63)	49,37 (2,12)	0,68	0,5207	a	a	a
4-INS	49,84 (1,66)	43,67 (3,43)	45,26 (2,32)	10,83	0,0008*	b	a	a
sin intervalo final								
1-LEC	44,09 (2,38)	42,23 (3,00)	43,21 (1,22)	1,13	0,3463	a	a	a
2-REP	47,56 (2,93)	44,34 (2,65)	45,53 (2,56)	2,51	0,1096	a	a	a
3-REL	47,78 (2,20)	51,35 (1,38)	52,01 (2,38)	8,79	0,0022*	b	a	a
4-INS	45,57 (2,28)	45,94 (4,51)	48,43 (1,83)	1,17	0,2003	a	a	a

^aLas letras en común no son significativamente diferentes ($\alpha=0,05$).

*Existe una diferencia significativa entre lenguas.

**Existe una diferencia altamente significativa entre lenguas.

Según los resultados de la tabla 5, la métrica VarcoV logró discriminar las lenguas en tres de las cuatro tareas. En relación con la tarea 1-LEC, esta métrica reflejó un efecto significativo de lenguas tanto en los casos con cómputo del intervalo final ($F=23,69$; $p<,0001$) como sin cómputo de éste ($F=10,70$; $p=,0009$). En cambio, en la tarea 2-REP no hubo una diferencia significativa de lenguas ($p>,05$) en ninguno de los dos tratamientos del intervalo final. La tarea 3-REL, por su parte, reflejó un efecto de lengua cuando se incluyó el intervalo final ($F=6,44$; $p=,0078$) y cuando se excluyó este intervalo ($F=12,04$; $p=,0005$). Por último, la tarea 4-INS no presentó un efecto de tipo de lengua cuando se computó el intervalo final ($p>,05$), pero sí lo hizo cuando no se lo computó ($F=7,00$; $p=,0056$). En todos los casos en los que hubo una distinción significativa, la prueba *post hoc* separó el grupo de español (letra ‘a’), por un lado, y las dos variedades del inglés (letra ‘b’), por el otro.

Tabla 5: VarcoV y efecto del tipo de lengua con el método A

Tareas	ES	II	IE	F	p-valor	Post hoc ^a		
						ES	II	IE
con intervalo final								
1-LEC	41,22 (5,90)	58,52 (4,05)	54,33 (4,59)	23,69	<0,0001**	a	b	b
2-REP	53,11 (8,04)	57,32 (5,64)	56,11 (5,21)	0,80	0,4664	a	a	a
3-REL	51,82 (10,52)	66,17 (9,99)	67,53 (6,04)	6,44	0,0078*	a	b	b
4-INS	53,46 (10,13)	65,20 (9,91)	66,00 (10,39)	3,35	0,0580	a	a	a
sin intervalo final								
1-LEC	41,11 (6,47)	54,05 (3,74)	51,49 (6,02)	10,70	0,0009*	a	b	b
2-REP	52,47 (8,01)	57,10 (5,82)	55,84 (5,39)	0,95	0,4060	a	a	a
3-REL	45,96 (9,44)	64,37 (9,08)	64,31 (4,98)	12,04	0,0005*	a	b	b
4-INS	47,17 (8,19)	61,86 (9,83)	64,60 (9,99)	7,00	0,0056*	a	b	b

^aLas letras en común no son significativamente diferentes ($\alpha=0,05$).

*Existe una diferencia significativa entre lenguas.

**Existe una diferencia altamente significativa entre lenguas.

Habiendo mostrado los resultados de las métricas por separado, a continuación incluimos los gráficos de las métricas en conjunto a modo de visualizar las diferencias rítmicas de los grupos de informantes según el método A.

Método A con intervalo final

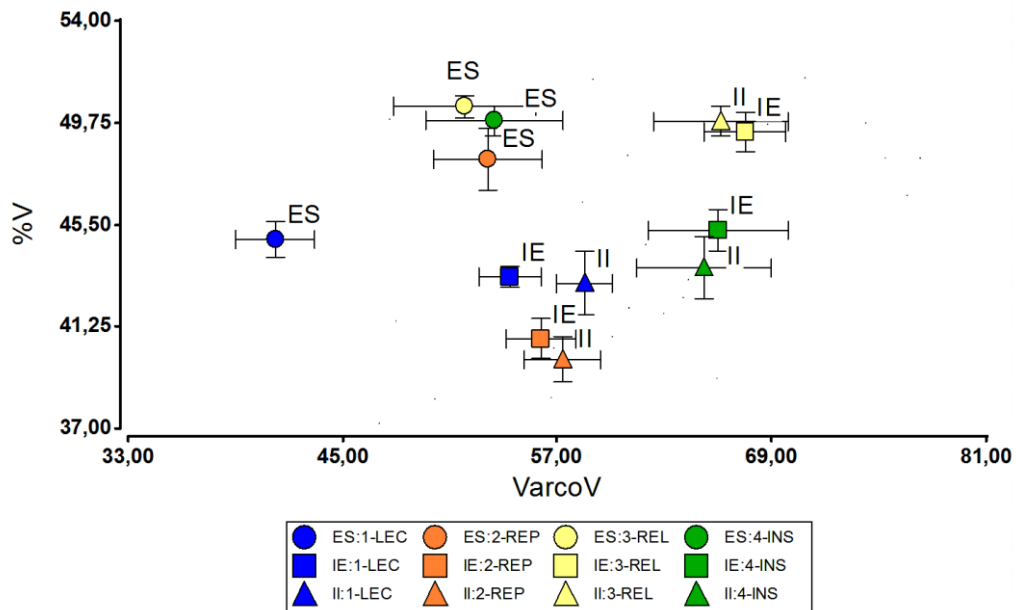


Gráfico 1: El ritmo según el método A con cómputo del intervalo final.

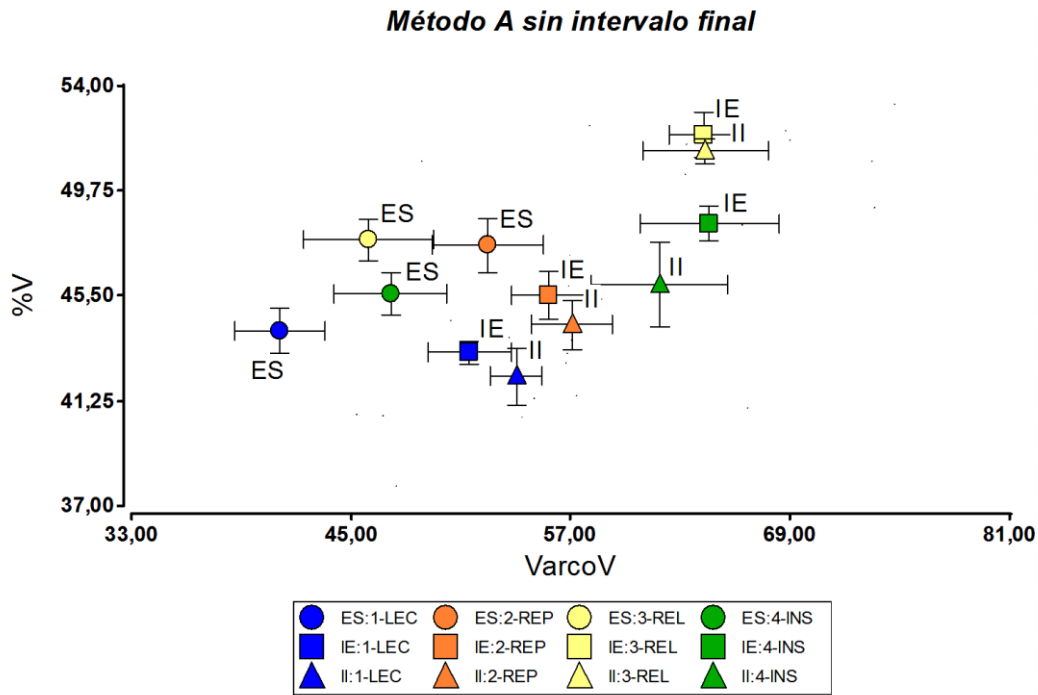


Gráfico 2: El ritmo según el método A sin cómputo del intervalo final.

En los gráficos 1 y 2 podemos observar que los valores métricos para el español se ubican en la zona izquierda de los gráficos, mientras que los valores para las dos variedades del inglés se hallan en dirección hacia la derecha. Por otra parte, la mayor separación entre el español y el inglés se visualiza en los datos con cómputo del intervalo final, aunque, recordemos, la discriminación significativa entre lenguas solamente se dio en algunas tareas.

3.2. Resultados según el método B

La tabla 6 contiene los valores de la métrica %V según el método B. Al igual que los resultados del método A, esta métrica diferenció las lenguas solamente en algunas tareas. Cuando se incluyó el intervalo final, hubo un efecto significativo de lenguas en las tareas 2-REP ($F=15,82$; $p=,0001$) y 4-INS ($F=9,54$; $p=,0015$). En cambio, cuando el intervalo final se excluyó, la métrica %V solamente reflejó una diferencia de lenguas en la tarea 3-REL ($F=11,19$; $p=,0007$). En los casos en los que las lenguas se diferenciaron de manera significativa, la prueba *post hoc* separó el grupo ES (letra 'b') de los grupos II e IE (letra 'a').

Tabla 6: % V y efecto de tipo de lengua con el método B

Tareas	ES	II	IE	F	p-valor	Post hoc ^a		
						ES	II	IE
con intervalo final								
1-LEC	44,86 (1,93)	43,54 (4,08)	43,27 (1,14)	5,11	0,5069	a	a	a
2-REP	48,64 (3,45)	41,03 (2,49)	41,91 (2,22)	15,82	0,0001*	b	a	a
3-REL	50,76 (1,33)	50,00 (1,67)	49,61 (2,31)	0,73	0,4937	a	a	a
4-INS	50,40 (2,17)	44,28 (3,23)	45,66 (2,75)	9,54	0,0015*	b	a	a
sin intervalo final								
1-LEC	43,82 (2,51)	42,41 (3,35)	43,31 (1,25)	0,55	0,5842	a	a	a
2-REP	48,29 (2,75)	45,29 (2,38)	46,76 (2,46)	2,45	0,1148	a	a	a
3-REL	47,37 (2,37)	51,88 (1,75)	52,67 (2,59)	11,19	0,0007*	b	a	a
4-INS	45,54 (2,31)	46,52 (4,25)	49,45 (2,11)	3,12	0,0687	a	a	a

^aLas letras en común no son significativamente diferentes ($\alpha=0,05$).

*Existe una diferencia significativa entre lenguas.

Siguiendo con los resultados del método B, la tabla 7 muestra los valores de la métrica VarcoV. Esta métrica demuestra un efecto significativo del tipo de lengua en todas las tareas y en los dos tratamientos del intervalo final. Con respecto a la tarea 1-LEC, la diferenciación de las lenguas fue altamente significativa tanto con cómputo del intervalo final ($F=60,60$; $p<,0001$) como sin éste ($F=31,29$; $p<,0001$). En la tarea 2-REP, la significación fue un tanto menor ($F=6,24$; $p=,0087$ con intervalo final y $F=7,59$; $p=,0041$ sin intervalo final). La tarea 3-REL diferenció las lenguas de manera significativa en los datos con cómputo del intervalo final ($F=8,14$; $p=,0030$) y de manera altamente significativa cuando no se incluyó este intervalo ($F=18,37$; $p<,0001$). Por último, la tarea 4-INS también mostró una distinción entre lenguas de manera significativa ($F=8,76$; $p=,0022$ con cómputo del intervalo final y $F=13,77$; $p=,0002$ sin cómputo de este intervalo). En todos los casos, la prueba *post hoc* logró discriminar los informantes del español (letra 'a') de los informantes del inglés (letra 'b').

Tabla 7: VarcoV y efecto de tipo de lengua con el método B

Tareas	ES	II	IE	F	p-valor	Post hoc ^a		
						ES	II	IE
con intervalo final								
1-LEC	34,81 (2,17)	57,96 (4,34)	53,72 (5,39)	60,60	<0,0001**	a	b	b
2-REP	42,28 (7,01)	53,63 (7,44)	52,24 (4,96)	6,24	0,0087*	a	b	b
3-REL	46,88 (10,35)	64,52 (10,53)	64,82 (7,35)	8,14	0,0030*	a	b	b
4-INS	43,28 (7,12)	61,69 (12,41)	61,99 (8,39)	8,76	0,0022*	a	b	b
sin intervalo final								
1-LEC	34,52 (2,21)	51,43 (4,40)	51,45 (6,31)	31,29	<0,0001**	a	b	b
2-REP	41,67 (5,73)	52,82 (7,06)	52,04 (4,96)	7,59	0,0041*	a	b	b
3-REL	40,12 (6,82)	61,36 (9,49)	61,23 (5,85)	18,37	<0,0001**	a	b	b
4-INS	36,02 (6,32)	58,43 (13,19)	59,55 (7,40)	13,77	0,0002*	a	b	b

^aLas letras en común no son significativamente diferentes ($\alpha=0,05$).

*Existe una diferencia significativa entre lenguas.

**Existe una diferencia altamente significativa entre lenguas.

Luego de haber mostrado los resultados del método B con las métricas por separado, presentamos a continuación los valores de las métricas %V y VarcoV en forma conjunta por medio de gráficos. En el gráfico 3 mostramos las métricas con cómputo del intervalo final, mientras que en el gráfico 4 lo hacemos sin el cómputo de este intervalo.

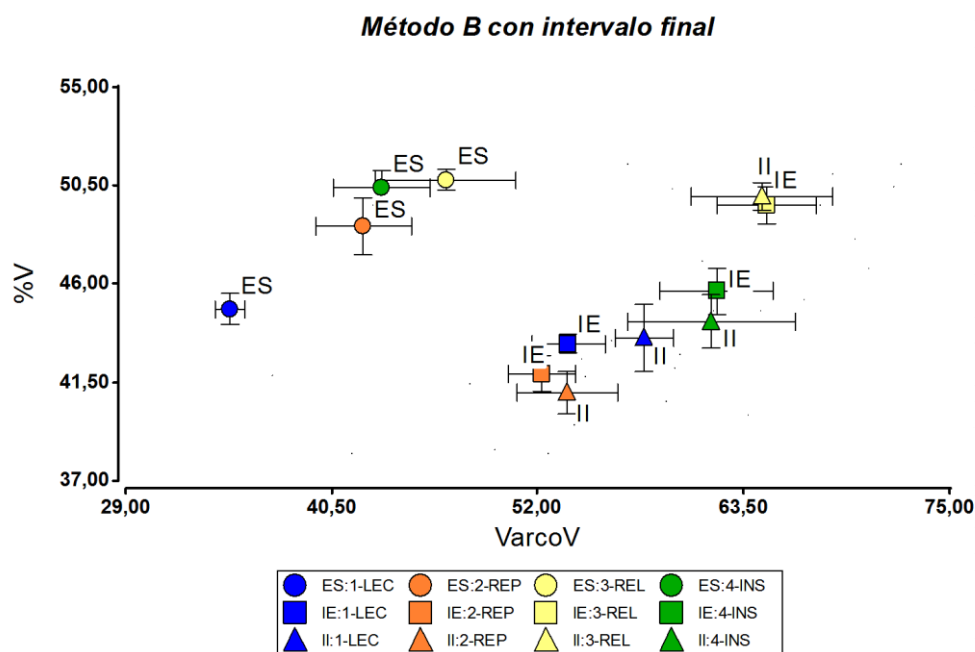


Gráfico 3: El ritmo según el método B con cómputo del intervalo final.

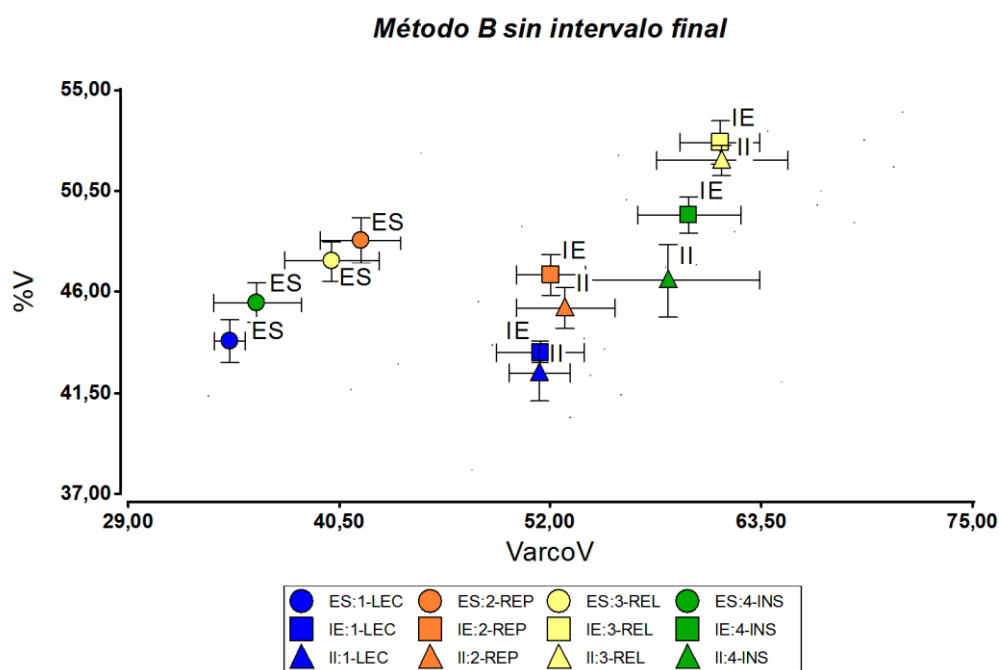


Gráfico 4: El ritmo según el método B sin cómputo del intervalo final.

Los gráficos 3 y 4 muestran que los valores del español se concentran en la zona izquierda de los gráficos, mientras que los dialectos del inglés se agrupan en la zona de la derecha. Cabe resaltar que esta ubicación de las dos lenguas refleja una clara separación entre un ritmo silábico y un ritmo acentual.

4. Análisis y discusión

La presente investigación permite discutir cuestiones metodológicas sobre cómo dar cuenta del ritmo de habla en las lenguas. Por un lado, contar con herramientas acústicas para distinguir lenguas como el español y el inglés sienta las bases para el análisis de lenguas o variedades lingüísticas que no cuentan con una descripción rítmica acabada. Por ejemplo, si después de adoptar los mismos criterios metodológicos una lengua o variedad lingüística arroja valores más cercanos al español y no al inglés, estaríamos en presencia de un ritmo silábico y aportaríamos datos al conocimiento de la clasificación lingüística. Por otro lado, saber que ciertas decisiones en la metodología de la medición del ritmo son productivas para la diferenciación rítmica habilita la comparación de la producción lingüística en hablantes de segundas lenguas en estudios longitudinales o transversales, tanto en contextos de adquisición natural como en contexto de instrucción. A continuación, hacemos un resumen de los resultados e incluimos discusiones en torno a los tres aspectos metodológicos que hemos abordado: las métricas, los métodos para las pausas y la inclusión o exclusión del intervalo final.

4.1. Sobre las métricas %V y VarcoV

La métrica %V resultó ser menos productiva para diferenciar las lenguas. En ninguno de los métodos se registró una diferencia de lenguas para la 1-LEC, probablemente porque las lecturas no se diferenciaron lo suficiente entre las lenguas como para ser reflejadas en esta métrica. Cuando se computó el intervalo final, las lenguas mostraron una diferencia significativa en las tareas 2-REP y 4-INS en ambos métodos. En cambio, cuando se excluyó el último intervalo, las lenguas se diferenciaron solamente en 3-REL también en ambos métodos. El análisis *post hoc* siempre agrupó los datos del español, por un lado, y del inglés de Inglaterra y Estados Unidos, por el otro. Atribuimos esta falta de diferenciación sistemática de la métrica %V a la influencia que tiene la velocidad de habla o tempo en las distintas tareas de elicitación. Tal como se ha señalado en la literatura (Dellwo & Wagner 2003, Loukina *et al.* 2013), esta métrica es sensible al estilo de habla y, por lo tanto, la discriminación rítmica de las lenguas deja de ser categórica. En nuestros datos podemos observar que las lenguas se asemejan en la lectura por el hecho de no demostrar una diferencia estadísticamente significativa, mientras que en las demás tareas la diferenciación depende generalmente de la inclusión o exclusión del intervalo final.

Por otra parte, la métrica VarcoV mostró un efecto del tipo de lengua en casi todos los casos del método A y en todas las tareas del método B. La ventaja de VarcoV por sobre %V se debe a la normalización del habla (Fuchs 2016, White & Mattys 2007). Es decir, al neutralizar el efecto de la velocidad del habla producida por los distintos estilos de habla que genera una lectura o repetición en comparación con tareas espontáneas, se logra medir la estructura rítmica de mejor manera y, por lo tanto, las lenguas logran diferenciarse significativamente.

4.2. Sobre los métodos A y B de Correlatore

Para el análisis de las lenguas, la métrica %V no mostró diferencias entre los dos métodos. Esta métrica solo logró distinguir las lenguas en las tareas 2-REP y 4-INS con cómputo del intervalo final y solamente en la tarea 3-REL sin cómputo de este intervalo. La métrica VarcoV, en cambio, presentó una diferencia entre los dos métodos. Por su parte, en los datos del método A, VarcoV solamente discriminó las lenguas en las tareas 1-LEC, 3-REL y 4-INS (y no en 2-REP), mientras que en los datos del método B esta métrica demostró una diferencia significativa entre el español y las dos variedades del inglés en todos los casos. En consecuencia, estas diferencias entre los métodos de Correlatore nos llevan a concluir que el método B presenta ventajas a la hora de diferenciar tipologías rítmicas. Si bien para Loukina *et al.* (2009) no se registran diferencias significativas entre los dos tratamientos de las pausas, nuestros datos arrojan evidencia para proponer el empleo del método B en Correlatore, porque se tienen en cuenta las pausas de cada *textgrid* y se computan los promedios de cada secuencia de intervalos entre pausas. Los resultados de nuestra investigación permiten confirmar la sugerencia de Mairano (2011: 50) sobre una mejor clasificación rítmica por medio de este método.

4.3. Sobre el cómputo del intervalo final

La inclusión o no de este intervalo implicó resultados distintos. La información de los segmentos finales resulta más relevante para el empleo de la métrica %V. Considerando que los intervalos que preceden una pausa están afectados por una forma de reforzamiento manifestado por la prolongación o acortamiento (Fougeron & Keating 1997), explicamos esta mayor distinción entre el español y el inglés con cómputo del intervalo final como consecuencia de que el inglés presenta mayor alargamiento al final de frase que el español (Ortega-Llebaria & Prieto 2007, Prieto *et al.* 2012, Wightman *et al.* 1992). Así, podemos reforzar la idea de que los segmentos finales dan cuenta de las particularidades del ritmo, tal como se ha sugerido en la literatura (Grabe & Low 2002, Kireva & Gabriel 2015, White & Mattys 2007).

La métrica VarcoV, por su parte, logró discriminar las lenguas mayormente cuando no se computó el intervalo final. Para el método A, la exclusión del intervalo final reflejó un efecto significativo de lenguas en mayores instancias. Con respecto al método B, si bien las lenguas se diferenciaron significativamente tanto en los casos con cómputo del intervalo final como sin éste, la ventaja de la exclusión del intervalo final en el método B se debió a que hubo un nivel de significación levemente mayor en las tareas 2-REP, 3-REL y 4-INS, expresado por valores mayores del estadístico *F* y menores del *p*-valor. Interpretamos estos datos como la capacidad de la métrica VarcoV de reflejar la estructura rítmica de manera más productiva cuando se excluye el intervalo final. Asimismo, podemos entender el alargamiento final como un fenómeno un tanto independiente que ayuda a diferenciar las lenguas con la métrica %V, pero que para VarcoV provoca una leve semejanza entre las lenguas, ya que la exclusión del intervalo final permite una diferenciación un tanto más significativa.

5. Conclusión

Asumiendo que el ritmo es la consecuencia de distintas propiedades fonético-fonológicas, el español y el inglés representan dos tipologías rítmicas distintas –el ritmo silábico y el ritmo acentual, respectivamente– que son medibles acústicamente por medio de las métricas rítmicas de duración. Nuestros datos han indicado que la variedad de español elegida (patagónico) y las dos variedades del inglés (británico y estadounidense) reflejan una diferencia estadísticamente significativa, dependiendo del tipo de métrica y del tratamiento de las pausas y los intervalos finales. Con respecto a las métricas empleadas, la normalización de los datos ha demostrado ser una ventaja para diferenciar las dos lenguas. Es decir, la métrica VarcoV ha logrado diferenciar de manera categórica los dos tipos de ritmo y no así la métrica %V, porque esta última está afectada por la velocidad del habla. En lo que refiere al tratamiento de las pausas, nuestros datos han demostrado que cuando las pausas se tienen en cuenta, las lenguas se diferencian en mayor medida. Por lo tanto, el método B de Correlatore resulta ser más productivo para caracterizar ritmos distintos. Por último, el cómputo del intervalo final permite mayores casos de diferenciación rítmica cuando se emplea la métrica %V, mientras que para VarcoV el grado de significación resulta más alto cuando se excluye

este intervalo. En suma, nuestros datos arrojan evidencia a favor de la métrica VarcoV con exclusión de los intervalos finales por medio del método B. Así, esperamos haber realizado un aporte metodológico a los estudios acústicos del ritmo de habla, a modo de poder distinguir lenguas con tipologías rítmicas diferentes como el español en contraste con dos variedades principales del inglés.

Referencias

- Abercrombie, David (1967). *Elements of General Phonetics*. Edinburgh University Press.
- Allen, George D. (1975). Speech rhythms: Its relation to performance universals and articulatory timing. *Journal of Phonetics* 3 (2): 75–86.
- Arvaniti, Amalia (2009). Rhythm, Timing and the Timing of Rhythm. *Phonetica* 66 (1–2): 46–63. <https://doi.org/10.1159/000208930>
- Arvaniti, Amalia (2012). The usefulness of metrics in the quantification of speech rhythm. *Journal of Phonetics* 40 (3): 351–373. <https://doi.org/10.1016/j.wocn.2012.02.003>
- Beckman, Mary E. & Janet B. Pierrehumbert (1986). Intonational structure in Japanese and English. *Phonology* 3: 255–309.
- Bertinetto, Pier Marco (1981). *Strutture prosodiche dell'italiano: accento, quantità, sillaba, giuntura, fondamenti metrici*. Firenze: Presso l'Accad. della Crusca.
- Bertinetto, Pier Marco (1989). Reflections on the Dichotomy “Stress” vs. “Syllable-timing”. *Revue de Phonétique Appliquée* 91-93: 99–130.
- Boersma, Paul & David Weenink (2015). *Praat: doing phonetics by computer*. Programa versión 5.4.08. www.praat.org.
- Borzone de Manrique, Ana María & Angela Signorini (1983). Segmental durations and the rhythm in Spanish. *Journal of Phonetics* 11: 117–128.
- Dasher, Richard & Dwight Bolinger (1982). On pre-accentual lengthening. *Journal of the International Phonetic Association* 12: 58–69.
- Dauer, Rebecca (1983). Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics* 11 (1): 51–62.
- Dauer, Rebecca (1987). Phonetic and phonological components of language rhythm. *11th International Congress of Phonetic Sciences*. Estonia: Tallinn, 447–450.
- Dellwo, Volker & Petra Wagner (2003). Relations between language rhythm and speech rate. En Maria-Josep Solé, Daniel Recasens i Vives & Joaquin Romero (eds.) *Proceedings of International Congress of Phonetic Science*. Barcelona: Casual Productions, 471–474.
- Dellwo, Volker (2006). Rhythm and Speech Rate: A Variation Coefficient for deltaC. En Pawel Karnowski & Imre Szigeti (eds.) *Language and language-processing*. Frankfurt am Main: Peter Lang, 231–241.
- Di Rienzo, Julio A., Fernando Casanoves, Mónica G. Balzarini, Laura A. González, Elena M. Tablada & Carlos W. Robledo (2008). *InfoStat*. Versión 2008. Universidad Nacional de Córdoba, Argentina: Grupo InfoStat, FCA.
- Equipo de Audacity (2015). *Audacity (R): Free Audio Editor and Recorder*. Versión 2.1.1. <http://old.audacityteam.org/>.
- Fougeron, Cécile & Patricia A. Keating (1997). Articulatory strengthening at edges of prosodic domains. *The Journal of the Acoustical Society of America* 101 (6): 3728–3740.
- Fuchs, Robert (2016). *Speech rhythm in varieties of English: Evidence from educated Indian English and British English*. Singapore: Springer.

- Gabriel, Christoph & Elena Kireva (2014). Prosodic Transfer in Learner and Contact Varieties. *Studies in Second Language Acquisition* 36 (02): 257–281. <https://doi.org/doi:10.1017/S0272263113000740>
- Grabe, Esther & Ee Ling Low (2002). Durational Variability in Speech and the Rhythm Class Hypothesis. En Natasha Warner & Carlos Gussenhoven (eds.) *Papers in Laboratory Phonology* 7. Berlin: Mouton de Gruyter, 515–546.
- Kireva, Elena & Christoph Gabriel (2015). Rhythm properties of a contact variety: Comparing read and semi-spontaneous speech in Argentinean Porteño Spanish. En Elizabeth Delais-Roussarie, Mathieu Avanzi & Sophie Herment (eds.) *Prosody and Language in Contact: L2 Acquisition, Attrition and Languages in Multilingual Situations*. Berlin: Springer, 149–168.
- Langus, Alan, Jacques Mehler & Marina Nespors (2017). Rhythm in language acquisition. *Neuroscience & Biobehavioral Reviews* 81: 158–166. <https://doi.org/10.1016/j.neubiorev.2016.12.012>
- Lehiste, Ilse (1973). Rhythmic units and syntactic units in production and perception. *The Journal of the Acoustical Society of America* 54: 1228–1234.
- Loukina, Anastassia, Greg Kochanski, Burton Rosner, Elinor Keane & Chilin Shih (2011). Rhythm measures and dimensions of durational variation in speech. *The Journal of the Acoustical Society of America* 129 (5): 3258–3270. <https://doi.org/10.1121/1.3559709>.
- Loukina, Anastassia, Greg Kochanski, Burton Rosner, Elinor Keane & Chilin Shih (2013). What determines duration-based rhythm measures: text or speaker? *Laboratory Phonology* 4 (2): 339–382.
- Loukina, Anastassia, Greg Kochanski, Chilin Shih, Elinor Keane & Ian Watson (2009). Rhythm measures with language-independent segmentation. *Proceedings of Interspeech 2009*, 1931–34.
- Mairano, Paolo & Antonio Romano (2010). Un confronto tra diverse metriche ritmiche usando Correlatore. En Stephan Schmid, Michael Schwarzenbach & Dieter Studer (eds.) *La dimensione temporale del parlato (Proc. of the V National AISV Congress, University of Zurich, Collegiengebaude)*. Zurich, Switzerland, 79–100.
- Mairano, Paolo & Antonio Romano (2011). Rhythm metrics for 21 languages. En *ICPhS XVII*. Hong Kong, 1318–1321.
- Mairano, Paolo (2011). *Rhythm typology: acoustic and perceptive studies*. Tesis doctoral, University of Turin.
- O'Rourke, Erin (2008). Speech rhythm variation in dialects of Spanish: applying the pairwise variability index and variation coefficients to Peruvian Spanish. *Proc. Fourth Conf. on Speech Prosody*, Campinas, Brasil. 431–434.
- Ortega-Llebaria, Marta & Pilar Prieto (2007). Stress and focus in Spanish and Catalan: patterns of duration and vowel quality. En Pilar Prieto, Joan Mascaró & Maria-Josep Solé (eds.) *Segmental and Prosodic Issues in Romance Phonology. Current Issues in Linguistic Theory Series*. Amsterdam/Philadelphia: John Benjamins, 155–175.

- Pamies Bertrán, Antonio (1999). Prosodic typology: on the dichotomy between stress-timed and syllable-timed languages. *Language Design* 2: 103–130.
- Pike, Kenneth Lee (1945). *The intonation of American English*. Ann Arbor: University of Michigan Press.
- Prieto, Pilar, Maria Del Mar Vanrell, Lluïsa Astruc, Elinor Payne & Brechtje Post (2012). Phonotactic and phrasal properties of speech rhythm. Evidence from Catalan, English, and Spanish. *Speech Communication* 54 (6): 681–702. <https://doi.org/10.1016/j.specom.2011.12.001>.
- Ramus, Franck, Marina Nespor & Jacques Mehler (1999). Correlates of linguistic rhythm in the speech signal. *Cognition* 73: 265–92. [https://doi.org/10.1016/S0010-0277\(99\)00058-X](https://doi.org/10.1016/S0010-0277(99)00058-X).
- Roach, Peter (1982). On the distinction between ‘Stress-timed’ and ‘Syllable-timed’ Languages. En David Crystal (ed.) *Linguistic controversies*. London: Edward Arnold. 73–79.
- White, Laurence & Sven L. Mattys (2007). Calibrating rhythm: First language and second language studies. *Journal of Phonetics* 35 (4): 501–522. <https://doi.org/10.1016/j.wocn.2007.02.003>.
- Wiget, Lukas, Laurence White, Barbara Schuppler, Izabelle Grenon, Olesya Rauch & Sven L. Mattys (2010). How stable are acoustic metrics of contrastive speech rhythm? *The Journal of the Acoustical Society of America* 127 (3): 1559–1569. <https://doi.org/10.1121/1.3293004>.
- Wightman, Colin W., Stefanie Shattuck-Hufnagel, Mari Ostendorf & Patti J. Price (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *The Journal of the Acoustical Society of America* 91 (3): 1707–1717.