



PIER MARCO BERTINETTO & CHIARA BERTINI

Modelización del ritmo y estructura silábica, con aplicación al italiano

"contributo al volume in onore di Carmen Pensado Ruiz"

1 La hipótesis «control vs. compensación».

Los estudios sobre la estructura rítmica de las lenguas naturales encontraron nuevo impulso en años recientes desde la propuesta de Ramus & Nespors & Mehler (1999), en seguida adoptada y desarrollada por otros autores, en particular por Low & Grabe & Nolan (2000) y Grabe & Low (2002). Hubo también aplicaciones al italiano y sus dialectos (Russo & Barry 2003; 2004; Schmid 2004; Mairano & Romano 2006).

La discusión se concentró sobre la búsqueda de un algoritmo capaz de representar con éxito las características rítmicas de las lenguas, dando forma concreta a la intuición de Pike (1945) en cuanto al contraste entre lenguas iso-accentuales y lenguas isosilábicas («syllable- vs. stress-timed»). Los sucesivos estudios han aclarado que la interpretación literal de esta dicotomía no encuentra confirmación en la realidad fonética. A pesar de esto, es opinión casi general que la intuición de Pike es fundamentalmente correcta, en el sentido de que las lenguas naturales se distinguen en relación a su orientación rítmica. No faltaron, por lo demás, intentos de reinterpretar esta oposición en formas nuevas,

relacionando un conjunto de comportamientos fonéticos con un conjunto de características fonológicas (Bertinetto 1981; Dauer 1983; Bertinetto 1989).

Según la hipótesis propuesta por Bertinetto & Vékás (1991), este complejo de rasgos fonéticos y fonológicos podría encontrar un punto de convergencia en el contraste entre lenguas «de control» y lenguas «de compensación». La idea de base, ya señalada en Bertinetto & Fowler (1989), inspirada por el trabajo de Fowler (1977), es que las lenguas pueden diferenciarse según la manera en que los gestos vocálicos y consonánticos se combinan en la cadena articulatoria. Una lengua «de control» ideal tendría que concebirse como una lengua en la cual todos los segmentos reciben un idéntico esfuerzo articulatorio y tienen (idealmente) la misma duración. Por supuesto, esto es imposible, a causa de la variedad de puntos y modos de articulación; pero esta hipótesis tiene plausibilidad, si consideramos que las lenguas se comportan en efecto de manera diferente en relación a la combinación de gestos vocálicos y consonánticos. Algunas lenguas admiten un nivel de coarticulación más grande que otras (Goldstein & Chitoran & Selkirk 2007).¹ Esto se asemeja a propuestas recientes (O'Dell & Nieminen 1999; Barbosa 2006), donde las contribuciones respectivas de los osciladores silábico y acentual se modelan algorítmicamente. En efecto, estas últimas propuestas pueden integrarse con la hipótesis de «control vs. compensación» (desde ahora: CC). Por supuesto, CC quiere describir el comportamiento fonotáctico (que con mucha aproximación podríamos llamar «silábico»), que afecta a la sobre-ordenada alternancia acentual (o sea, el nivel considerado por O'Dell & Nieminen y Barbosa) y, a la vez, es afectado por ella. Cuál de los dos componentes es el predominante, es algo que queda – por el momento – sin aclarar. Bertinetto & Bertini (*en preparación*) proponen un modelo explícito del ritmo de las lenguas naturales que incluye crucialmente el papel de dos niveles: (i) el nivel silábico, donde tiene lugar la acción de los osciladores consonántico y vocálico acoplados (Goldstein & Chitoran & Selkirk 2007); (ii) el nivel oracional, donde se efectúa la

¹ Por supuesto, los tipos rítmicos no se limitan a dos, como en la hipótesis de Pike. En particular, muchos autores han hablado de «mora-timing», además de «stress-timing» y «syllable-timing». En la concepción de CC aquí propuesta, sin embargo, el tipo moraico podría constituir la forma extrema del tipo «de control». De toda manera, la distinción entre «control» y «compensación» tiene que concebirse como una gradación continua sobre un eje polarmente orientado, más que como una verdadera dicotomía.

acción de los osciladores acentual y silábico acoplados (O'Dell & Nieminen 1999). En este trabajo consideramos solamente el primer nivel.

Si la hipótesis de CC es correcta, eso tendría que comportar algunas importantes consecuencias desde el punto de vista de las características articulatorias y prosódicas. Entre las posibles predicciones, podemos enumerar las siguientes:

- articulatoriamente (como ya se ha dicho), tendría que haber una mayor coarticulación – es decir superposición – de los gestos vocálicos y consonánticos en las lenguas «de compensación»; lo que produce, por consiguiente, menor correspondencia de fase entre el oscilador consonántico y el oscilador vocálico;

- prosódicamente, las duraciones de los segmentos tendrían que regularse de manera diversa: en las lenguas «de control», los segmentos deberían manifestar una mayor incompresibilidad, a pesar de las variaciones en la velocidad de elocución. Esto no quiere decir, por supuesto, que estas lenguas no puedan tolerar fuertes variaciones de velocidad, sino que el nivel de coarticulación entre gestos adyacentes no puede sobrepasar un cierto límite, sobre todo en cuanto a los segmentos vocálicos. Por consiguiente, la reducción de timbre en las vocales átonas tendría que encontrar un límite más alto que en el caso de las lenguas «de compensación», donde (en los casos extremos) los segmentos vocálicos pueden hasta dar la impresión de desaparecer del espectrograma, a pesar de que sus consecuencias perceptivas permanezcan, precisamente como efecto de la co-producción con las consonantes adyacentes (Bertinetto & Fowler 1989).

Naturalmente, estas diferencias se manifiestan como gradaciones en un continuo, más bien que como contraposiciones nítidas. De toda manera, si la hipótesis general es correcta, se puede suponer que la duración de la sílaba en las lenguas «de control» está más directamente relacionada (en comparación con las lenguas «de compensación») con el número de segmentos que la componen, a causa de la menor coarticulación. Esto produce una consecuencia aparentemente paradójica: es decir, tenemos que esperar una mayor compensación de la duración intra-silábica en las lenguas «de compensación» que en las lenguas «de control», a pesar de que las primeras correspondan al tipo iso-accentual de la oposición tradicional y las últimas al tipo iso-silábico (donde precisamente se suponía una duración silábica constante).

La Figura 1 muestra una representación muy esquemática de las tendencias descritas. Si la hipótesis CC es correcta, una lengua «de

control» ideal tendría que mostrar un mismo nivel de fluctuación de la duración para las vocales y las consonantes (es decir, su colocación en el plano tendría que acercarse a la bisectriz), mientras que las lenguas «de compensación» tendrían que mostrar una fluctuación relativamente más fuerte entre las vocales en comparación con las consonantes.

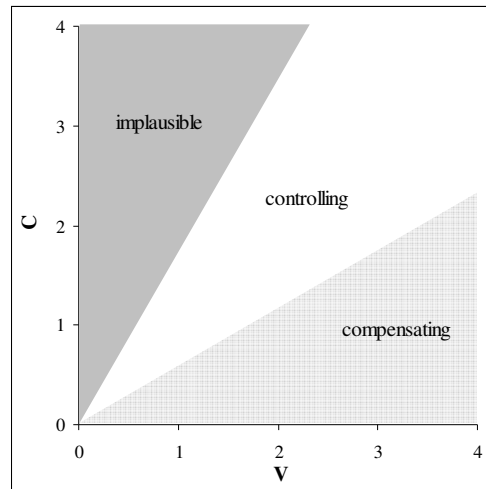


Figura 1. Representación esquemática de la colocación, en el plano cartesiano, de las lenguas «de control» y «de compensación».

Si todo esto es correcto, hay una consecuencia inmediata e importante: la concepción hasta ahora adoptada respecto al tema de la isocronía, incluso en los estudios recientes, presenta un elemento problemático. En el modelo de Ramus como en el de Grabe & Low, se analiza el comportamiento de los intervalos vocálicos y consonánticos, sin considerar los segmentos que los componen. Como consecuencia, no existe la posibilidad de distinguir entre algunas situaciones que podrían producirse en las lenguas naturales. Por ejemplo, si una lengua supuestamente «de control» tiene un nivel de complejidad silábica bastante grande (por supuesto, dentro de los límites permitidos para una lengua de este tipo), los intervalos consonánticos tendrían que presentar una marcada fluctuación en los valores de duración, como en el tipo «de compensación». Sin embargo, si estos valores se ponen en relación con el número de segmentos componentes, la diferencia frente al tipo «de compensación» tendría que afirmarse bastante claramente.

En un trabajo reciente (Bertinetto & Bertini 2008), los autores de esta contribución han propuesto un nuevo modelo, denominado «Control/Compensation Index» (CCI), basado sobre el siguiente algoritmo (inspirado en el modelo PVI de Grabe & Low), donde m significa ‘número de intervalos’ (vocálicos o consonánticos), d ‘duración’ (en segundos) y n ‘número de segmentos entre los intervalos pertinentes’:

$$CCI = \frac{100}{m-1} \sum_{k=1}^{m-1} \left| \frac{d_k}{n_k} - \frac{d_{k+1}}{n_{k+1}} \right|$$

La idea de considerar el número de segmentos constituyentes no es totalmente nueva: véase el trabajo de Rouas & Farinas (2004). Hay, sin embargo, diferencias importantes. En primer lugar, Rouas & Farinas tratan cada vocal como un núcleo de sílaba independiente, mientras que en el modelo CCI los intervalos vocálicos se dividen por el número de vocales constituyentes. Esto provoca una diferencia en el caso de las sinalefas, que componen intervalos vocálicos complejos; por lo demás, los *glides* se tratan como elementos consonánticos en ambos modelos. En segundo lugar, el índice de Rouas & Farinas es relativamente estático, en cuanto se basa en la duración media de los segmentos vocálicos (D_v) y de los segmentos consonánticos (D_c) en cada oración, posiblemente combinados con el número medio de segmentos consonánticos en cada sílaba (N_c). El modelo CCI, por el contrario, es dinámico, en cuanto disfruta de las ventajas del modelo PVI de Grabe & Low, que (desde nuestro punto de vista) se muestra superior al de Ramus en su capacidad de monitorizar las continuas fluctuaciones de la duración de los intervalos. La semejanza de nuestro modelo con el de Rouas & Farinas (desde ahora RF) se limita, por lo tanto, al hecho de que se introduce un índice de complejidad fonotáctica, expresado por el número de segmentos consonánticos en la sílaba (RF) o por el número medio continuamente variable de segmentos vocálicos y consonánticos en los respectivos intervalos (CCI). La complejidad intra-silábica es, en definitiva, el núcleo esencial del contraste entre lenguas «de control» y «de compensación».²

² Tenemos que precisar que, en la aplicación del modelo RF que se presentará en este trabajo, hemos elegido sólo una de sus versiones, aquella en la que se combinan, como coordenadas del plano cartesiano, D_c y D_v . El modelo RF se presenta también en otra versión, con D_c y N_c sobre los dos ejes. Sin embargo, esta última versión no permitiría una comparación fiel con

La Figura 2 muestra la colocación respectiva sobre el plano cartesiano, producida por una selección de algoritmos: el modelo CCI, el de Ramus (en dos versiones: ΔV y $\%V$)³, el de Grabe & Low (también en dos versiones: PVI y nPVI)⁴, el de Dellwo (otra vez en dos versiones: VarcoV y Varco $\%V$) y el modelo RF. Para comparar mejor los resultados, los valores vocálicos y consonánticos de ΔV , $\%V$, nPVI, Varco y RF se han dividido por 10. Es importante observar que sólo el modelo CCI – donde ambos tipos de intervalos, vocálicos y consonánticos, se dividen por el número de segmentos – permite una interpretación rigurosa de su colocación en relación a la bisectriz: para los otros modelos, la colocación en el plano es más sugestiva que precisa, porque no se puede saber exactamente cómo interpretar los valores numéricos (especialmente los que se refieren a las consonantes). Los datos en la figura se basan sobre las producciones semi-espontáneas de diez hablantes de la variedad de Pisa del italiano, accesibles en la base de datos AVIP/API:⁵ las líneas ovales de la figura (basadas en la desviación estándar relativa al porcentaje calculado separadamente para vocales y consonantes⁶) indican la dispersión espacial de los diez hablantes en cada computación. Los criterios de constitución del corpus están descritos en Bertini & Bertinetto (*en prensa*). Es importante notar que la parte final de cada oración – desde la última vocal acentuada – ha sido cortada, porque su comportamiento rítmico es marcadamente diferente. En total, los datos se basan sobre 2991 intervalos vocálicos y 2995 intervalos consonánticos: mucho más que en los demás trabajos de modelización rítmica y bastante para conseguir una considerable estabilidad estadística.

nuestro modelo. Además, las figuras en Rouas & Farinas (2004) muestran que la diferencia entre dichas versiones es mínima.

³ El índice ΔV calcula la desviación estándar de los intervalos vocálicos, mientras que el índice $\%V$ indica el porcentaje total de intervalos vocálicos en comparación con los intervalos consonánticos.

⁴ El índice PVI es como el índice CCI, con la diferencia de que en PVI no se tiene en cuenta el número de segmentos fonemáticos. En el índice nPVI hay un recurso matemático para neutralizar la diferencia de velocidad en el interior de la oración.

⁵ <http://www.cirass.unina.it/ricerca/studi%20parlato/raccolta%20corpora/api/api.htm>

⁶ La desviación estándar relativa al porcentaje se calcula de la manera siguiente: $100 \cdot \frac{\sigma}{x_m}$ (donde σ indica la desviación estándar y x_m el valor medio).

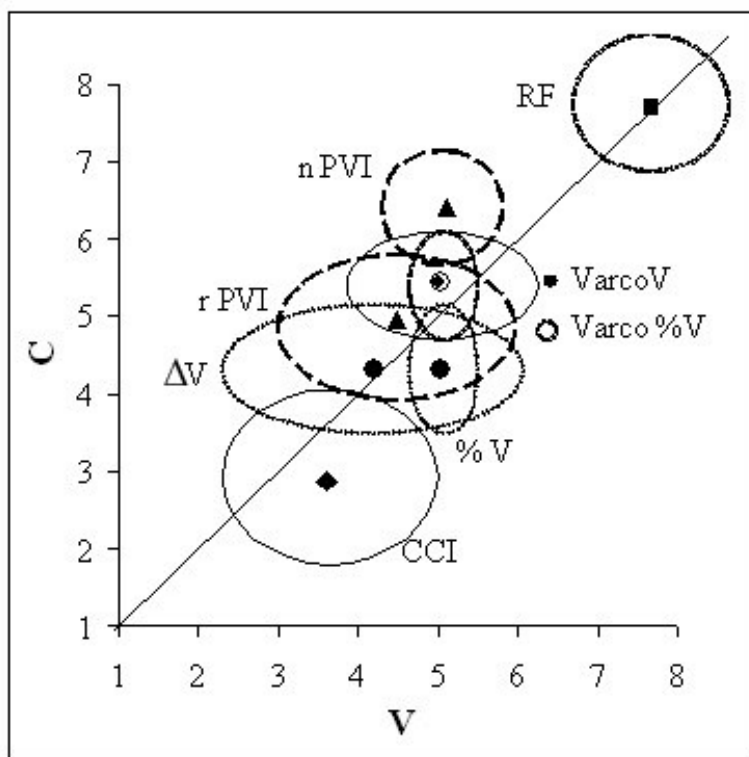


Figura 2. Colocaciones respectivas de las producciones lingüísticas de 10 hablantes pisanos producidas para los seis modelos considerados. Los valores de ΔV , $\%V$, nPVI, Varco y RF se han dividido por 10.

Por supuesto, la validez de un método concebido para medir las tendencias rítmicas de las lenguas naturales depende crucialmente de su capacidad para agrupar correctamente los idiomas prototípicamente pertenecientes a uno u otro grupo rítmico fundamental. Por el momento, no podemos producir estos resultados: tenemos la intención de extender nuestro análisis a datos de otras lenguas, comparando no sólo hablantes diferentes, sino también estilos de habla diferentes. Podemos sin embargo observar que el modelo CCI produce, en general, resultados perfectamente coherentes e interpretables, a la luz de lo que sabemos sobre las tendencias prosódicas del italiano. Véase, en relación a eso, el citado trabajo de Bertini & Bertinetto (*en prensa*).

Un resultado muy relevante es la fuerte influencia del factor «velocidad de elocución» sobre el comportamiento rítmico. La importancia de este elemento ha sido subrayada, en particular, por

los trabajos de Dellwo (por ej., 2004). Las Figuras 3a-b muestran la disposición de los datos producidos para los algoritmos antes considerados. Las tres clases de velocidad (medida en sílabas / segundo) dividen las producciones de los hablantes según el criterio siguiente: (I) $4.8 < 7.0$ (promedio: 6,2) (II) $\geq 7,0 \leq 8,1$ (promedio: 7,6) (III) $> 8,1$ (promedio: 9,2). En este caso, las elipses se limitan a agrupar los datos pertenecientes a cada modelo, para facilitar la comprensión; se debe notar que en la Figura 3b las clases I de Varco%V y II de VarcoV se superponen totalmente (además, I y II de Varco%V son muy vecinas):

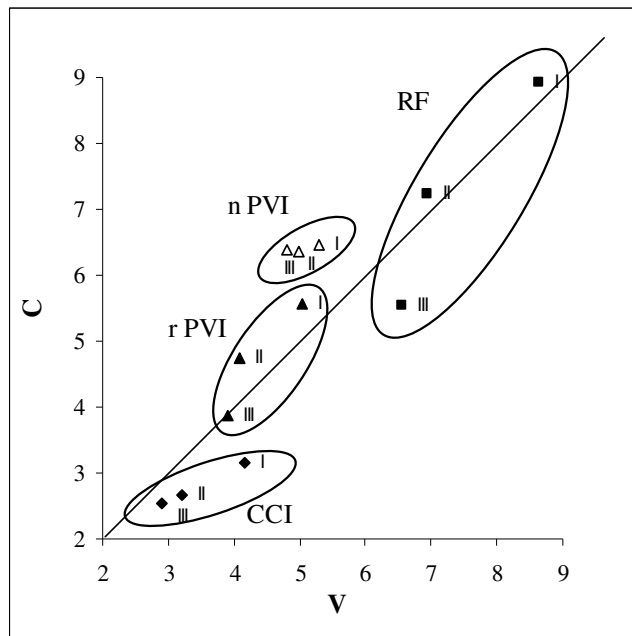


Figura 3a. Comportamiento de cuatro diversos modelos rítmicos en relación a la velocidad de elocución (clase I = lento; clase III = veloz). Algoritmos CCI, RF, PVI y nPVI.

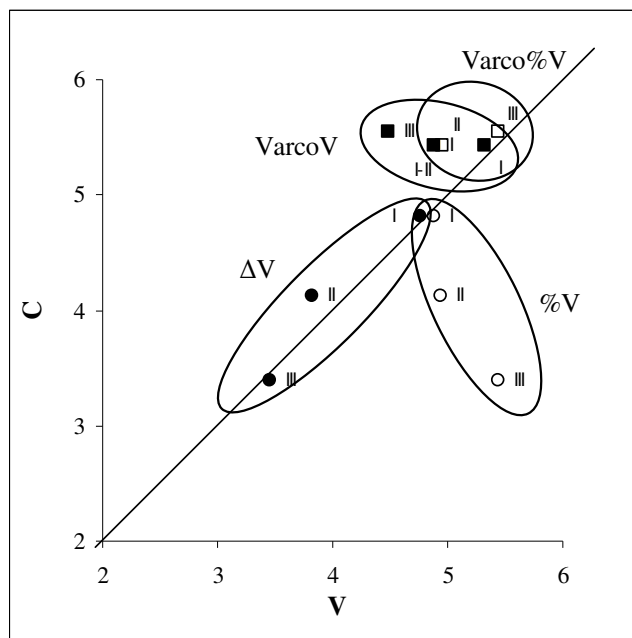


Figura 3b. Comportamiento de otros tres modelos rítmicos en relación a la velocidad de elocución (clase I = lento; clase III = veloz). Algoritmos VarcoV, Varco%V, ΔV y %V.

Un elemento importante para apreciar la eficacia del modelo CCI se manifiesta en el hecho de que la proximidad de las clases de velocidad II y III, afirmada icónicamente por CCI y nPVI y confirmada por la tendencial convergencia señalada por las computaciones estadísticas, parece más plausible que la indicación contraria sugerida por otros modelos, consistente en una proximidad relativamente mayor entre las clases I y II.⁷ En cuanto a

⁷ Según el cálculo del *t*-test (véase la Tabla 1), en CCI la clase I se distingue significativamente de la II, mientras que las II y III no están separadas una de la otra. Aplicando el test de Mann-Whitney, la indistinción entre II y III se mantendría sólo para las consonantes, y esto se puede explicar perfectamente en relación al comportamiento ideal de una lengua «de control». En estas lenguas, el aumento de velocidad de habla tendría que mostrar un efecto tendencialmente igual para vocales y consonantes; en los niveles superiores de velocidad, sin embargo, las consonantes tendrían que encontrar el límite de incompresibilidad más pronto que las vocales.

Con los otros modelos empleados, la situación se presenta de manera diferente y menos fácilmente inteligible. Aquí se consideran sólo los contrastes entre clases limítrofes, no los contrastes entre I y III, que son mucho menos iluminadores (la Tabla 1 los indica sólo para completar la información). Los modelos ΔV , rPVI y RF muestran un contraste

la diferente orientación espacial de %V, hay que apuntar que proviene de su concepción misma, y no sería (en sí misma) un factor de penalización de este modelo.⁸

2 La contribución de los componentes silábicos

Es importante notar que el modelo CCI, igual que los de Ramus, Grabe & Low y Dellwo, no se basa en la noción de sílaba: los intervalos consonánticos pueden a menudo incluir una coda y un onset silábico. La influencia de la sílaba es sólo en apariencia más fuerte en el caso de RF: los autores hablan de «pseudo-syllables» en relación a las secuencias de consonantes, pero estas secuencias no se separan en onset y coda: se trata, una vez más, de intervalos consonánticos no silabeados. A pesar de que la falta de atención a la sílaba esté en conflicto con la idea tradicional del ritmo (véase la denominación de «syllable-timing»), esto no constituye un problema para nuestra concepción (la hipótesis CC), en la cual el centro de gravedad se coloca más bien en el dominio de la fonotáctica que en el de la sílaba. Sin embargo, es interesante verificar el efecto de los componentes silábicos en relación al modelo CCI. El objetivo de esta contribución es, precisamente, el de aclarar este punto.

Con esta finalidad, hemos calculado dos índices consonánticos: CCI/*on* par las consonantes de onset, y CCI/*co* para las consonantes de coda. El conjunto de estos dos cálculos se designa CCI/*sil*. Como bien se comprende, el hecho de considerar

significativo completo (vocales y consonantes) entre las clases I y II, y sólo para las consonantes entre II y III; en VarcoV y nPVI, no hay ningún contraste significativo (esto se comprende bien para modelos como estos, concebidos para neutralizar las diferencias de velocidad, que sin embargo tienen un papel muy importante en el desarrollo del ritmo); por fin, en Varco%V sólo hay contraste para las vocales entre II y III. En Bertinetto & Bertini (2008) y Bertinetto & Bertini (*en preparación*) hemos analizado puntualmente todas estas diferencias; en este trabajo se ha elegido comparar CCI sólo con RF.

⁸ En cierta manera, %V puede considerarse como la imagen especular de los otros modelos. Si el aumento de velocidad se realiza comparativamente más sobre las consonantes que sobre las vocales, el porcentaje de estas últimas por supuesto aumenta (véase la nota 3).

separadamente los onsets y las codas introduce una cierta discontinuidad en el interior de cada tipo de intervalo, particularmente en el caso de las codas. La fonotáctica del italiano presenta un onset en casi cada sílaba, mientras que las codas tienen una distribución más restringida. Con muy pocas excepciones, las codas del italiano pueden estar constituidas por las sonantes /l r N/ (donde N indica el archifonema nasal) y por el primer elemento de las geminadas (incluso, en este caso, los segmentos obstruyentes). Además, hemos considerado como posible coda el fonema /s/ delante de otra consonante, a pesar de que su estatuto prosódico sea problemático (Bertinetto 1999). La diferencia entre onset y coda en italiano es entonces considerable: los onsets se suceden bastante regularmente, no obstante la posible presencia de codas. Estas, por el contrario, en nuestro corpus se encuentran en el 42% de las sílabas (frente al 58% de sílabas abiertas) y, por consiguiente, constituyen una secuencia relativamente discontinua. Este hecho no sólo plantea problemas en la aplicación del algoritmo CCI (que presupone continuidad local de las secuencias pertinentes), sino también disminuye su potencia estadística. El primer problema es inevitable; para el segundo propondremos una solución más adelante.

Con estas precauciones, podemos verificar el resultado de las computaciones producidas por el algoritmo CCI/*sil*, considerando separadamente los dos componentes consonánticos de la sílaba. La Figura 4 muestra la colocación de las tres clases de velocidad antes definidas. A pesar de que la orientación sea la misma en los tres casos, hay diferencias. Si fuera posible imaginar dos lenguas irreales, el «italiano/*on*» y el «italiano/*co*», el primero parecería más conforme al ideal de las lenguas «de control» (colocándose a lo largo de la bisectriz), mientras que el segundo se desplazaría un poco más hacia el territorio de las lenguas «de compensación», con su fluctuación más fuerte en la duración de las vocales que en la de las consonantes. Pero, por supuesto, esto no es nada más que un artefacto experimental. Lo que podemos aprender concretamente de las diferentes colocaciones de los datos en la Figura 4 es más bien el hecho de que los onsets, debido a su mayor riqueza fonotáctica, tienen una variabilidad, es decir una fluctuación momentánea de la duración, más grande que las codas. En apariencia, eso podría depender del número de elementos componentes: en italiano, el onset puede incluir hasta tres elementos, mientras que la coda no puede incluir más de un elemento (con muy raras excepciones no autóctonas). En realidad, esta no puede ser la explicación del

problema, porque en el modelo CCI, como en sus derivados «silábicos» CCI/*on* y CCI/*co*, la duración de cada intervalo consonántico se divide por el número de elementos constituyentes. La causa de la diferencia, al menos en parte, radica en la diversa naturaleza fonética de las consonantes, a pesar de que las codas incluyan casi cada tipo de segmentos en razón de la presencia de los primeros miembros de las geminadas. La proporción relativa de sonantes es en efecto más alta en las codas que en los onsets; además, hay una presencia más abundante de *glides* (segmentos de muy breve duración) en los onsets que en las codas.⁹ De toda manera, la menor variabilidad/fluctuación de la duración de las consonantes de coda parece constituir un hecho aclarado, probablemente como consecuencia de su duración media más breve, como se puede notar en la Tabla 2 del Apéndice. Si, por otra parte, consideramos que las consonantes de coda muestran paradójicamente – como se verá más tarde – una mayor sensibilidad estadística a los contrastes de velocidad, a pesar de su menor duración, se puede concluir que el verdadero motor del ritmo es la dinámica de las variaciones locales, más bien que los valores medios de duración. Volveremos sobre esto en las conclusiones.

⁹ Tenemos que observar aquí que, en la medición de los dos miembros de las geminadas, hay un inevitable elemento de aproximación. Como no hay una medida clara para dividir las geminadas en sus dos componentes fonemáticos, el criterio adoptado en nuestra investigación ha consistido en dividir por dos la duración total de la geminada. Esto, por supuesto, introduce en los datos de CCI/*sil* una cierta artificialidad: una artificialidad que no existe en el caso de CCI, donde no se distingue entre coda y onset (la fórmula se basa únicamente en el número de segmentos contenidos en los intervalos). De toda manera, no es evidente el motivo por el cual dicha artificialidad tendría que afectar de manera diferente a los onsets y a las codas. La diferencia de variabilidad señalada antes entre onsets y codas es, por consiguiente, independiente del criterio artificial adoptado para la división de las geminadas.

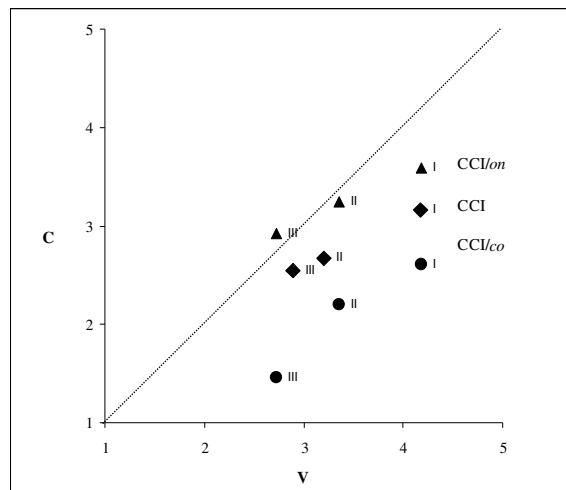


Figura 4. Colocación en el plano cartesiano del las tres clases de velocidad con respecto al algoritmo general (CCI) y a los algoritmos que se refieren a los componentes silábicos (CCI/on, CCI/co).

En efecto, la discriminación estadística entre las clases de velocidad se presenta de manera diferente en el modelo CCI de base y en los modelos CCI/sil. En CCI, hay indistinción entre las clases II y III; en CCI/on, por el contrario, las tres clases se distinguen (a pesar de que, entre II y III, la discriminación sobre el eje consonántico sea estadísticamente marginal: .052; véase la nota 7 y la Tabla 1 del Apéndice). En el caso de CCI/co, sin embargo, no es aconsejable calcular directamente el nivel de significatividad, por razones de insuficiente potencia estadística en la distribución de las codas al nivel de la oración (es decir: considerando cada oración en sí misma). Para superar el problema se ha intentado la solución siguiente. Para cada hablante se han acumulado todas las oraciones - o, según el caso, las oraciones incluidas en la misma clase de velocidad - constituyendo así una especie de macro-oración. Como resultado el número de codas aumenta marcadamente en el interior de la porción de habla considerada.

Como paso intermedio en este camino, se ha calculado el efecto de la acumulación en el modelo CCI básico, con la intención de comparar preliminarmente este resultado con datos ya analizados. Llamamos CCI-cum el modelo «cumulativo» general y CCI-cum/sil los modelos análogos que se refieren a los componentes silábicos. Para la comparación se ha elegido el modelo RF: el cual, como se describió en §1, es intrínsecamente estático y, por consiguiente, no se ve afectado por el problema de la

discontinuidad entre los intervalos pertinentes. Además, este modelo muestra una considerable sensibilidad a las variaciones de velocidad (véase la Tabla 1). La Figura 5 muestra que el resultado de los modelos CCI y CCI-cum es muy semejante: la única diferencia consiste en el hecho de que la dispersión de los diez hablantes (indicada por las elipses) es más grande con CCI-cum, como consecuencia del menor número de observaciones.¹⁰ El modelo RF-cum, por el contrario, produce resultados bastante diferentes del correspondiente modelo de base, a pesar de su naturaleza estática; la cual, en principio, tendría que constituir un ventaja en caso de acumulación, en cuanto los modelos estáticos no se ven afectados por la sucesión local de los intervalos. Esto indica que el modelo CCI tiene, quizás sorprendentemente, una mayor estabilidad.

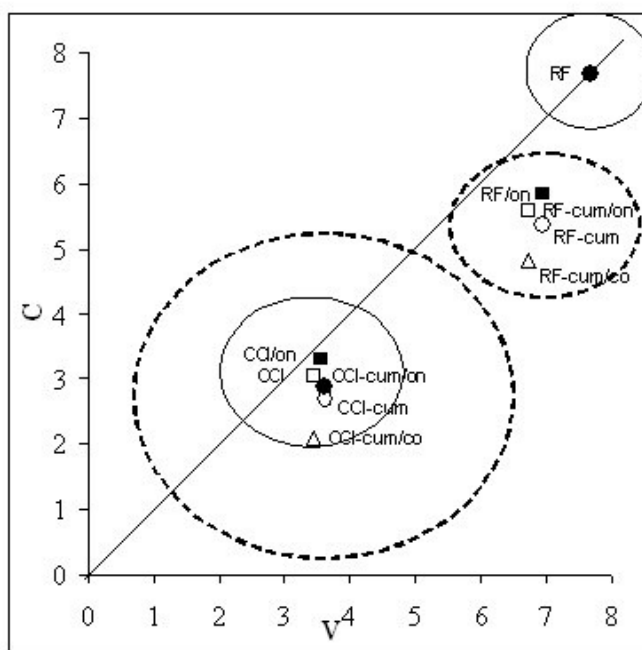


Figura 5. Colocación de los modelos generales (CCI, CCI/on, RF, RF/on,; marcas en negro, líneas continuas) y de los modelos cumulativos (CCI-cum, CCI-cum/on, CCI-cum/co, RF-cum, RF-cum/on, RF-cum/co; marcas en blanco, líneas discontinuas). Los valores de RF, que indican duraciones medias, están divididos por 10, para poder representarlos en un mismo gráfico.

¹⁰ En teoría tendría que haber 10 observaciones en cada clase de velocidad; en realidad hay menos, porque no todos los hablantes están representados en cada clase.

Las Figuras 6a-b muestran el efecto de los modelos CCI-*cum* (6a) y RF-*cum* (6b) sobre las tres clases de velocidad de habla. Aquí se puede ver cómo, en general, la acumulación reduce la variación sobre todo en el eje consonántico, mientras que para las vocales el efecto es menor. Desde el punto de vista estadístico, hay diferencias significativas sólo entre las clases I y II: en CCI-*cum*, la única diferencia se halla en las vocales, mientras que en RF-*cum* hay significatividad para las consonantes también. Por lo tanto, en ambos modelos se reduce la significatividad con respecto a los modelos no-cumulativos de base. Por consiguiente, tenemos que utilizar con cuidado los datos cumulativos, ya que introducen variaciones no inocuas (véanse los datos en la Tabla 1).

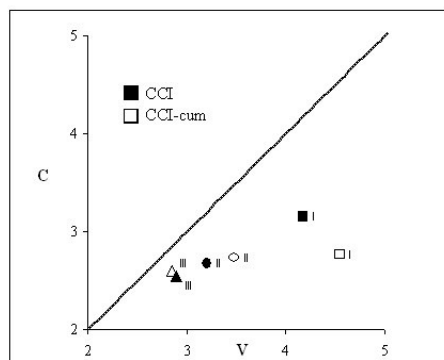


Figura 6a. Efecto de la velocidad de habla según los modelos general CCI y cumulativo CCI-*cum*.

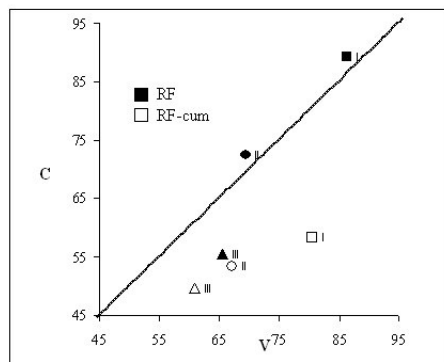


Figura 6b. Efecto de la velocidad de habla según los modelos general RF y cumulativo RF-*cum*.

Con esta precaución en la mente, podemos interpretar mejor el resultado generado por la aplicación del modelo *CCI-cum/sil*, que nos permite conseguir una adecuada potencia estadística para las codas. Las Figuras 7a-b muestran la colocación respectiva de los datos calculados con *CCI-cum/on* y *CCI-cum/co*; por comparación, véanse los datos de *CCI/sil* en la Figura 4. Se nota sobre todo la casi ausencia de variación en el eje consonántico. Desde el punto de vista estadístico, y considerando sólo las consonantes,¹¹ en *CCI-cum/on* las clases de velocidad no son distintas, mientras que en *CCI-cum/co* las consonantes se distinguen mutuamente sólo entre las clases II y III. Esto indica que en italiano hay una diferencia prosódica importante entre consonantes de onset y de coda: las primeras hallan más temprano su límite de incompresibilidad, mientras que las segundas, a pesar de su menor duración media, muestran fluctuaciones relativas locales más grandes precisamente entre los niveles de velocidad más altos.¹² La menor variabilidad de las consonantes de onset en función de la velocidad se nota, en efecto, también en el modelo no-cumulativo *CCI-on*, donde la significatividad estadística es constantemente menor para las consonantes que para las vocales (Tabla 1).

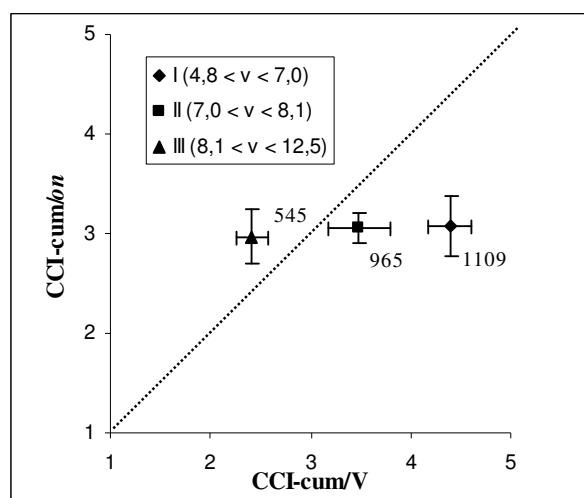


Figura 7a. Efecto de la velocidad de habla según el algoritmo *CCI-cum/on*.

¹¹ Para las vocales, los resultados estadísticos son por supuesto los mismos descritos ya en el comentario de la Figura 6.

¹² El adjetivo *local* tiene que interpretarse aquí con prudencia, porque en italiano no hay codas en cada sílaba.

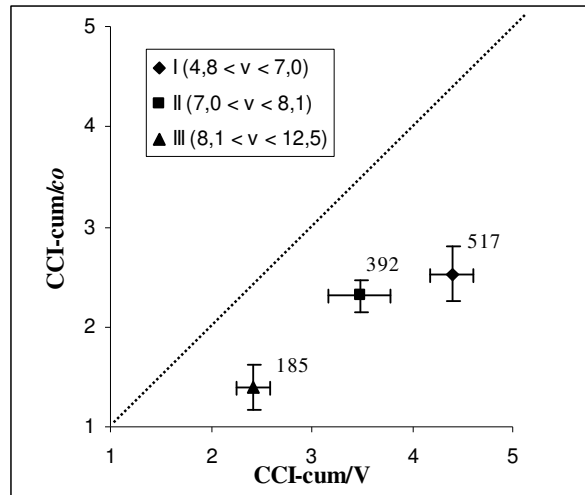


Figura 7b. Efecto de la velocidad de habla según el algoritmo CCI-cum/co. (Se indica para cada clase de velocidad el número de intervalos medidos.)

Para completar la comparación con el modelo RF-cum, véanse ahora las Figuras 8a-b. Aquí se nota una sensibilidad mucho más grande en relación a los componentes silábicos. Hay significatividad constante para vocales y consonantes en todas las comparaciones estadísticas, con la excepción parcial de las consonantes entre las clases II y III en RF-cum/on, donde el valor es sólo marginalmente significativo. Esta fuerte sensibilidad a las variaciones estadísticas entre las clases de velocidad se nota también en el modelo «silábico» no-cumulativo RF/on, donde (como se ve en la Tabla 1), los contrastes entre clases de velocidad son constantemente significativos para vocales y consonantes.

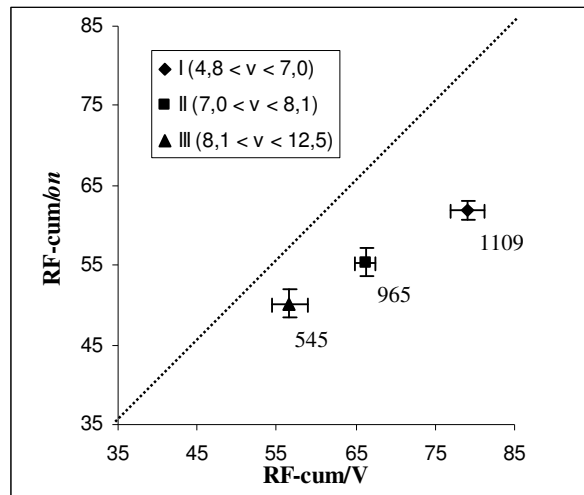


Figura 8a. Efecto de la velocidad de habla según el algoritmo RF-cum/on. En la abscisa los valores pertenecientes a las vocales.

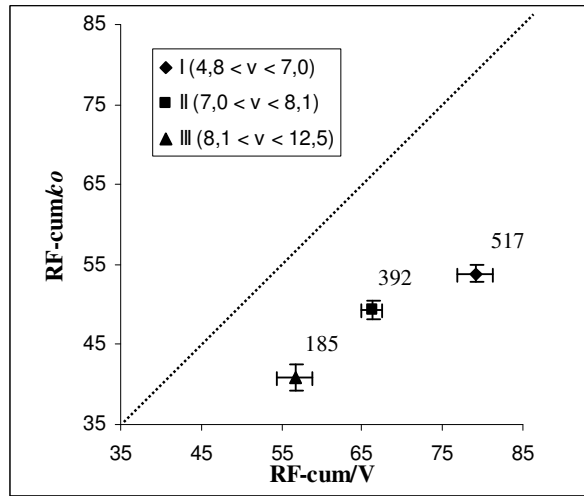


Figura 8b. Efecto de la velocidad de habla según el algoritmo RF-cum/co. En la abscisa los valores pertenecientes a las vocales. (Se indica para cada clase de velocidad el número de intervalos medidos.)

3 Discusión

En §2 se han presentado argumentos para evaluar el papel de los componentes silábicos de onset y coda en el desarrollo del ritmo. Para apreciar mejor los resultados se ha comparado el modelo CCI, dinámico, con el modelo RF, intrínsecamente estático. Además, para observar mejor el comportamiento de las codas (debido a su relativa discontinuidad en la cadena fónica del italiano), se ha realizado una versión especial de los dos algoritmos, consistente en la acumulación de las oraciones producidas para cada hablante, formando así macro-oraciones. Este recurso, realizado para obtener secuencias más largas, aumenta el número de elementos considerados en cada secuencia (consonantes de onset o de coda), permitiendo un mejor control estadístico; pero, por otra parte, esto reduce el número de observaciones y con ello la potencia estadística general (sólo 10 hablantes, cuando hay datos de todos, en el caso de las diferentes clases de velocidad). Por supuesto, las indicaciones obtenidas a través de los modelos cumulativos tienen que ser manejadas con prudencia. En resumen:

El modelo RF se ha mostrado:

- muy sensible - quizás demasiado - al efecto de la acumulación y de las variaciones de velocidad, pero no particularmente sensible a estas últimas en su variante cumulativa (Figuras 5 y 6b, y Tabla 1);

- muy sensible en cambio, en sus variantes «silábicas», a las variaciones de velocidad, así que las diferencias entre las clases se imponen con bastante evidencia, sobre todo para las codas (Figuras 8a-b y Tabla 1).

Estos hechos pueden interpretarse como una señal de relativa inestabilidad. Por un lado, la acumulación disminuye las diferencias entre clases de velocidad; por otro, las variantes «silábicas» se muestran muy sensibles a esas diferencias, incluso en los modelos cumulativos (RF-cum/on y RF-cum/co). Esto podría alimentar dudas en cuanto a la eficacia de este algoritmo estático de modelización del ritmo.

El modelo CCI, por su parte, se muestra:

- mucho menos sensible al efecto de la acumulación, sobre todo en cuanto a la media general (Figura 5; la dispersión de los sujetos, en cambio, se enfatiza) y, en cierta manera, también en relación a las clases de velocidad (Figura 6a), lo que indica una considerable estabilidad; y eso a pesar de su carácter dinámico, que

en principio podría constituir un problema como consecuencia de la discontinuidad de los intervalos pertinentes (especialmente las codas);

- sensible en manera bastante diferenciada, en sus variantes «silábicas», a las variaciones observables en las consonantes de onset y de coda (Figuras 7a-b).

Estos resultados sugieren algunas reflexiones. Ante todo, se puede deducir que el modelo CCI permite, por su mayor estabilidad y fiabilidad, una mejor interpretación del papel desarrollado por los componentes silábicos en la dinámica del ritmo. La discusión siguiente, por lo tanto, se basará exclusivamente en las indicaciones de este modelo.

Si consideramos las duraciones medias indicadas en la Tabla 3 del Apéndice, podremos notar que los valores medios de cada componente silábico, entre las diferentes clases de velocidad, están casi exactamente distanciados. Entonces, el diferente valor de significatividad de los contrastes estadísticos no depende de una supuesta variación de distancia relativa entre las clases, sino de la dinámica interna al ritmo en su continua fluctuación local, subrayada (sobre todo para las vocales) por la alternancia entre sílabas acentuadas y no acentuadas. Esto es un factor que los modelos estáticos no pueden iluminar. Para cualquiera de estos modelos, la efectiva sucesión de los acentos en la cadena hablada y las consiguientes perturbaciones locales de las duraciones segmentales, no tienen ninguna relevancia. Lo que podemos aprender de los resultados generados para el modelo CCI (en su versión básica, no-cumulativa) es, por el contrario, que el aumento de la velocidad de habla disminuye drásticamente las fluctuaciones locales, así que no hay diferencia estadística entre las clases II y III.

En cuanto a los componentes silábicos, nuestros análisis demuestran dos hechos interesantes a nivel consonántico. En primer lugar, el modelo CCI/*sil* muestra una menor fluctuación local de la duración en los onsets que en las codas, a pesar de la menor duración media de las codas (Figura 4, Tabla 2). En segundo lugar, si consideramos específicamente los modelos CCI-*cum/sil*, hay contraste significativo sólo para las codas y sólo entre las clases II y III (Tabla 1). Esto indica que, al aumentar la velocidad (y al disminuir consiguientemente la duración), las consonantes de coda, a pesar de su duración media más breve, manifiestan una potencialidad de fluctuación local más grande que las consonantes de onset, las cuales alcanzan más rápido su límite de incompresibilidad (Figura 7). Esto nos revela una característica

importante del ritmo en italiano que sólo se puede notar a nivel intra-silábico.

Tenemos ahora que preguntarnos sobre el verdadero papel de la sílaba en el desarrollo del ritmo en italiano y en las lenguas naturales en general. En nuestra opinión, la respuesta puede articularse en dos pasos.

En un primer momento, notamos que, según las observaciones precedentes, la sílaba cuenta bastante, porque las tendencias rítmicas no son las mismas en los diferentes componentes. Con alta velocidad de habla, las consonantes de coda parecen tener una mayor flexibilidad en comparación con las de onset, a pesar de su menor variedad fonética (a causa de la abundante presencia de sonorantes) y de su menor duración. Aquí se podría formular una objeción: el modelo CCI se basa, crucialmente, sobre las variaciones locales, pero las codas pueden tener una presencia muy escasa en muchas oraciones, por efecto de las restricciones fonotácticas del italiano. Esto es verdad. Sin embargo, la comparación con los modelos RF-cum/sil nos ofrece una interesante confirmación indirecta: el contraste entre las clases II y III es muy significativo en RF-cum/co y sólo marginalmente significativo en RF-cum/on (Tabla 1). Esto indica, por supuesto, que la diferencia de comportamiento entre onset y coda no es un artefacto de los modelos CCI/sil, porque se observa también en los modelos RF-cum/sil.

En un segundo paso, notamos que el papel desarrollado por los componentes silábicos es un efecto secundario de la estructura fonotáctica del italiano, con sus restricciones de recurrencia fonemática. La sílaba, en sí misma, no es una unidad primaria de la fonología, sino el resultado de las presiones fonotácticas básicas: un epifenómeno, como han propuesto algunos autores (Dressler & Dziubalska-Kořaczyk 1994; Dziubalska-Kořaczyk 2002; Ohala & Kawasaki 1997; Vennemann 1988; 1994). De eso hay confirmaciones experimentales (por ej., Bertinetto et al. 2007). En otras palabras: no es la sílaba lo que modela la articulación del ritmo, sino la sílaba misma es articulada por la secuencia regulada de los segmentos fonemáticos, según la fonotáctica de cada lengua. La sílaba es entonces generada por la misma fuerza que produce el ritmo. Las regularidades observadas en italiano, en el comportamiento de onset y codas, no están en el origen del ritmo, sino que constituyen un elemento característico y esencial del ritmo en esta lengua; es decir, uno de sus componentes, sea cual sea el auténtico origen del ritmo en las lenguas naturales.

4 Apéndice

| | | v1 vs. v2 | v2 vs. v3 | v1 vs. v3 |
|------------|------------|------------------|------------------|------------------|
| Ramus | ΔV | 0.016 | n.s. | 0.004 |
| | ΔC | 0.000 | 0.000 | 0.000 |
| | %V | n.s. | 0.001 | 0.000 |
| PVI | nPVI-V | n.s. | n.s. | 0.018 |
| | nPVI-C | n.s. | n.s. | n.s. |
| | rPVI-V | 0.002 | n.s. | 0.002 |
| | rPVI-C | 0.000 | 0.000 | 0.000 |
| VARCO | V | n.s. | n.s. | 0.009 |
| | C | n.s. | n.s. | n.s. |
| | %V | cf. Ramus %V | cf. Ramus %V | cf. Ramus %V |
| RF | V | 0.000 | n.s. | 0.000 |
| | C | 0.000 | 0.000 | 0.000 |
| RF-cum | V | 0.000 | n.s. | 0.000 |
| | C | 0.021 | n.s. | 0.004 |
| RF/on | V | 0.000 | 0.002 | 0.000 |
| | C/on | 0.000 | 0.000 | 0.000 |
| RF/co | V | - | - | - |
| | C/co | - | - | - |
| RF-cum/on | V | 0.000 | 0.002 | 0.000 |
| | C/on | 0.009 | n.s. (0.055) | 0.000 |
| RF-cum/co | V | cf. RF-cum/on V | cf. RF-cum/on V | cf. RF-cum/on V |
| | C/co | 0.011 | 0.001 | 0.000 |
| CCI | V | 0.000 | n.s. | 0.000 |
| | C | 0.002 | n.s. | 0.000 |
| CCI-cum | V | 0.019 | n.s. | 0.001 |
| | C | n.s. | n.s. | n.s. |
| CCI/on | V | 0.000 | 0.014 | 0.000 |
| | C | 0.036 | n.s. (0.052) | 0.000 |
| CCI/co | V | - | - | - |
| | C | - | - | - |
| CCI-cum/on | V | 0.006 | 0.015 | 0.000 |
| | C | n.s. | n.s. | n.s. |
| CCI-cum/co | V | cf. CCI-cum/on V | cf. CCI-cum/on V | cf. CCI-cum/on V |
| | C | n.s. | 0.015 | 0.032 |

Tabla 1. Comparaciones estadísticas con el *t*-test entre las clases de velocidad.

| | C/on | V | C/co |
|----------------|------|------|------|
| MEDIA | 62,2 | 69,8 | 51,0 |
| DESV. EST. | 34,1 | 41,8 | 23,7 |
| N ^a | 2619 | 3076 | 1094 |

Tabla 2. Duración media de los componentes silábicos en el corpus utilizado.

| | Clase I | | | Clase II | | | Clase III | | |
|----------------|---------|------|------|----------|------|------|-----------|------|------|
| | C/on | V | C/co | C/on | V | C/co | C/on | V | C/co |
| MEDIA | 70,0 | 78,0 | 56,1 | 60,0 | 66,7 | 48,5 | 50,4 | 59,7 | 42,2 |
| DESV. EST. | 36,7 | 40,1 | 26,0 | 32,0 | 48,5 | 21,1 | 27,5 | 27,4 | 18,4 |
| N ^a | 1215 | 1274 | 520 | 1037 | 1127 | 396 | 569 | 675 | 158 |

Tabla 3. Duración media de los componentes silábicos en las tres clases de velocidad.

5 Referencias bibliográficas

- Barbosa, Plinio (2006): *Incursões em torno do ritmo da fala*. Campinas: Pontes.
- Bertinetto, Pier Marco (1981): *Strutture prosodiche dell'italiano. Accento, quantità, sillaba, giuntura, fondamenti metrici*. Firenze: Accademia della Crusca.
- Bertinetto, Pier Marco (1989): «Reflections on the dichotomy 'stress- vs. syllable timing'». *Revue de phonétique appliquée* 91/93, 99-130.
- Bertinetto, Pier Marco (1999): «La sillabazione dei nessi /sC/ in italiano. Un'eccezione alla tendenza 'universale'». In: Paola Benincà & Alberto Mioni & Laura Vanelli, eds., *Fonologia e morfologia dell'italiano e dei dialetti d'Italia. Atti del XXXI Congresso della Società di Linguistica Italiana, Padova, 25-27 settembre 1997*. Roma: Bulzoni, 71-96.
- Bertinetto, Pier Marco & Chiara Bertini (2008): «On modeling the rhythm of natural languages: the 'Control/Compensation Index'». In: *Proceedings of Speech Prosody 2008 (Campinas, 6-10 May 2008)*.
- Bertinetto, Pier Marco & Chiara Bertini (*en preparación*): «Towards a unified predictive model of natural language rhythm».
- Bertinetto, Pier Marco & Carol A. Fowler (1989): «On sensitivity to durational modifications in Italian and English». *Rivista di linguistica* 1, 69-94.
- Bertinetto, Pier Marco & Sylwia Scheuer & Katarzyna Dziubalska-Kołodziej & Maddalena Agonigi (2007): «Intersegmental cohesion and syllable division in Polish». In: Jürgen Trouvain & William J. Barry, eds., *Proceedings of the 16th International Congress of Phonetic Sciences*. Saarbrücken: Universität des Saarlandes, 1953-1956.
- Bertinetto, Pier Marco & Domokos Vékás (1991): «Controllo vs. compensazione: sui due tipi di isocronia». In: Emanuela Magno Caldognetto & Paola Benincà & Laura Vanelli, eds., *L'interfaccia tra fonologia e fonetica*. Padova: Unipress, 155-162.
- Bertini, Chiara & Pier Marco Bertinetto (*en prensa*): «Prospezioni sulla struttura ritmica dell'italiano basate sul corpus semispontaneo AVIP/API». In: *Comunicación presentada en el congreso AISV 2007, Cosenza (Diciembre 2007)*.
- Dauer, Rebecca M. (1983): «Stress-timing and syllable-timing reanalyzed». *Journal of phonetics* 11, 51-62.
- Dellwo, Volker (2004): «Rhythm and speech rate: A variation coefficient for ΔC ». In: P. Karnowski & I. Szigeti, eds., *Language and language-processing*. Frankfurt am Main: Lang, 231-241.
- Dressler, Wolfgang U. & Katarzyna Dziubalska-Kołodziej (1994): «Evidence of syllabic writing against the phonological syllable». In: Wolfgang U. Dressler & Martin Prinzhorn & John R. Rennison, eds., *Phonologica 1992. Proceedings of the 7th International Phonology Meeting*. Torino: Rosenberg & Sellier, 65-76.

- Dziubalska-Kořaczyk, Katarzyna (2002): *Beats-and-Binding Phonology*. Frankfurt am Main: Lang.
- Goldstein, Louis & Ioana Chitoran & Elisabeth Selkirk (2007): «Syllable structure as coupled oscillator modes: Evidence from Georgian vs. Tashlhiyt Berber». In: Jürgen Trouvain & William J. Barry, eds., *Proceedings of the XVIth International Congress of Phonetic Sciences: ICPHS XVI, 6-10 August, Saarbücken, Germany*. Saarbrücken: Universität des Saarlandes, 241-244.
- Grabe, Ester & Ee Ling Low (2002): «Durational variability in speech and the rhythm class hypothesis». In: *Papers in Laboratory Phonology 7*. Berlin: Mouton de Gruyter, 515-546.
- Low, Ee Ling & Ester Grabe & Francis Nolan (2000): «Quantitative characterizations of speech rhythm: syllable-timing in Singapore English». *Language and speech* 43, 377-401.
- Mairano, Paolo & Antonio Romano (2006): «Lingue isosillabiche e isoaccentuali: misurazioni strumentali su campioni di italiano, francese, inglese e tedesco». In: Veronica Giordani & Valentina Bruseghini & Piero Cosi, eds., *Atti del III Convegno Nazionale AISV Scienze vocali e del linguaggio*. Trento: EDK Editore, CD-rom.
- O'Dell, Michael L. & Tommi Nieminen (1999): «Coupled oscillator model of speech rhythm». In: John Ohala et al., eds., *Proceedings of the XIVth International Congress of Phonetic Sciences: ICPHS 99, San Francisco, 1-7 August, 1999*. Berkeley, CA: University of California, vol. 2, 1075-1078.
- Ohala, John J. & Haruko Kawasaki-Fukumori (1997): «Alternatives to the sonority hierarchy for explaining segmental sequential constraints». In: Stig Eliasson & Ernst H. Jahr, eds., *Language and its ecology: Essays in memory of Einar Haugen*. Berlin: Mouton de Gruyter, 343-365.
- Pike, Kenneth L. (1945) [1947]: *The intonation of American English*. Ann Arbor: University of Michigan Press.
- Ramus, Franck & Marina Nespors & Jacques Mehler (1999): «Correlates of linguistic rhythm in the speech signal». *Cognition* 73, 265-292.
- Rouas, Jean-Luc & Jérôme Farinas (2004): «Comparaison de méthodes de caractérisation du rythme des langues». In: *Actes du colloque MIDL 2004 «Identification des langues et des variétés dialectales par les humains et par les machines» (Paris, 29-30 nov. 2004)*. http://www.limsi.fr/MIDL/actes/session%20II/Rouas&Farinas_MIDL2004.pdf.
- Russo, Michela & William J. Barry (2003): «Isocronia soggettiva o oggettiva? Relazioni tra tempo articolatorio e quantificazione ritmica». In: Federico Albano Leoni & Francesco Cutugno & Massimo Pettorino & Renata Savy, eds., *Atti del convegno nazionale Il parlato italiano*. Napoli: M. D'Auria Editore, CD-rom.

- Russo, Michela & William J. Barry (2004): «Interaction between segmental structure and rhythm. A look at Italian Dialects and Regional Standard Italian». *Folia linguistica* 38, 277-296.
- Schmid, Stephan (2004): «Une approche phonétique de l'isochronie dans quelques dialectes italo-romans». In: Trudel Meisenburg & Maria Selig, eds., *Nouveaux départs en phonologie. Les conceptions sub- et suprasegmentales*. Tübingen: Narr, 109-124.
- Vennemann, Theo (1988): *Preference laws for syllable structure and the explanation of sound change: with special reference to German, Germanic, Italian, and Latin*. Berlin: Mouton de Gruyter.
- Vennemann, Theo (1994): «Universelle Nuklearphonologie mit epiphänomenaler Silbenstruktur». In: Karl H. Ramers & Heinz Vater & Henning Wode, eds., *Universale phonologische Strukturen und Prozesse*. Tübingen: Niemeyer, 7-54.

