EL RITMO Y LA ENTONACIÓN EN LA LECTURA DEL CASTELLANO

COMUNICACIÓN PRESENTADA AL IX ENCUENTRO DE LA SEPLN DE SANTIAGO DE COMPOSTELA (1993)

Josep Martí & Francesc Gudayol
Escola Superior d'Enginyeria Electrònica i Informàtica La Salle
Universitat Ramon Llull. Barcelona

I. INTRODUCCIÓN

Objetivo del presente estudio

Con este trabajo se pretende caracterizar los rasgos fundamentales de la prosodia del castellano por lo que hace referencia al ritmo y a la entonación en el marco de una lectura formal de un texto escrito. El estudio se realiza dentro del marco de la investigación de las reglas prosódicas que aporten una mayor naturalidad a los conversores texto-voz en castellano. En principio, hemos descartado la modalidad del lenguaje espontáneo ya que en los conversores texto-voz normalmente se pretende provocar un efecto de lectura oral de una máquina a partir de texto escrito. De momento no se pretende llegar a unos resultados definitivos, cosa que extrapola el marco de esta experimentación, sino definir y cuantificar los rasgos fundamentales de la evolución del tono y de la duración de los fragmentos de voz en función de su situación dentro de la frase.

El interés se centra en la medida de datos concretos fácilmente cuantificables para que puedan ser introducidos en forma de reglas objetibables de un conversor texto-voz. Para una síntesis de voz con un elevado grado de naturalidad es muy recomendable partir de un análisis morfológico y sintáctico del texto escrito. De momento no disponemos de un analizador automático a partir del texto. Intentaremos, por tanto, buscar reglas que puedan establecerse a partir de datos léxicos más sencillos y fácilmente detectables. Básicamente partiremos de una segmentación del texto en grupos fónicos comprendidos entre signos de puntuación. En caso de que estos segmentos sean excesivamente largos se subdividirán previamente de acuerdo con unos criterios sencillos como los sugeridos por los diseñadores del conversor AMIGO (RODRÍGUEZ, 1992). Los grupos fónicos a su vez se subdividen en grupos acentuales. Cada grupo acentual está constituido por una palabra tónica y las palabras átonas que la preceden.

Nuestro trabajo se basa en hipótesis previas suficientemente estudiadas en otros idiomas para completar su definición y cuantificación en el caso de la lengua castellana. Difícilmente podremos hallar reglas totalmente definidas y que se apliquen a la totalidad de los casos previstos. La prosodia de una lengua, aún tratándose de una lectura formal del mismo texto, se presta a numerosas variantes entre unos y otros locutores; con diferentes soluciones que un oyente normal considera igualmente correctas. Se trata por tanto de encontrar los modelos que presentan una frecuencia estadística suficiente para que no puedan ser considerados como poco naturales en el caso de una lectura.

Metodología

El estudio se ha hecho sobre una base de 8 informantes masculinos, de edades comprendidas entre 20 y 48 años, que han leído un mismo texto de estilo periodístico (ver tabla II) en una sala debidamente tratada con material absorbente. Cada informante se ha preparado para una lectura correcta y ha recibido la indicación de hacerla con la máxima naturalidad. La grabación se ha realizado sobre cinta DAT y con un micrófono SHURE-PROLOGUE 8L. El análisis de tiempos y curvas entonativas se ha realizado sobre una SPARK STATION IPX SUN y con el tratamiento WAVES+ de la casa Entropic.

El estudio rítmico se ha realizado a base de mediciones temporales de los 40 grupos acentuales del texto (ver anexo final) para cada uno de los 8 locutores, lo que representa un conjunto de 320 grupos acentuales. Se ha medido también la duración de todas las vocales tónicas, una por cada grupo acentual, que en total son también 320 vocales tónicas. Entre las vocales átonas se ha seleccionado un número suficiente: 31 por informante, es decir: 248 vocales átonas en total.

El estudio de la entonación se ha centrado en la medición de la frecuencia fundamental (F0) durante la emisión de cada una de las vocales tónicas. Para cada vocal se ha extraído por separado el valor medio y el valor incremental de F0 durante su elocución. Los valores de la frecuencia fundamental se han normalizado respecto al valor medio de todas las mediciones de cada locutor, en vistas a eliminar los efectos del tono de voz propio de cada informante.

La duración total del texto para cada lector y su velocidad de pronunciación en sílabas por segundo es la que se puede ver en la tabla I. La duración media es de 21,4 s. y la velocidad media 6,0 sílabas/s. La muestra resulta muy variada ya que va desde casos con un ritmo lento de lectura (5,1 sílabas/s) hasta casos de lectura que ya se podría calificar de rápida (7,4 sílabas/s). Los resultados finales se refieren a los valores medios, de forma que serían representativos de un caso de lectura a velocidad normal.

Tabla I. Duración total y velocidad de locución del texto estudiado.

Informante	Tiempo total (s)	Velocidad (sílabas/s)
L.P.D.	18,9	6,7
F.G.	24,7	5,1
J.M.	24,7	5,1
J.L.P.	23,7	5,3
J.R.	22,8	5,5
A.T.	22,1	5,7
J.T.	17,2	7,3
A.M.	17,1	7,4
Valor medio:	21,4	6,0

II. EL RITMO

Las pausas

En el texto estudiado (Cf. tabla II) se confirma la uniformidad de todos los informantes en marcar las pausas indicadas por signos ortográficos. Además cuando el fragmento es excesivamente largo hay una tendencia a añadir pausas en puntos muy específicos. Particularmente todos los informantes menos uno añaden una pausa antes de la frase copulativa "y en los que había...". Algunos informantes añaden una pausa antes de algunas frases explicativas como: "que daba vueltas" o "dándole vueltas a esa palanca". En el caso del último grupo fónico previo al interrogante se han observado pausas antes del verbo "cambió" o antes del grupo acentual "por las de ser". En algún caso estas pausas no se han llegado a materializar como espacios de silencio sino únicamente como alargamiento de la sílaba inmediatamente anterior. Parece ser como si todos los informantes estuvieran de acuerdo en considerar el período excesivamente largo, pero a la hora de decidir la posición de la pausa hubiera disparidad de criterios. Este resultado nos favorece a la hora de establecer reglas para las pausas, que siempre podrán tener una cierta flexibilidad, sin perder la naturalidad de la locución.

Tabla II. Distribución de las pausas

Grupo fónico	Duración media (ms)	Desviación típica (ms)	Nº de informantes	Nº de sílabas	Nº de grupos acentuales
Hubo una época,	293	155	8	6	2
yo debía tener	221	-	1	6	3
unos cinco años,	446	172	8	6	2
en que me encantaban los conductores de tranvías,	580	320	8	15	3
de aquellos tranvías de entonces,	326	202	8	9	3
que no tenían puertas	376	188	7	7	3
y en los que había una palanca de freno muy grande	240	201	4	17	5
que daba vueltas.	931	336	8	5	2
A mí siempre me emocionaba ver al conductor	270	132	2	14	5
dándole vueltas a esa palanca.	1191	309	8	10	4
Y un poco más tarde	268	145	3	6	2
cambió sus aspiraciones	278	202	2	8	2
por las de ser premio Nobel,	437	45	2	8	3

En la tabla II se presenta un estudio completo de todas las pausas observadas, con su valor medio, su desviación típica y el número exacto de informantes que las han aplicado en cada caso. Por tanto, la subdivisión de los grupos fónicos más extensos no se ha realizado por igual en todos los casos, pero si que se cumple una cierta regla empírica de no alargar los grupos fónicos más allá de unas 24 sílabas, ni más allá de unos 9 grupos acentuales. En

la tabla se indican las pausas que siguen a cada uno de los períodos indicados. Se da también el número de sílabas fonéticas y de grupos acentuales correspondientes a cada fragmento. Hay que tener en cuenta que aquí se mencionan todas las pausas obtenidas, pero cada locutor ha realizado menos pausas de las aquí recogidas, incluyendo en un mismo grupo fónico dos períodos consecutivos de los indicados en la tabla. Claramente se observa que la pausa correspondiente a un punto es mucho más extensa que las demás con una duración en torno a un segundo. Las pausas que no han sido marcadas por todos los informantes son las más cortas y normalmente no pasan de los 300 ms. Mención a parte merece la pausa situada inmediatamente antes del interrogante "¿no?" que no ha sido realizada más que por dos informantes, debido a la inclusión de la frase anterior i de la interrogación dentro del mismo grupo fónico, como si hubiera un rechazo a la creación de un grupo fónico monosilábico.

Duración de los grupos acentuales

Como ya es bien sabido, el ritmo de locución de una lengua puede ser de dos tipos: ritmo acentual o ritmo silábico (CRISTAL 88). En el primer caso se da una constancia de la duración de los grupos acentuales contabilizados a partir del tiempo transcurrido entre dos acentos consecutivos. En el segundo caso se da más bien una constancia de las sílabas.

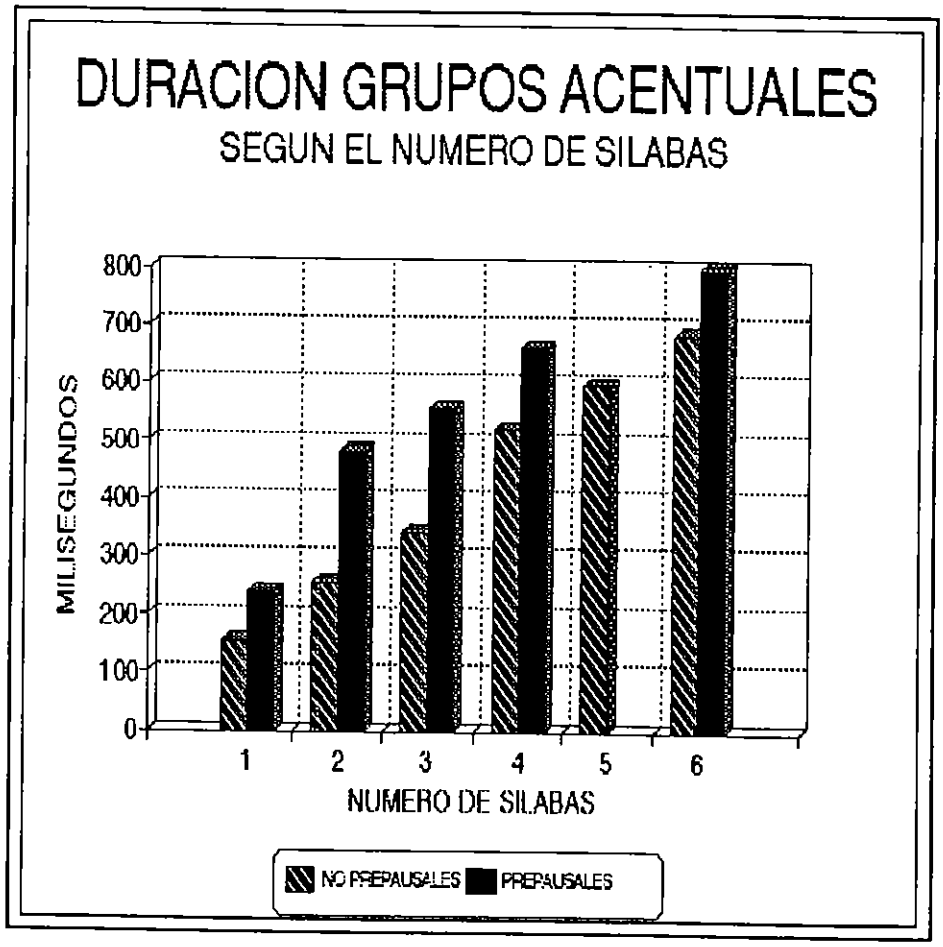


Fig. 1 Duración de los grupos acentuales en función del número de sílabas. (No se dispone de datos para el caso de grupos prepausales con 5 sílabas)

Tradicionalmente se ha considerado que el castellano tiene un ritmo preferentemente silábico aunque en algunos estudios del castellano sudamericano se ha llegado a conclusiones

ntradictorias con este presupuesto (BORZONE, 1983), (TOLEDO, 1988).

En nuestro caso, para la segmentación de los grupos acentuales, hemos partido de la ssificación que da Quilis (QUILIS, 1990) en las páginas 225-230 de las palabras tónicas átonas de la lengua castellana. La medición de cada grupo acentual se ha hecho a partir de da una de las palabras tónicas más las diferentes partículas átonas que le preceden. La uración media de cada grupo acentuado para los 8 informantes es la que puede verse en el nexo final. Claramente se constata que la duración es muy fluctuante según el número de llabas, lo cual descarta toda posibilidad de considerar el ritmo acentual para el castellano. a dependencia entre la duración de los grupos acentuales sigue una ley de carácter muy neal, según puede comprobarse en el gráfico de la figura 1. En el mismo gráfico puede bservarse claramente como la ley es distinta según se trate de grupos prepausales o no repausales. La diferencia entre los dos casos tiende a ser constante e independiente del úmero de sílabas. Si caracterizamos los dos casos por sus respectivas rectas de regresión se obtienen los siguientes resultados:

Grupos no prepaus.: $y = 42 + 109.x$ con un factor de regresión $R^2 = 0,987$
Grupos prepausales: $y = 212 + 105.x$ con un factor de regresión $R^2 = 0,926$

En donde "y" corresponde a los ms del grupo acentual y "x" al número de sílabas del grupo.

La separación entre ambas rectas es muy constante y se puede cuantificar a partir de las dos ordenadas en el origen, que difieren en 170 ms. Esta duración es la que hay que atribuir prioritariamente al incremento de la última sílaba por su carácter prepausal. Los respectivos factores de regresión son extremadamente altos y la pendiente en ambos casos es la misma, lo cual permite deducir una ley fácilmente cuantificable respecto a la diferenciación entre grupos prepausales y no prepausales. Este incremento afecta básicamente a la última sílaba del grupo acentual y tanto a la vocal como a las consonantes.

La duración media de las sílabas no prepausales se puede deducir del anexo final, dividiendo cada grupo no prepausal por su número de sílabas. El resultado obtenido por este método es de 124 ms por sílaba con una desviación típica de 24 ms. Este valor se ha obtenido a partir de 28 grupos acentuales y de cada uno de los 8 informantes. El valor medio de las sílabas prepausales obtenido sobre 12 grupos acentuales y los 8 informantes es de 238 ms con una desviación típica de 66 ms. Por tanto el incremento representa un factor de 1,92 (casi el doble) que habrá que distribuir entre la vocal y las consonantes de la última sílaba prepausal. La diferencia entre sílabas prepausales y no prepausales es de 114 ms, algo inferior a los 170 ms de diferencia entre los respectivos grupos acentuales prepausales y no prepausales. La discrepancia probablemente habrá que atribuirla a la penúltima sílaba antes de la pausa que sufriría también un ligero aumento en su duración.

Recordemos que no todos los informantes coinciden en la posición de las pausas en el texto. En los lugares indicados por coma o punto todos los locutores coinciden en la realización de las pausas. Igualmente sucede inmediatamente antes de la conjunción "y". En el caso de frases explicativas no separadas por signos de puntuación se da una cierta disparidad de realizaciones: mientras unos realizan pausas de una duración considerable, otros las eliminan. Con todo, aun en el caso de supresión de la pausa, se da un efecto de elongación de la última sílaba antes de la frase explicativa, de manera muy semejante al de las pausas con silencio. Esta situación la hemos observado al inicio de las frases explicativas:

"que daba vueltas" y "dándole vueltas a esa palanca", pero también antes del verbo "cambió" de la frase: "y un poco más tarde cambió sus aspiraciones por las de ser premio Nobel". Este caso se explicaría por la tendencia a introducir una pausa en medio de un grupo fónico relativamente largo. Los grupos acentuales inmediatamente anteriores a estos casos se han contabilizado por tanto como prepausales.

Duración de las vocales

Trabajando sobre el mismo texto hemos procedido a la medición de la duración de las vocales, distinguiendo el caso de vocales prepausales y no prepausales. Entendemos por vocales prepausales aquellas que corresponden a la última sílaba antes de una pausa. Estas sílabas, a su vez, pueden ser tónicas o átonas. El caso más frecuente para el castellano es el de las sílabas átonas correspondientes a palabras llanas acentuadas en la penúltima sílaba.

Tabla III. Duración de las vocales átonas no prepausales.

Vocal:	|a|	|e|	|i|	|o|	|u|
Duración (ms):	48	48	39	51	54
Desviación típica:	11	12	8	11	9
nº de casos:	40	48	23	31	24

Tabla IV. Duración de las vocales tónicas no prepausales.

Vocal:	|a|	|e|	|i|	|o|	|u|
Duración (ms):	71	64	67	67	61
Desviación típica:	10	16	14	15	19
nº de casos:	72	79	48	56	15

Tabla V. Duración de las vocales prepausales.

Vocal:	|a|	|e|	|i|	|o|	|u|
Duración (ms):	128	94	102	102	-
Desviación típica:	34	24	9	18	-
nº de casos:	31	24	8	16	-

Las duraciones promediadas de las vocales en cada situación estudiada pueden verse en las tablas III, IV y V. Para el caso de las vocales prepausales no se ha notado una diferencia significativa entre tónicas y átonas por lo que hace referencia a su duración. Estos resultados se comparan en el gráfico de la figura 2. No disponemos de datos para el caso de la vocal "u" en situación prepausal.

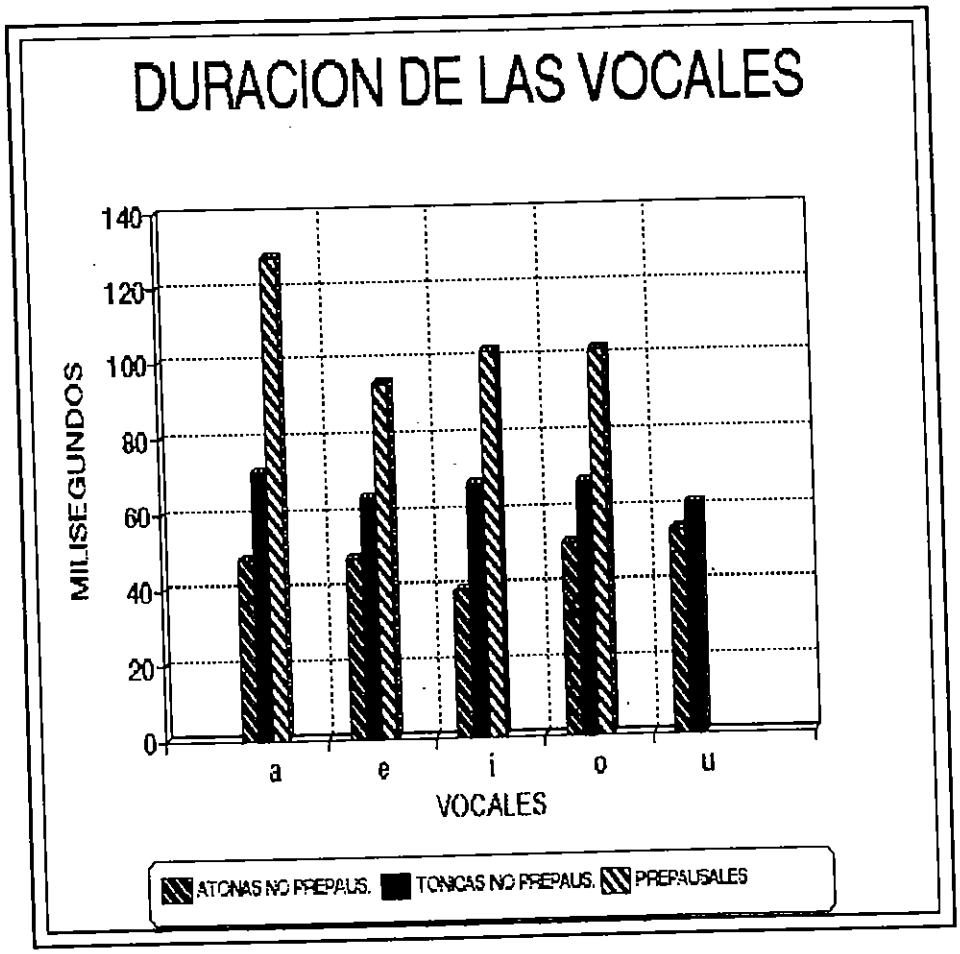


Fig. 2 Gráfico comparativo de la duración de las vocales,

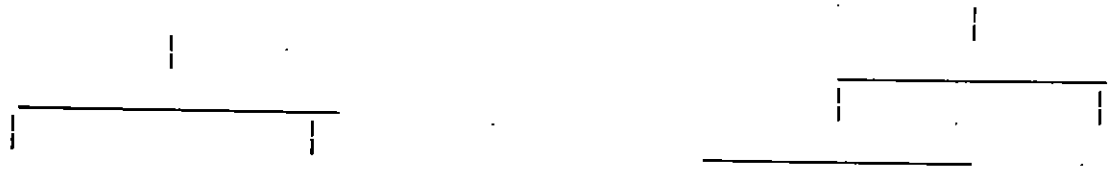
Resumiendo aún más los resultados obtenidos, podemos establecer un baremo general de alargamiento de las vocales respecto al caso de las vocales átonas y no prepausales. El factor de alargamiento sería 1,38 para las vocales tónicas y 2,23 para las vocales en situación prepausal. De todas formas, cabría mencionar que para el caso de la vocal "a" se ha notado un efecto más fuerte de alargamiento que se podría concretar en los factores 1,48 para una "a" tónica i 2,67 para una "a" prepausal. La elongación debida a la posición prepausal ha resultado extremadamente alta, lo que explica un efecto de enmascaramiento del efecto acentual antes de las pausas. Somos conscientes de que este resultado difiere notablemente del estudio realizado sobre el castellano argentino (BORZONE, 83), donde se encontraron diferencias menos acentuadas.

III. LA ENTONACIÓN

Una hipótesis de partida

Frecuentemente se considera que la entonación de un texto está estrechamente ligada a la estructura sintáctica del mismo (CHOMSKY & HALLE 1968). Con todo la correspondencia entre prosodia y sintaxis no es biunívoca; de forma que otros muchos factores, no siempre bien conocidos, intervienen también en la estructura prosódica (MARTIN, 1982, 1986). Entre estos factores cabría resaltar el efecto del ritmo de la elocución, que suponemos particularmente importante para el caso de la lectura que aquí nos interesa. .

La estructura entonativa se construye a partir de unas unidades básicas que serían los grupos acentuales, los cuales se estructurarían de una forma ordenada dentro de la unidad más amplia constituida por los grupos fónicos situados entre dos pausas consecutivas. Las pausas vienen determinadas básicamente por los signos de puntuación, pero se dan también otras circunstancias que ocasionan pausas más o menos sistemáticas. En francés se ha sugerido un sistema (MARTIN, 1982) relativamente sencillo de establecer esta estructura a partir de un contraste de pendientes de F0 entre el último grupo acentual y los anteriores, y entre los grupos acentuales yuxtapuestos con la misma dependencia jerárquica. El punto de partida viene dado siempre por el signo de la pendiente del último grupo acentual, los demás se construyen por oposiciones sucesivas entre elementos de un mismo nivel jerárquico. La estructura jerárquica no obedece estrictamente a reglas sintácticas, sino que responde también a condicionantes rítmicos. El modelo sugerido por Ph. MARTIN responde a esquemas entonativos como los de la figura 3.



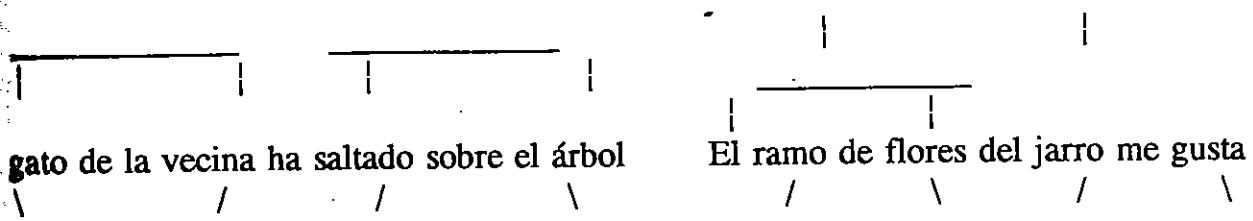
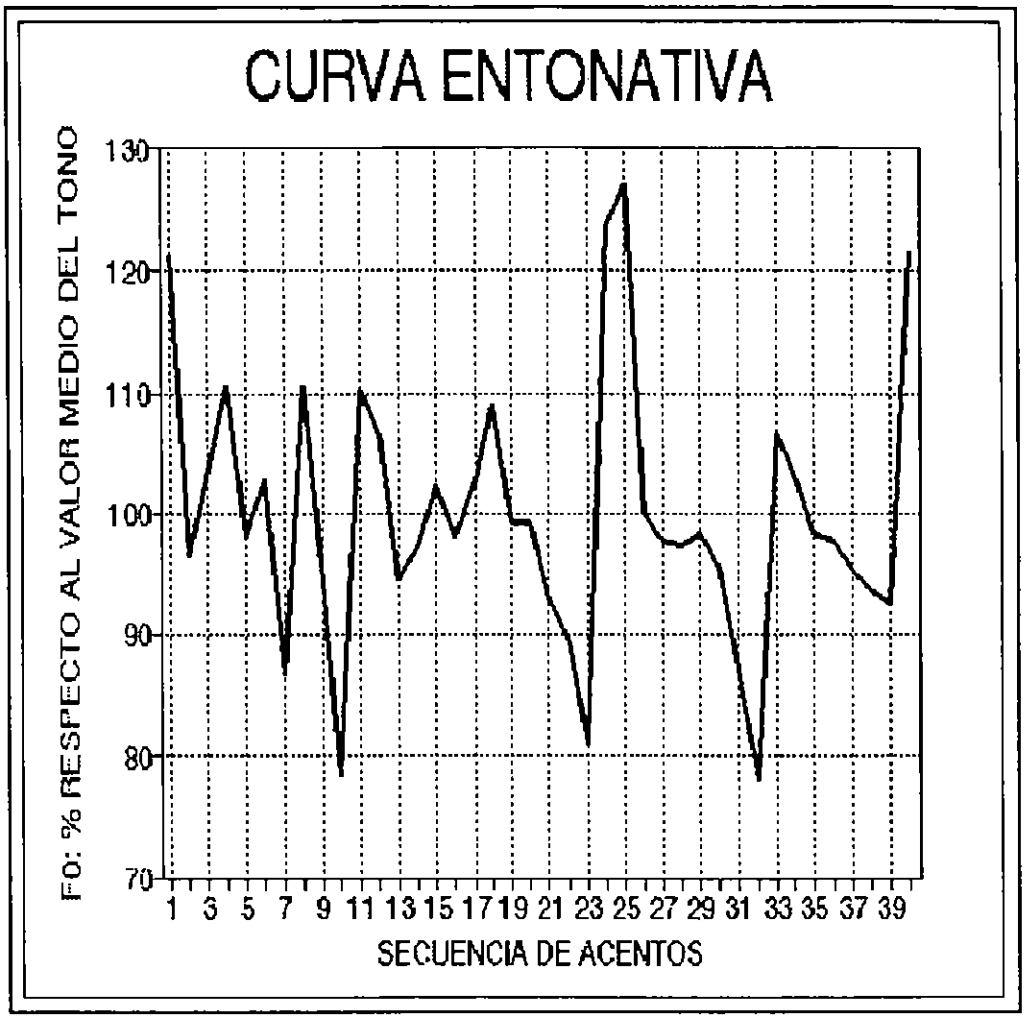


Fig. 3. Modelos entonativos según Ph. MARTIN.

Partiendo de esta hipótesis y siguiendo la misma pauta sugerida para el francés (MARTIN, 1982) y el italiano (MARTIN, 1977) hemos realizado las mediciones absolutas e incrementales en el intervalo correspondiente a cada vocal tónica del corpus considerado. Es evidente que los cambios en la curva melódica se extienden igualmente al resto de fonemas sonoros dentro del grupo acentual; pero hemos considerado que su valor más significativo es el correspondiente a la vocal tónica. En cualquier caso esta reducción se puede interpretar como una necesidad de normalización de las medidas a efectos comparativos entre grupos fónicos.

La declinación

La medición del valor absoluto de F0 para cada vocal tónica es un buen recurso para una estimación de la declinación a lo largo de un grupo fónico. Es bien conocida la polémica existente entre partidarios y detractores de la curva declinativa global de F0 a lo largo de cada grupo fónico (LADD, 1988). Si se examinan las curvas melódicas del castellano difícilmente se aprecian formas declinativas claramente identificables o cuantificables como se ha propuesto por parte de algunos autores refiriéndose a la lengua inglesa (PIERREHUMBERT, 1981) (KUTIK, 1983) (COOPER, 1985). De todas formas si que se detecta una cierta tendencia global decreciente de las sucesivas fluctuaciones melódicas. Los resultados obtenidos considerando la media de los valores absolutos de F0 en las vocales tónicas normalizados respecto al valor medio del F0 de cada parlante son los que se pueden apreciar en el anexo. Es interesante analizar también la forma gráfica (Fig. 4) de estas curvas declinativas normalizadas y promediadas.



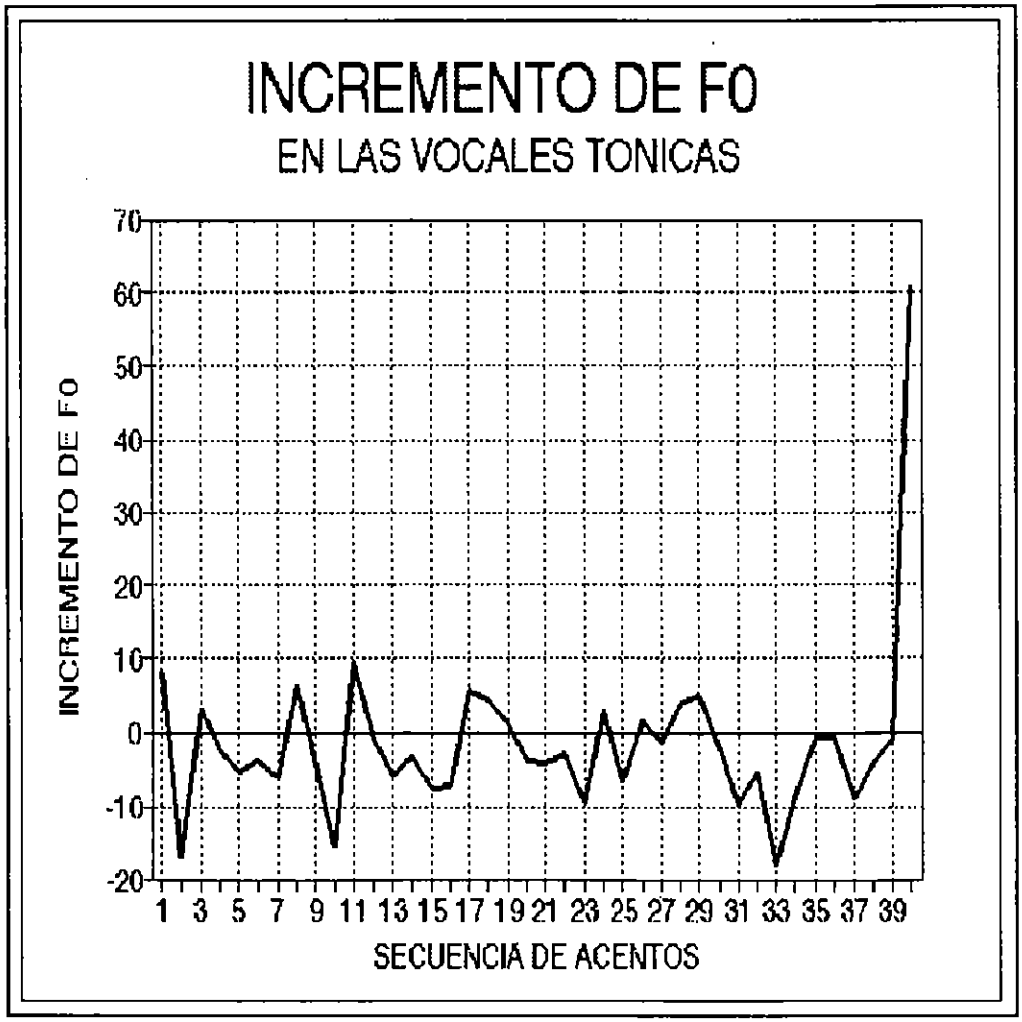


Fig. 4 Valor medio de las curvas declinativas normalizadas respecto al valor medio del F0 de cada parlante.

Si consideramos esta curva por grupos fónicos se obtienen

los resultados de la tabla VI, en la cual no se aprecia una ley uniforme de variación de F0 por sílaba o por grupo fónico; pero sí una clara tendencia monótona decreciente dentro de cada grupo acentual. Si excluimos la variación del primer grupo fónico, que pretende marcar un énfasis inicial del texto leído, nos podríamos quedar con un intervalo del ritmo de declinación comprendido entre -2,1 % y -0,6 % por sílaba pronunciada. Si contabilizamos las variaciones por grupos acentuales se da una mayor dispersión, lo que confirmaría también aquí la tendencia al ritmo silábico del castellano por encima de un ritmo acentual.

Fig. 5. Valor promediado de los incrementos de F0 durante las vocales tónicas.

Tabla VI. Declinación por grupos fónicos normalizados respecto a una F0 de 100.

GRUPO FÓNICO	Variación total	Variación por sílaba	Variación por grupo acentual	nº de sílabas	nº de grupos acentuales
Hubo una época,	-24	-4,0	-12,0	6	2
yo debía tener unos cinco años,	-24	-2,0	-4,8	12	5
en que me encantaban ... tranvías,	-32	-2,1	-10,7	15	3
de aquellos tranvías de entonces	-16	-1,8	-5,3	9	3
que no tenían puertas	-4	-0,6	-1,3	7	3
y en los que había una ... vueltas.	-28	-1,3	-4,0	22	7

A mi siempre me ... palanca.	-49	-2,0	-5,4	24	9
Y un poco más tarde ... premio Nobel,	-14	-0,6	-2,0	22	7

En la figura 5 puede verse también el valor incremental de F0 en todas las vocales tónicas, promediado respecto de todos los informantes. En este gráfico, si exceptuamos el interrogante final, se observa un predominio de las pendientes negativas respecto de las positivas, lo que confirma la tendencia declinativa a lo largo de toda la elocución.

Estructura sintáctica y entonación

En principio, con esta experiencia no se pretende realizar un estudio completo de la correspondencia entre la estructura sintáctica y las fluctuaciones melódicas de los grupos acentuales. Con todo, podemos sacar algunos indicios que podrían orientar un estudio posterior más extenso en esta línea. Se confirma una marcada tendencia al contraste en el signo de las pendientes respecto a la pendiente final del grupo fónico. En los grupos fónicos enunciativos el último grupo acentual es siempre decreciente, mientras que los anteriores situados al mismo nivel jerárquico tienden a ser crecientes. El último grupo fónico del texto analizado, debido a su forma interrogativa presenta una pendiente decreciente en todos los grupos acentuales, en virtud de su oposición a la pendiente fuertemente creciente del interrogante final. El fenómeno tiene, con todo, sus excepciones. Por ejemplo la frase parentética "yo debía tener unos cinco años" tiene una pendiente decreciente en todas sus vocales tónicas excepto la primera. El fenómeno se podría explicar por el efecto de reducción sistemática del tono de voz en los incisos parentéticos (KUTIK, 1983).

Otro fenómeno que destaca claramente es la presencia de dos grupos acentuales decrecientes sucesivos antes de cada una de las pausas y no uno solo como sugería la hipótesis inicial de trabajo. Parece como si en el caso de la lectura existiese una cierta preplanificación de las pausas que adelantaría sus efectos. El fenómeno se ha constatado sistemáticamente antes de cada una de las pausas detectadas por intervalos de silencio, y también delante de las pausas virtuales detectadas por la reducción del ritmo de pronunciación de la última sílaba.

IV. CONCLUSIONES

Con este trabajo se ha realizado un estudio orientativo previo sobre el ritmo y la entonación propios de la lectura de la lengua castellana en vistas a la obtención de las reglas prosódicas básicas que darían una naturalidad suficiente a un conversor texto-voz.

Como punto de partida se confirma la utilidad de trabajar la prosodia dentro de los grupos fónicos separados por pausas. De cara a la generación automática de modelos prosódicos de un conversor texto-voz, no se aprecian efectos importantes de ritmo o entonación que impliquen una relación entre grupos fónicos distintos. La duración máxima de los grupos fónicos observados es de 24 sílabas o 9 grupos acentuales. La duración de las pausas más normales fluctúa entre 1 segundo y 300 ms. Las pausas menos sistemáticas insertadas para dividir grupos fónicos largos son algo más cortas: entre 100 y 200 ms. En

unos casos estas pausas menos sistemáticas se substituyen por una elongación importante la última sílaba, que se comporta como una pausa virtual.

Básicamente se descarta la posibilidad de un ritmo acentual para el castellano. Tampoco se puede hablar de un ritmo estrictamente silábico por la gran diferencia detectada entre la duración de las sílabas prepausales y no prepausales. Sin embargo se aprecia una cierta constancia de la duración para cada uno de los dos tipos de sílabas. Estadísticamente, la diferencia entre los dos grupos corresponde a un factor 1,92. Naturalmente se dan otros factores, no estudiados aquí, que afectan a la duración de la sílaba, como son el número y el tipo de fonemas, el carácter tónico o átono y la forma abierta o cerrada.

Se ha estudiado particularmente la duración intrínseca de las vocales y su variación de acuerdo con las modalidades átona, tónica y prepausal. La proporción respectiva resulta de 1 : 1,38 : 2,23. Para el caso de la vocal "a" se ha observado una mayor sensibilidad a estos efectos, lo que se manifiesta por la proporción 1 : 1,48 : 2,67.

Dentro de cada grupo fónico se ha detectado una declinación de F0 que fluctúa entre los valores de -2,1% y -0,6% por sílaba, de forma que tiende a mantenerse constante dentro de cada grupo fónico. No se ha observado una declinación entre grupos fónicos sucesivos.

Dentro de cada grupo fónico se aprecia un efecto de contraste sistemático entre la pendiente de los dos últimos grupos acentuales y la de los anteriores. En las frases enunciativas los dos grupos acentuales finales son claramente decrecientes y los demás tienden a ser crecientes. En la frase interrogativa se manifiesta el mismo contraste, pero con los signos de las pendientes invertidos. De momento no se ha observado una relación clara entre la estructura sintáctica y la estructura prosódica, para lo cual se precisaría un estudio más detallado. Debido a la variedad de opciones entonativas, el hecho de trabajar únicamente con los valores medios de las pendientes puede enmascarar ciertos efectos, únicamente detectables en el estudio individual de cada informante.

AGRADECIMIENTOS

Este trabajo se ha realizado en el marco del proyecto PB90 - 0704 de la DGICYT y también gracias a la generosa colaboración de la O.N.C.E. y de la Empresa Ciberveu S.A. que dan soporte a un grupo de investigación sobre tecnologías de voz aplicadas a minusválidos radicado en el Departamento de Acústica de Ingeniería La Salle.

ANEXO: RELACIÓN DE RESULTADOS PROMEDIADOS POR GRUPOS ACENTUALES

Nº	Grupo acentual	nº sílabas fonéticas	Duración (ms) (grupo tónico)	Duración (ms) (vocal tónica)	F0 normalizada (vocal tónica)	Incremento F0 (vocal tónica)
1	Hubo	2	177	74	121	+8,1
2	una época,	4	642	100	97	-17
3	yo	1	143	63	104	+3,2
4	debía	3	245	60	111	-2,4
5	tener	2	261	81	98	-5,3
6	unos cinco	4	516	61	103	-3,9
7	años,	2	409	76	87	-6,0
8	en que me encantaban	6	762	62	110	+6,4
9	los conductores	5	610	67	94	-4,4
10	de tranvías,	4	712	102	78	-15,5
11	de aquellos	3	386	58	110	+9,7
12	tranvías	3	401	73	107	-1,0
13	de entonces,	3	680	78	94	-5,8
14	que no	2	227	70	97	-3,3
15	tenían	3	320	63	102	-7,5
16	puertas	2	499	50	98	-7,1
17	y en los que había	6	601	72	103	+5,4
18	una	2	165	39	109	+4,3
19	palanca	3	323	63	99	+1,0
20	de freno	3	388	61	99	-3,8
21	muy grande	3	524	74	93	-4,2
22	que daba	3	307	77	89	-3,1
23	vueltas.	2	554	74	81	-9,5
24	A mí	2	195	74	124	+3,1
25	siempre	2	343	48	127	-6,8
26	me emocionaba	5	580	78	100	+1,6
27	ver	1	166	73	98	-1,6
28	al conductor	4	630	96	97	+3,9
29	dándole	3	356	71	98	+5,0
30	vueltas	2	365	65	95	-2,1
31	a esa	2	247	61	87	-9,6
32	palanca.	3	465	79	78	-5,3
33	Y un poco más	4	493	69	107	-18
34	tarde	2	315	62	103	-8,5
35	cambió	2	285	56	98	-0,5
36	sus aspiraciones	6	800	47	98	-0,8
37	por las de ser	4	552	68	95	-9,0
38	premio	2	254	57	94	-4,0
39	Nobel,	2	312	85	93	-1,0
40	¿no?	1	240	150	122	+61

REFERENCIAS

BORZONE, A.M. & SIGNORINI, A. (1983). "Segmental duration and rythm in Spanish". Journal of Phonetics, 11, pp. 117-128

CHOMSKY, N. & HALLE, M. (1968). "The sound pattern of English". Harper and Row. New York.

COOPER, W.E.; EADY, S.J. & MULLER, P.R. (1985). "Acousticals aspects of contrastive stress in question-answer contexts". J.A.S.A., 77 pp. 2142-2156.

CRYSTAL, T.H. & HOUSE, A.S. (1988). "Segmental durations in connected-speech signals: Syllabic stress". J.A.S.A. 83, pp. 1574-1558.

KUTIK, E.J.; COOPER, W.E. & BOYCE, S. (1983). "Declination of fundamental frequency in speaker's production of parenthetical and main clauses". J.A.S.A., 73 pp. 1731-1738.

LADD, D.R. (1988). "Declination reset and hierarchical organisation of utterances". J.A.S.A., 84 pp. 530-544.

MARTIN, Ph. (1977). "Une théorie pour l'intonation de l'italien". Université Libre de Bruxelles. Rapport d'activitées de l'Institut de Phonétique (1977) pp. 95-113.

MARTIN, Ph. (1982). "Phonetic realisations of prosodic contours in French". Speech Communication, vol. 1 pp. 283-294.

MARTIN, Ph. (1986). "Structure prosodique et structure rythmique pur la synthèse". 15èmes. Journées d'études sur la parole". Aix en Provence. 27-30 mai.

PIERREHUMBERT, J. (1981). "Synthesizing intonation". J.A.S.A. 70 pp. 985-995.

QUILIS, A. & HERNANDEZ ALONSO, C. (1990). "Lingüística española aplicada a la terapia del lenguaje". Ed. Gredos. Madrid.

RODRIGUEZ, M.A.; GREGORIO, J.; MACARRON, A, & MONZON, L. (1992). "AMIGO: un conversor texto-voz para el español". VIII congreso SEPLN. (sept. 1992). Boletín nº 13 SEPLN pp. 389-400.

TOLEDO, G.A. "El ritmo en el español. Estudio fonético con base computacional". Ed. Gredos. Madrid.