

Evaluación acústica y análisis prosódico de la voz esofágica

M. Heliodora Cuenca Villarín, hcuenca@us.es

Marina Barrio Parra, mmbpp@us.es

Antonio Páez Moriana, labfon1@siff.us.es

Universidad de Sevilla

El objeto de este estudio ha sido realizar un análisis acústico de la voz erigmofónica o esofágica en 10 sujetos sometidos a una laringectomía total y contribuir de esta manera a la caracterización del habla esofágica en español. Se ha explorado la realización de la curva melódica en diversos tipos de enunciados, así como un análisis temporal de sílabas acentuadas/inacentuadas y distribución de pausas, factores relacionados con un habla más natural.

Los sujetos fueron clasificados en tres grupos mediante un test de percepción según su habilidad para producir voz esofágica, atendiendo a criterios de inteligibilidad y aceptabilidad. Las grabaciones de los tres grupos de informantes realizadas en un Mini Disc Sony MZ-R3, constituyen un corpus que incluye dos producciones de "a" sostenida y lectura de un texto. El análisis se llevó a cabo por medio del Multidimensional Voice Program (MDVP) de Kay Elemetrics, y del Praat del Instituto de Ciencias Fonéticas de Ámsterdam.

El MDVP se empleó para algunas de las variables estudiadas tradicionalmente en fonética acústica: frecuencia fundamental (F0), tiempo máximo de fonación (TMF), intensidad, y espectrograma de banda estrecha de la "a" sostenida; los datos obtenidos han sido comparados con aquellos publicados para voces normales también analizadas con el MDVP. El Praat resultó más conveniente para analizar la curva melódica, intensidad, duración, y velocidad de habla o "speech rate"; los datos registrados en los tres grupos permiten relacionar aceptabilidad perceptiva con los datos objetivos de los parámetros acústicos.

Introducción

Los laringectomizados que se comunican mediante voz esofágica o erigmofonía tienen que utilizar el aire contenido inyectado en el esófago como fuente de energía sonora, la cual hace vibrar el segmento faringoesofágico, en lugar de las cuerdas vocales. En los últimos años se ha incorporado a la evaluación de la voz una nueva herramienta: el análisis acústico. El análisis acústico permite obtener datos objetivos frente a la subjetividad del diagnóstico perceptivo del profesional. El oído percibe la voz como un todo y el análisis acústico

separa los distintos componentes de manera que pueden ser valorados cuantitativamente.

A la hora de abordar el análisis acústico de la voz erigmofónica nos encontramos con diversos problemas propios de una onda más o menos irregular y de corta duración. El estudio acústico aplicado a las voces patológicas se basa sobre todo en la cuantificación de la cantidad de ruido indeseable en la señal acústica, o en la alteración de la periodicidad en la vibración de las cuerdas vocales (Rosique *et al.*, 2003: 362)

Este estudio aborda el análisis de algunas variables estudiadas tradicionalmente en fonética clínica: duración (Tiempo Máximo de Fonación (TMF) y sílabas por minuto (spm)), intensidad, frecuencia fundamental (F0) y espectrografía. Pero además hemos incluido un estudio más lingüístico, el análisis prosódico, en un intento de aunar dos disciplinas distintas pero complementarias. La prosodia desempeña un papel más importante cuando la calidad auditiva del habla y la inteligibilidad del habla es inferior a la normal (Van Rossum, 2005:7). En el análisis prosódico las variables estudiadas son también duración, intensidad, frecuencia y pausas.

El objetivo principal del trabajo es aportar datos para posibles comparaciones con otros estudios similares y ofrecer con el análisis prosódico de la voz erigmofónica un enfoque lingüístico complementario.

Sujetos

Para el estudio se han seleccionado 10 hombres de entre 23 que habían sido sometidos a una laringectomía total y pertenecientes a la Asociación Sevillana de Laringectomizados¹. Los sujetos son de la provincia de Sevilla o llevan muchos años viviendo en ella. Fueron seleccionados por un grupo de personas acostumbradas al habla esofágica siguiendo un criterio de inteligibilidad y aceptabilidad según su habilidad para producir habla esofágica. Todos los sujetos analizados utilizan el habla esofágica de forma habitual y relativamente satisfactoria. Con edades comprendidas entre 53 y 77 años y con una media de 66, el tiempo transcurrido desde la operación oscilaba entre 2 y 23 años, y una media de 7. Asimismo, los seleccionados fueron clasificados a su vez en tres grupos. El grupo I,

¹ Este estudio no hubiera sido posible sin la inestimable colaboración de la Asociación Sevillana de Laringectomizados.

de alto nivel, lo constituyen los informantes 1, 2 y 3; en el grupo II con un nivel medio, están los informantes 4, 5, 6 y 7; y finalmente en el III, con un nivel bajo, los informantes 8, 9 y 10. Además hemos utilizado un grupo que hemos llamado mixto y que incluye dos sujetos del grupo I y dos del grupo III (informantes 1, 2, 8 y 9). Pese a esta clasificación perceptiva, se observan diferencias, incluso dentro del grupo, basadas principalmente en el ritmo, la articulación de los sonidos, la dosificación del aire en la producción de voz, la entonación y el F0.

Procedimiento

Las grabaciones se llevaron a cabo en la Asociación Sevillana de Laringectomizados y en el Hospital Virgen del Rocío de Sevilla, de forma individual y en salas acondicionadas, procurando que estuvieran aisladas acústicamente. Se utilizó una grabadora digital portátil de Minidisc Sony MZ-R3 con un micrófono profesional cardioide Shure SM58, dispuesto a una distancia de la boca de unos 15 cm sobre un pie y en ligero ángulo de aproximadamente 30° de elevación por la parte posterior, para evitar en lo posible el ruido producido por el aire que entra y sale del estoma. A continuación se pasó la señal al ordenador con una frecuencia de muestreo de 44.100 Hz y una cuantización de 16 bits, para su almacenamiento y posterior análisis. Los programas de análisis utilizados son el MDVP (*Multi-Dimensional Voice Program*) modelo 4305 de Kay Elemetrics, y el Praat v. 4.4.22 de Paul Boersma y David Weenink del *Instituto de Ciencias Fonéticas* de la Universidad de Ámsterdam.

El material utilizado consiste en grabaciones de la /a/ sostenida y la lectura de un texto que les proporcionamos, diseñado *ad hoc*. Se les pidió a los informantes que produjeran una /a/ sostenida durante todo el tiempo que les fuera posible y en un tono y volumen confortable. Seguidamente se les facilitó el texto para su lectura, del que hemos seleccionado para este trabajo una oración declarativa o asertiva, y otra interrogativa: *Mi calle es algo chica* y *¿Quién puede seguir ese ritmo?*

Análisis

El MDVP es una herramienta que ha demostrado su eficiencia para el análisis automático de voces patológicas. Sin embargo, en el caso

de la voz erigmofónica obtuvimos pocos resultados², por lo que incorporamos el programa Praat. El MDVP sólo obtuvo resultados completos con 4 muestras (de 20) pertenecientes a 3 informantes (de 10), mientras el Praat obtuvo resultados con 18 muestras de nueve informantes³. La causa de estas limitaciones de los programas en general es debida a la irregularidad de la onda y a la baja frecuencia de la misma. La extracción automática de la frecuencia fundamental (perceptivamente tono), por un lado considera la onda como un todo y no sólo el pico inicial, por lo que no puede encontrar dos ciclos similares y falla a la hora de calcular la Frecuencia Fundamental. Por otro lado, en lo que respecta a la baja frecuencia, el programa está configurado de tal manera que no puede detectar pulsos glotales que estén separados entre sí más de 20 ms (siendo la menor frecuencia registrada de 50 Hz), (Ladefoged, 2003: 82). Con el Praat conseguimos extraer un mayor número de resultados, bajando el mínimo del rango de frecuencias, pero sin modificar los valores estándares del umbral de la sonoridad. Cuando no hubo resultados con ninguno de los programas, la extracción se realizó manualmente. De cualquier forma, hemos incluido los valores de F0 obtenidos con el MDVP y con el Praat.

Duración

TMF (Tiempo Máximo de Fonación) /a/: en milésimas de segundo. En la voz laríngea se consideran valores anormales los inferiores a 15 seg. (Jackson-Menaldi, 2002: 213). Se midió la duración de la /a/ sostenida y se calculó la media de los tres grupos de hablantes, por separado y conjuntamente.

TMF en cadena hablada: se midieron las producciones de los “eructos” (golpes de voz) del grupo mixto, durante la lectura del texto, para su comparación con el TMF de la /a/ sostenida.

Intensidad

Medido en decibelios (dB) relativos al umbral de la audición. Para un adulto normal la intensidad de fonación durante la conversación en un nivel medio alto-fuerte está entre 75-80 dB (Jackson-Menaldi,

² Disponemos del MDVP modelo 4305 (versión para MS-DOS) por lo que desconocemos el resultado que obtendría con las nuevas versiones.

³ La explicación sobre la diferencia entre ambos programas está basada en el algoritmo de extracción del F0. Dicha explicación se encuentra en el manual de Praat bajo el epígrafe: “Voice 5. Comparison with other programs”

2002:226). La intensidad de la voz esofágica estará siempre por debajo de la voz laríngea normal. Los factores que inducen una disminución de la intensidad son un soporte respiratorio inadecuado y la ausencia de cierre glótico. Hemos calculado la intensidad media de la totalidad y de los grupos y hemos comparado contornos o curvas de intensidad.

Frecuencia Fundamental

Medido en hercios (Hz). Estrictamente hablando el tono es una propiedad auditiva (algo que se percibe), no es una propiedad acústica, ni por tanto un aspecto de la onda que se pueda medir. Desde un punto de vista práctico cuando se habla del tono de la voz, normalmente se habla de la velocidad o ritmo a la que los pulsos glotales se repiten, es decir, de la frecuencia fundamental o F0 que es el correlato acústico del tono (Ladefoged 2003:75). Rosique *et al.* (1999: 131) registran la media de la frecuencia fundamental del hablante laríngeo en 120,30 ($\pm 15,57$).

Espectrografía

Frente al análisis cuantitativo expuesto se encuentra el análisis cualitativo que ofrece la espectrografía. Permite explorar el contenido espectral de una señal y la presencia y evolución temporal de las zonas de los formantes. Una estructura formántica clara y adecuada se relaciona con un buen sistema resonador, con la ausencia de escape aéreo y con la periodicidad normal de movimiento de las cuerdas vocales (Núñez Batalla, 1999: 31).

Se calcularon las medias de los formantes de todos los sujetos y se han comparado con los valores aportados por la literatura para la voz esofágica en español. Asimismo, se comentarán algunas características de la voz esofágica en la espectrografía de banda estrecha.

Análisis Prosódico

Se han seleccionado dos enunciados, representativos de dos modalidades: declarativa o asertiva, e interrogativa, a partir de la lectura del texto por el grupo mixto y se ha realizado un análisis de los parámetros prosódicos: F0, intensidad, duración y pausas.

Resultados y discusión

TMF

Robbins recoge que Berlin halló duraciones medias de 1,3 seg. en un grupo de hablantes esofágicos “pobres” y de 2,8 en un grupo de “buenos” hablantes (Robbins et al., 1984:209). En el mismo artículo Robbins et al. registraron sus propias medias de 1,9 ($\pm 0,7$) con un número de 15 informantes. Según afirma Bellandese (1998: 2), estos valores están de acuerdo con los aportados por Sedory et al. (1989). La siguiente tabla muestra los resultados obtenidos en este estudio:

| | Todos | Grupo I | Grupo II | Grupo III |
|-----|-----------------------|-----------------------|-----------------------|-----------------------|
| TMF | 1,265 ($\pm 0,345$) | 1,248 ($\pm 0,275$) | 1,134 ($\pm 0,248$) | 1,457 ($\pm 0,466$) |

Tabla 1. Medias de los TMF de los diez informantes en conjunto y separados por grupos.

El grupo III, en contra de lo que se podría prever, alcanza los valores más altos. Según algunos autores, el TMF es mayor en individuos con mejor calidad vocal en su conjunto y los TMF más cortos están relacionados con un habla menos fluida y menos inteligible⁴. Sin embargo, nuestros resultados muestran que la subida de la media en el grupo III es debida a que los informantes 8 y 10 tienen fluidez pero están en este grupo por su calidad vocal general.

A primera vista los valores que hemos registrado son bastante inferiores a los recogidos en la literatura. Sin embargo, también registramos el tiempo de fonación dentro de la cadena hablada para comprobar si los tiempos se correspondían con los que acabamos de exponer de la /a/ sostenida. Se seleccionó el grupo mixto para ello. Los resultados muestran que el hablante es capaz de producir valores superiores a los de la /a/ sostenida, por lo que podemos pensar que la dosificación es más efectiva en la cadena hablada.

Hemos incluido las sílabas por minuto, relacionadas también con la fluidez. Según el estudio de Filter y Hyman, la media se sitúa en 135 spm. (± 31) (Robbins et al., 1984:203). Ahora bien, tenemos que tener en cuenta el número de pacientes, la habilidad para producir habla y que se trata de otro idioma.

⁴ *A practical guide to post-laryngectomy vocal and pulmonary rehabilitation* - 4th Ed. Hilgers, Frans J. M., (Ed.)
<http://www.hoofdhalskanker.info/ProvoxWeb/index.htm>

| Nº Informante | /a/ 1 | /a/ 2 | Texto | spm |
|---------------|---------|---------|---------------------------|-----|
| 1 | 1,462 s | 1,625 s | 1,820 s (10 síl., 6 pal.) | 201 |
| 2 | 1,099 s | 0,997 s | 1,509 s (7 síl., 2 pal.) | 214 |
| 8 | 1,930 s | * | 1,864 s (8 síl., 3 pal.) | 171 |
| 9 | 0,897 | 0,917 | 1,129 s (5 síl., 1pal.) | 97 |

Tabla 2. TMF y sílabas por minuto (spm) del grupo mixto. Entre paréntesis, las sílabas y las palabras pronunciadas

Las sílabas por minuto de un hablante normal en el mismo contexto son 321 según nuestros resultados⁵.

Intensidad

Nuestros resultados demuestran estar en la media de los resultados obtenidos por Damste (Heuillet-Martin, 2003: 274), que sitúa la intensidad media entre 55 y 65 dB (con posibles variaciones de 10 dB). Sin embargo son muy inferiores a los obtenidos por Robbins (1984: 207) que sitúan la media en 73,8 ($\pm 5,0$).

| | Todos | Grupo I | Grupo II | Grupo III |
|------------|---------------------------|---------------------------|---------------------------|---------------------------|
| Intensidad | 61,030 ($\pm 7,017$) | 63,802 ($\pm 6,799$) | 62,614 ($\pm 8,213$) | 56.144 ($\pm 2,139$) |

Tabla 3. Medias de intensidad de la /a/ sostenida de los diez informantes en conjunto y separados por grupos.

Las curvas de intensidad muestran tantas irregularidades y diferencias entre sí que no nos pararemos a describirlas en este trabajo, sólo comentaremos que las curvas producidas por el sujeto 2 son las que presentan mayor similitud con las de una voz normal: curva rectilínea con inicio en ascenso, valle y descenso. Además hemos observado que el 40% de los sujetos produce al final de la curva un ligero ascenso de la intensidad como en un intento de aprovechar hasta la última reserva de aire. Incluso encontramos subida de la frecuencia fundamental en la onda del informante 1; la subida de intensidad provoca subida de frecuencia.

⁵ Este trabajo pertenece a otra investigación más amplia que llevamos a cabo y que incluye un grupo control de voces normales.

Frecuencia Fundamental (F0)

El tono es difícil de percibir auditivamente por la irregularidad de los impulsos y la incidencia del timbre. El F0 de la voz esofágica oscila entre los 25 y los 200 Hz por segundo. La frecuencia media es de 66 Hz en hombres y 86 Hz en mujeres (Heuillet-Martin, 2003: 273). Robbins *et al.* establecen la media en 65,3 ($\pm 30,5$) (Robbins, 1984: 205). Rosique *et al.* registran 77,29 ($\pm 25,43$) (Rosique, 1999: 131). La media para los informantes 1 y 2 con el MDVP es de 91.25 (± 4.03) y con el Praat de 75 (± 12.4).

Hemos analizado el F0 del grupo mixto, ya que en un primer momento fueron los únicos informantes que con el MDVP, al menos en alguna de sus aes, mostraban todos los resultados numéricos sin error. Sin embargo los resultados mostrados no eran del todo satisfactorios con respecto a lo apreciado visualmente, en lo que se refiere a los informantes 8 y 9, por lo que se procedió a analizar la señal con el Praat. Los datos no presentaron grandes variaciones, por lo que finalmente se extrajo manualmente el F0 de aquellas ondas cuyos valores aparentemente no coincidían con lo percibido visual y auditivamente. En el caso del informante 8, el MDVP calculó un valor muy aproximado al obtenido manualmente, no así el Praat. Sin embargo, ambos programas fallaron con el informante 9: o no obtuvimos resultados o no se correspondían con la extracción manual.

| Informantes | Manual | MDVP | Praat |
|-------------|---------------|---------------|---------------|
| 1 /a/ 1 | | 89 (121-66) | 58 (107-53) |
| 1 /a/ 2 | | 91 (109-66) | 77 (94 -51) |
| 2 /a/ 1 | | 88 (111-67) | 77 (86 -58) |
| 2 /a/ 2 | | 97 (109 -68) | 88 (104 -83) |
| 8 /a/ 1 | 153.47 (sel.) | 169 (249-120) | 228 (380 -40) |
| 9 /a/ 1 | 63.88 (sel.) | 363 (519-252) | 273 (276-197) |
| 9 /a/ 2 | 45.41 (sel.) | --- | --- |

Tabla 4. Medias de la frecuencia fundamental del grupo mixto. Entre paréntesis la máxima y la mínima.

Espectrografía

En general los formantes de las vocales son más elevados en el habla esofágica que en el habla normal, como demuestran Cervera *et al.* (2001:993), los cuales obtuvieron medias de 866 Hz (± 100) para F1 y 1498 (± 257) para F2 de la /a/ en un contexto CVCV. La siguiente tabla recoge nuestros resultados por grupos y con la totalidad de los informantes:

Beatriz Gallardo, Carlos Hernández y Verónica Moreno (Eds): *Lingüística clínica y neuropsicología cognitiva. Actas del Primer Congreso Nacional de Lingüística Clínica. Vol 2: Lingüística y evaluación del lenguaje*, coord. por C. Hernández y M. Veyrat. ISBN: 84-370-6576-3.

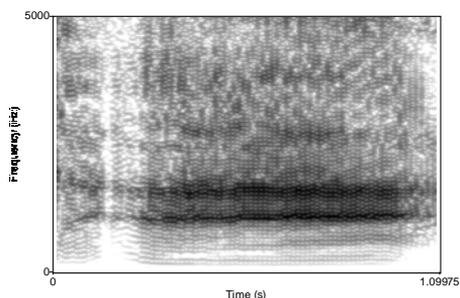
| | F1 | F2 | F3 |
|----------------|------------|-------------|-------------|
| Grupo 1 | 1060 (±47) | 1610 (±128) | 2837 (±244) |
| Grupo 2 | 1066 (±24) | 1744 (±127) | 2921 (±127) |
| Grupo 3 | 953 (±43) | 1540 (±186) | 2986 (±61) |
| Todos | 1030 (±63) | 1642 (±165) | 2915 (±162) |

Tabla 5. Medias de frecuencias de formantes en hercios de la /a/ sostenida, entre paréntesis la desviación estándar.

Nuestros datos resultaron ser bastante más altos que los recogidos por Cervera *et al.* Sin embargo no hay que olvidar que se trata de distintos contextos, y, como estos autores afirman, podemos encontrar diferencias ya que según con qué consonante coarticule la vocal, la frecuencia de los formantes variará. Sí es un dato común el de las frecuencias superiores a las normales.

La espectrografía de la voz esofágica presenta rasgos comunes a todos los informantes. En primer lugar la presencia de ruido que, según Núñez Batalla (1999:41), es producido por un flujo aéreo turbulento debido al cierre glótico incompleto, o por un comportamiento vibratorio irregular; y en segundo lugar una estructura formántica más o menos clara, ya que como explican Heuillet-Martin y Conrad (2003:269), los resonadores no han sufrido modificaciones (exceptuando la faringe) y proporcionan a la voz esofágica una estructura formántica parecida a la laríngea (2003: 275).

Rasgos más particulares, dependientes de la habilidad del hablante esofágico, son la presencia o ausencia de armónicos y la presencia, en hablantes con muy baja frecuencia fundamental, de pulsos producidos por la neoglotis. En la figura 1 hemos utilizado dos espectrogramas como ejemplo: el informante 2 presenta armónicos y los dos primeros formantes muy delimitados, sin embargo el informante 9 no presenta armónicos y los pulsos neoglóticos están muy marcados (son aún más visibles en el espectrograma de banda ancha).



Beatriz Gallardo, Carlos Hernández y Verónica Moreno (Eds): *Lingüística clínica y neuropsicología cognitiva. Actas del Primer Congreso Nacional de Lingüística Clínica. Vol 2: Lingüística y evaluación del lenguaje*, coord. por C. Hernández y M. Veyrat. ISBN: 84-370-6576-3.

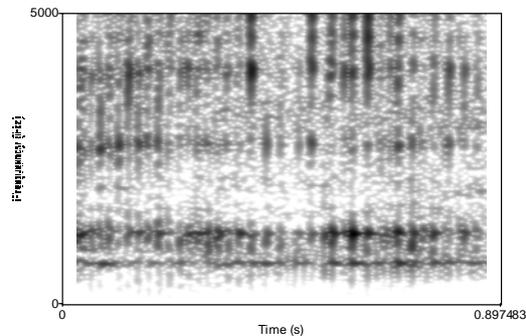


Fig. 1. Ejemplos de espectrogramas de banda estrecha (43 Hz) de los informantes 2 y 9.

Análisis Prosódico

La segmentación se llevó a cabo por medio de información audio-visual. Debemos destacar la enorme dificultad que hemos encontrado para analizar determinados parámetros, sobre todo la frecuencia fundamental de las sílabas antes de pausa, pues, probablemente debido a la falta ya de reserva de aire, la onda era muy irregular y el programa erróneamente terminaba por interpretarla como una onda periódica de muy alta frecuencia; de ahí que dichos datos hayan sido desestimados en nuestro análisis.

1. Duración

La prosodia contribuye a estructurar un enunciado, ayudando al oyente en su procesamiento del mensaje (Van Rossum 2005:10-11). Por medio de las fronteras prosódicas, los oyentes localizan las principales fronteras de dominio sintáctico.

Los rasgos asociados a las fronteras prosódicas forman una jerarquía, con el alargamiento final como rasgo principal, y al que se añaden el uso de pausas y F0. Los hablantes con voz esofágica pueden encontrar limitaciones en el grado de alargamiento final, pues merma aún más su reducido suministro de aire. En este estudio, por tanto, hemos analizado la organización temporal del enunciado asertivo del grupo mixto, concentrándonos en la relación entre acento y duración, así como en el uso del alargamiento final como rasgo delimitador de dominios prosódicos.

En la siguiente tabla se exponen los valores de duración para las vocales del primer grupo fónico esperado de la oración declarativa en habla laríngea. Se han marcado las sílabas acentuadas por medio de un

asterisco, y las sílabas en posición prepausal en negrita. La variabilidad entre hablantes se observa en el diverso número y posición de las pausas de inyección de nuestros informantes.

| Sílabas | Informante 1 | Informante 2 | Informante 8 | Informante 9 |
|---------|--------------|--------------|--------------|--------------|
| Mi | 67 | 92 | 156 | 119 |
| ca* | 122 | 98 | 98 | 159 |
| lle | 108 | 125 | 48 | 121 |
| es | 166 | 114 | 273 | 156 |
| al* | 93 | 87 | 90 | 106 |
| go | 70 | 154 | 72 | 61 |
| chi* | 66 | 73 | 130 | 160 |
| ca | 70 | 183 | 98 | 156 |

Tabla 6. Duración vocálica (en ms) del primer grupo fónico esperado en la oración declarativa. *Mi calle es algo chica*.

Como puede apreciarse en la tabla anterior, la explotación de la duración como rasgo caracterizador del acento o delimitador de dominios prosódicos ha sido más o menos consistente en nuestros hablantes. Los informantes 1, 8, y 9 sistemáticamente utilizan la duración para marcar la vocal acentuada. Sin embargo, el informante 2 destaca por medio de la duración las vocales en sílaba final de palabra tanto en posición prepausal como no prepausal.

En el resto de grupos fónicos que han conformado este enunciado, el comportamiento detectado ha sido, en general, el mismo.

2. Intensidad (Rango 30-500dB)

A continuación trataremos de corroborar la disminución de intensidad al final de palabra (Heuillet-Martin y Conrad, 2003: 274), comúnmente relacionada con el habla esofágica (Van Rossum, 2005:8), además de describir la relación que se establece entre intensidad y acento.

En el informante 1 se observa en el sintagma una tendencia decreciente en los valores de intensidad al mismo tiempo que la vocal en sílaba acentuada ha presentado valores superiores que en sílaba inacentuada.

El informante 2, al igual que el informante 1, presenta un crecimiento del valor de intensidad en la segunda sílaba del primer grupo en la sílaba acentuada de *calle*, y tras la pausa que sigue a *es* la tendencia es decreciente, independientemente de la acentuación.

| Sílabas | Informante | Informante | Informante | Informante |
|---------|------------|------------|------------|------------|
| | 1 | 2 | 8 | 9 |
| Mi | 60 | 50 | 43 | 46 |
| ca* | 72 | 54 | 52 | 53 |
| Lle | 59 | 52 | 55 | 47 |
| Es | 65 | 58 | 55 | 53 |
| Al* | 72 | 59 | 48 | 51 |
| Go | 62 | 58 | 44 | 35 |
| chi* | 53 | 49 | 48 | 45 |
| Ca | 46 | 52 | 50 | 42 |

Tabla 7. Intensidad vocálica (en dB) del primer grupo fónico esperado en la oración
Mi calle es algo chica.

Sin embargo, en los informantes 8 y 9 la condición de acentuada supone un crecimiento del valor de intensidad aunque no debemos olvidar que en el informante 8 la sílaba final rompe dicha inercia, si bien la diferencia quizás no sea significativa⁶.

Esta tendencia decreciente de los valores de intensidad es característica también de la voz laríngea.

Dorta (2005:100), así como Congosto (2005: 240) han observado para diversas variedades dialectales del español que en cada sintagma la intensidad desciende desde la primera vocal hasta la última, aunque no siempre de manera progresiva.

3. Frecuencia fundamental (Rango 30-300Hz)

En este apartado analizaremos, por un lado, la relación entre F0 e intensidad, pues según algunos autores, entre ellos Heuillet-Martin y Conrad (2005:173), la voz esofágica no puede disociar las variaciones entre la intensidad y el tono. Cuando la intensidad aumenta, la frecuencia también lo hace, e inversamente. Por otro lado, se describirá la configuración tonemática de la oración interrogativa pronominal enfática, pues la inflexión final o tonema es la parte más informativa de la unidad melódica (Cantero, 2002:122).

Las posibilidades entonacionales están relacionadas con la dosificación de la presión subneoglótica y con la tonicidad de la neoglotis. Sin embargo, Van Rossum (2005:101-103) ha observado que los hablantes con voz esofágica aumentan la intensidad en las altas frecuencias para transmitir un movimiento tonal ascendente, pero no suelen disminuir con efectividad la intensidad para transmitir un

⁶ Según Dorta y Hernández (2005:100), una diferencia de menos de 3 dB no se consideraría significativa.

movimiento tonal descendente, es decir, el esfuerzo vocal es manipulado en mayor grado cuando se ha producido una inflexión en anticadencia que en cadencia.

Excepto en el hablante 2 donde F0 e intensidad parecen comportarse de modo independiente, los informantes 1, 8 y 9 no disocian entre ambos parámetros. En el caso del informante 1, en lugar de ser un inconveniente en el momento de comunicarse, parece ser una estrategia eficiente en la transmisión de la inflexión tonemática.

Independientemente del número de pausas de inyección realizadas por cada uno de los hablantes, hemos advertido un perfil melódico similar.

En el enunciado interrogativo pronominal *¿Quién puede seguir ese ritmo?*, según Sosa (1999:144-148) o Quilis (1993:431), la parte más alta del contorno coincidiría con la primera sílaba acentuada, luego de la cual se produciría un descenso gradual para culminar con un tonema descendente. Ésta es la configuración entonativa que producen los informantes 2 y 8. La curva entonativa del informante 2 revela un comportamiento regular en cada uno de los grupos fónicos que lo conforman, es decir, una elevación de la frecuencia fundamental en la primera sílaba para descender progresivamente. El informante 8 también presenta un tonema descendente o en cadencia, pero con un uso menos consistente de la frecuencia fundamental. En la curva entonativa del informante 1, se observa cómo después del primer pico desciende en la sílaba postónica para elevarse suavemente en la primera sílaba del nuevo grupo fónico, inacentuada, y realizarse un segundo pico en la sílaba tónica, y a partir de ahí, a lo largo del último grupo fónico, descender progresivamente hasta el final. El campo tonal ha sido más amplio en el hablante 1. Los hablantes también han presentado un registro tonal más alto en el pretonema interrogativo que en el declarativo, tal y como describe Sosa (1999: 146) para voces laríngeas.

Respecto al informante 9, la escasez de datos disponibles nos impide describir su curva entonativa, si bien es verdad que por el valor frecuencial de la vocal final y de algunas de las sílabas anteriores, podemos intuir también aquí un tonema descendente.

4. Pausas

Las pausas localizadas en posiciones prosódicamente motivadas en una oración contribuyen a mejorar el reconocimiento del habla (Van Rossum, 2005: 7). Hemos interpretado como una pausa aquellos silencios que han excedido las 100 milésimas de segundo.

En este apartado hemos seleccionado dos informantes, los clasificados como 2 y 9. El informante 9 inserta numerosas pausas de inyección, de una mayor duración que las pausas de inyección del informante 2; algunas de ellas rompen sintagmas, como por ejemplo, *ese/ritmo* y llegan a presentar una duración mayor que las pausas en fronteras oracionales, y sistemáticamente el tiempo de emisión de voz es inferior al de silencio, lo cual justifica que se le asigne tal número.

Obsérvense a modo de ejemplo los datos relativos a la duración de las pausas en la oración interrogativa pronominal enfática *¿Quién puede seguir ese ritmo?*, que se ilustran en las siguientes tablas:

| Grupo fónico | Dur. Grupo fónico – pausa (ms) |
|--------------------------------|---------------------------------------|
| ¿Quién puede seguir ese ritmo? | 2100-240 |
| | 654- 387 |

Tabla 8. Valores de duración de emisiones fónicas y pausas para el informante 2 en la oración interrogativa.

| Grupo fónico | Dur. grupo fónico – pausa (ms) |
|---------------------|---------------------------------------|
| ¿Quién | 441 -1553 |
| Puede | 418 – 921 |
| Seguir | 690 – 974 |
| Ese | 537 – 1167 |
| Ritmo? | 630 – 1017 |

Tabla 9. Valores de duración de emisiones fónicas y pausas para el informante 9 en la oración interrogativa.

Asimismo, las pausas de inyección del informante 2 son de una duración inferior a la de las pausas sintácticas, mientras se observa la tendencia inversa en el informante 9. Como recuerda Van Rossum (2005: 124), las pausas esperadas deberían ser considerablemente más largas que las inesperadas, y de hecho las pausas de duración superior a 250 milisegundos se perciben con dificultad, y puesto que en el caso del informante 9 no sólo las pausas sintácticas, sino también las de inyección han superado con creces ese límite, los oyentes pueden no percibir la diferencia.

También es cierto que, como ha observado Van Rossum (2005: 125) para el holandés, el número de pausas internas desciende en las frases más cortas, y entonces el alargamiento final de grupo fónico ocurre con mayor frecuencia.

Conclusiones

El análisis acústico de voz esofágica ha revelado datos que no difieren significativamente de los hallados en la literatura, excepto en los referentes a la relación entre el TMF y la calidad de voz. En el análisis prosódico se han detectado ciertas características comunes entre nuestros informantes y hablantes con voz laríngea: el uso de inflexiones tonemáticas propias del tipo de enunciado analizado; la duración ha sido el correlato acústico asociado a la acentuación; y en el sintagma una tendencia decreciente de la intensidad aunque no siempre de manera progresiva.

No obstante, cada sujeto genera sus propios mecanismos o estrategias en la codificación prosódica de las estructuras sintácticas o unidades de información, de la misma manera que no emplean un método puro en la fonación, sino mixto y personalizado.

Por consiguiente los datos objetivos obtenidos a partir del análisis acústico de la /a/ sostenida no siempre han vaticinado la clasificación de nuestros informantes a nivel perceptivo, frente al análisis prosódico según el cual persistiría dicho orden, lo cual puede tener implicaciones para la rehabilitación vocal de estos pacientes. Sin embargo, creemos que en el caso de hablantes con un nivel muy bajo de inteligibilidad y aceptabilidad, dicho análisis acústico sería más indicativo de sus habilidades comunicativas.

Finalmente, estos resultados no pueden ser concluyentes, en primer lugar por haberse manejado un volumen reducido de datos; y en segundo lugar, por las características inherentes al corpus empleado, ya que las habilidades comunicativas reales del hablante laringectomizado pueden quedar distorsionadas por sus habilidades lectoras anteriores a la laringectomía.

Bibliografía

- Bellandese, Mary Holland (1998): *The relation between acoustic and perceptual characteristics of laryngeal, excellent tracheoesophageal and excellent esophageal speakers*. Tesis Doctoral. University of Connecticut.
- Cantero Serena, Francisco José (2002): *Teoría y análisis de la entonación*. Barcelona: Publicacions i Edicions UB.
- Cervera, Teresa, Miralles, José L. y González-Álvarez, Julio (2001): "Acoustical Analysis of Spanish Vowels Produced by Laryngectomized Subjects", *Journal of Speech, Language and Hearing Research*, 44: 988-996.
- Congosto, Yolanda (2005): "Una primera aproximación a la prosodia del habla urbana de Sevilla", *Estudios de Fonética Experimental*, 14: 225-246.

Beatriz Gallardo, Carlos Hernández y Verónica Moreno (Eds): *Lingüística clínica y neuropsicología cognitiva. Actas del Primer Congreso Nacional de Lingüística Clínica. Vol 2: Lingüística y evaluación del lenguaje*, coord. por C. Hernández y M. Veyrat. ISBN: 84-370-6576-3.

- Dorta, Josefa y Hernández, Beatriz (2005): “Análisis prosódico de un corpus de habla experimental: interrogativas absolutas con expansión en el objeto vs sin expansión”, *Estudios de Fonética Experimental*, 14: 67-123.
- Heuillet-Martin, Geneviève y Conrad, Liliane (2003): *Hablar sin laringe*. Barcelona: Lebón. Trad. de Carlota Faixa.
- Jackson-Menaldi, María Cristina (2002): *La Voz Patológica*. Buenos Aires: Editorial Panamericana.
- Ladefoged, Peter (2003): *Phonetic Data Analysis: an introduction to fieldwork and instrumental techniques*, Malde (MA): Blackwell, 2004.
- Núñez Batalla, Faustino y Suárez Nieto, Carlos (1999): *Espectrografía clínica de la voz*. Oviedo: Servicio de Publicaciones, Universidad de Oviedo.
- Quilis, Antonio (1993): *Tratado de fonología y fonética españolas*. Madrid: Gredos.
- Robbins, Joanne et al. (1984): “A comparative acoustic study of normal, esophageal, and traqueoesophageal speech production”, *Journal of Speech and Hearing Disorders*, 49: 202-210.
- Rosique, Mariano et al. (1999): “Análisis acústico de la voz en fistuloplastia fonatoria tras laringectomía total”, *Acta Otorrinolaringológica Española*, 50(2): 129-133.
- Rosique, Mariano et al. (2003): “Análisis discriminante aplicado a los formantes de las vocales castellanas en la fonación con prótesis y erigimofonía tras laringectomía total”, *Acta Otorrinolaringológica Española*, 54: 361-366.
- Sosa, Juan Manuel (1999): *La entonación del español. Su estructura fónica, variabilidad y dialectología*. Madrid: Cátedra.
- Van Rossum, Maya (2005): *Prosody in Alaryngeal Speech*. Tesis Doctoral. Utrech: LOT.