

Manuscript Number:

Title: EL ESPECTROGRAMA DE BANDA ESTRECHA COMO AYUDA PARA EL APRENDIZAJE DEL METODO GRABS DE ANALISIS PERCEPTUAL DE LA DISFONIA

Article Type: Artículo original

Keywords: Trastornos de la voz; análisis de la voz
Voice disorders; voice analysis

Corresponding Author: Dr. Faustino Nuñez-Batalla, M.D.

Corresponding Author's Institution: Hospital Universitario Central de Asturias

First Author: Faustino Nuñez-Batalla

Order of Authors: Faustino Nuñez-Batalla; Juan Pablo Diaz-Molina; Isabel García-López; Adriana Moreno-Méndez; María Costales-Marcos; Carla Moreno-Galindo; Pablo Martínez-Cambor

Abstract: La evaluación perceptual de la calidad vocal sigue siendo un importante método para evaluar los trastornos vocales. El método GRABS se ha consolidado como una escala frecuentemente utilizada para puntuar la severidad de una disfonía, pero no se ha publicado un protocolo estándar para guiarse. El entrenamiento es importante para alcanzar una buena concordancia en la calificación de sus parámetros entre distintos observadores, sin embargo, las referencias bibliográficas más citadas no describen orientaciones para su uso clínico, muestras a analizar o calibración. Este estudio investigó el efecto de voces patrón y la espectrografía de banda estrecha en el entrenamiento del GRABS de profesionales no expertos. Las voces de 107 pacientes fueron evaluadas por cuatro profesionales no expertos utilizando la escala GRABS en 2 sesiones, primero sin voces patrón ni espectrograma y 6 meses después con voces patrón y espectrograma de banda estrecha. Los resultados muestran que las voces patrón y el espectrograma ayudaron a mejorar la fiabilidad de los parámetros G, B, A y S. La concordancia entre los distintos observadores de acuerdo al estadístico k fue significativamente mayor con la adición de la información espectrográfica para los parámetros B y S. Este estudio demuestra que los profesionales no expertos mejoran significativamente sus puntuaciones tras el entrenamiento con voces patrón y la visualización del espectrograma de banda estrecha.

Suggested Reviewers:

Sr Director:

Le ruego que considere el artículo EL ESPECTROGRAMA DE BANDA ESTRECHA COMO AYUDA PARA EL APRENDIZAJE DEL METODO GRABS DE ANALISIS PERCEPTUAL DE LA DISFONIA para su publicación en el Acta Otorrinolaringologica Española como Artículo Original.

Atentamente,

Dr Faustino Nuñez Batalla

Título: EL ESPECTROGRAMA DE BANDA ESTRECHA COMO AYUDA PARA EL APRENDIZAJE DEL METODO GRABS DE ANALISIS PERCEPTUAL DE LA DISFONIA

Title: THE EFFECT OF ANCHORS AND VISIBLE SPEECH IN THE TRAINING OF THE GRABS METHOD OF PERCEPTUAL RATING OF PATHOLOGICAL VOICES

Autores: Núñez-Batalla F¹, Díaz-Molina JP¹, García-López I², Moreno-Méndez A³, Costales-Marcos M¹, Moreno-Galindo C¹, Martínez-Camblor P⁴.

- (1) Servicio de ORL del Hospital Universitario Central de Asturias.
- (2) Servicio de ORL del Hospital Universitario La Paz (Madrid).
- (3) Fonoaudióloga. Universidad Manuela Beltrán (Bogotá, Colombia).
- (4) CAIBER, Oficina de Investigación Biosanitaria, Oviedo, Asturias

Dirección para correspondencia: Faustino Núñez Batalla. Servicio de ORL. Hospital Universitario Central de Asturias. c/ Celestino Villamil s/n. 33006. Oviedo. España.

Número de Palabras: 4343.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30

RESUMEN: La evaluación perceptual de la calidad vocal sigue siendo un importante método para evaluar los trastornos vocales. El método GRABS se ha consolidado como una escala frecuentemente utilizada para puntuar la severidad de una disfonía, pero no se ha publicado un protocolo estándar para guiarse. El entrenamiento es importante para alcanzar una buena concordancia en la calificación de sus parámetros entre distintos observadores, sin embargo, las referencias bibliográficas más citadas no describen orientaciones para su uso clínico, muestras a analizar o calibración. Este estudio investigó el efecto de voces patrón y la espectrografía de banda estrecha en el entrenamiento del GRABS de profesionales no expertos. Las voces de 107 pacientes fueron evaluadas por cuatro profesionales no expertos utilizando la escala GRABS en 2 sesiones, primero sin voces patrón ni espectrograma y 6 meses después con voces patrón y espectrograma de banda estrecha. Los resultados muestran que las voces patrón y el espectrograma ayudaron a mejorar la fiabilidad de los parámetros G, B, A y S. La concordancia entre los distintos observadores de acuerdo al estadístico k fue significativamente mayor con la adición de la información espectrográfica para los parámetros B y S. Este estudio demuestra que los profesionales no expertos mejoran significativamente sus puntuaciones tras el entrenamiento con voces patrón y la visualización del espectrograma de banda estrecha.

31 Palabras clave: Trastornos de la voz, análisis de la voz

32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

SUMMARY: Perceptual evaluation of voice quality remains a key standard for judgment of vocal impairments. The GRABS method has become a commonly used scale for rating severity of dysphonia, but has no published, standardized protocol to follow. Training is important to reach good interrater agreement for its parameters, however, most often cited references for the GRABS provides no guidelines for clinical administration, speech material or rating calibration. This study investigated the effect of anchors (exemplars of voices) and visible speech (narrowband spectrography) in the training of the GRABS method by naive listeners. Four naive listeners evaluated 107 recorded pathological voices using the GRABS scale in 2 separate sessions, first without visible spectrogram and 6 months later, with anchors and narrowband spectrogram as additional information. The results show that anchors and visible speech helped to improve the reliability G, B, A and S parameters. The interrater agreement according to k statistics was significantly stronger with addition of spectrographic information for rating breathiness and strain. This study found that naive listeners showed significant improvement after training with external anchors (exemplars of voices) and narrowband spectrography.

Key words: Voice disorders, voice analysis

INTRODUCCION

La recomendación del National Center for Voice and Speech (www.ncvs.org) es que la evaluación de la voz patológica debe comenzar con la realización de un análisis espectrográfico para clasificarla con el fin de determinar los métodos de estudio más idóneos para ese caso en particular. Así, una voz con el suficiente grado de periodicidad (tipo 1) debería ser analizada por medio de los parámetros de perturbación a corto plazo (jitter, shimmer, NNE y HNR). Si la voz presenta aperiodicidad, subarmónicos y roturas de voz (tipo 2) no puede ser analizada de forma fiable con los parámetros antes mencionados, una perturbación mayor del 5% nos indica que estamos ante una voz de este tipo, que debe ser estudiada por un método de clasificación perceptual (GRABS) y por un método visual como el espectrograma. Las voces caóticas (tipo 3) solo pueden ser estudiadas por el momento mediante métodos perceptuales, en tanto no se apliquen de forma práctica las dimensiones que estudian los fenómenos caóticos (dimensión fractal, exponente de Lyapunov, etc). Por tanto, la clasificación GRABS es un método imprescindible para la valoración de la voz patológica, dado que es necesario aplicarla en un importante número de pacientes en los cuales las medidas de la perturbación a corto plazo no son fiables o no se pueden calcular¹. Dado que el espectrograma de banda estrecha es la representación gráfica de un sonido, su observación simultánea a la percepción de la disfonía puede ayudar tanto al aprendizaje como a la fiabilidad de un método subjetivo de clasificación de las voces patológicas que es el método GRABS. En este estudio se analiza la ayuda del espectrograma de banda estrecha en el aprendizaje de dicho método.

MATERIAL Y METODOS

Se estudiaron de forma retrospectiva un total de 107 muestras de voz correspondientes a otros tantos pacientes, 29 hombres y 78 mujeres, diagnosticados mediante videoestroboscopia de edema de Reinke.

La señal acústica se registró mediante el uso de la aplicación Voice Assessment del programa Dr Speech 3.0 para Windows 95. El ordenador utilizado fue un PC compatible Pentium-100 con una memoria RAM de 16 Mb. Para la digitalización de la señal vocal se instaló una tarjeta de sonido compatible con Windows de 16-bit de resolución y frecuencias de registro de 44100

1 (Sound Blaster 16). El micrófono es unidireccional dinámico. La frecuencia de muestreo fue de
2 44.100 Hz y se utilizó un micrófono de alta resolución frecuencial que se colocó a 10 cm de la
3 boca del paciente mientras realizaba la fonación de una vocal /e/ a intensidad y tonos
4 cómodos en una cámara sonoamortiguada. De la emisión el ordenador capta 3 segundos. Se
5 siguieron las recomendaciones del National Center for Voice and Speech².
6
7

8 **Análisis Espectrográfico de las Voces**

9
10 El análisis espectrográfico consistió en la realización de un espectrograma de banda estrecha a
11 partir de las voces digitalizadas que fueron entregadas a cada observador. Para ello se utilizó el
12 programa Praat³. Cada participante llevó a cabo el análisis espectrográfico de las voces a
13 estudio en su propio PC. Se configuró el programa Praat para obtener un espectrograma de
14 banda estrecha: tras cargar la voz que se pretende estudiar se abre la ventana "View and edit"
15 para abrir posteriormente la opción "spectrum" de su menú. En "Spectrogram settings" se
16 cambia la cifra de "Window length (s)" de 0.005 por 0.045. Una vez completado este paso ya
17 se observa en la ventana correspondiente el espectrograma de banda estrecha de la voz a
18 estudio. Para mejor observación del trazado se puede pasar la imagen mediante "Paint visible
19 spectrogram" a la ventana "Praat picture" desde donde se puede copiar el espectrograma para
20 imprimirlo.
21
22
23
24
25
26
27
28
29
30

31 **Escalas y patrones de referencia espectrográficos**

32 Se crearon escalas para cada ítem del GRABS (excepto la G) mediante patrones
33 espectrográficos que ilustran los extremos de severidad de las mismas con el fin de que los
34 observadores comparen el espectrograma de la voz que están analizando. En un extremo se
35 observa una voz normal para el parámetro y en el extremo contrario un ejemplo de máxima
36 severidad. El parámetro B se ilustra con la presencia de ruido en el espectrograma siguiendo
37 los grados de Yanagihara⁴, el parámetro R se ilustra con la presencia de subarmónicos⁵, el
38 parámetro A con la ausencia de trazado tanto de ruido como de armónicos en el
39 espectrograma⁶ y el parámetro S con la presencia de señal acústica de mucha intensidad
40 (ruido, armónicos y subarmónicos). En la figura 1A se muestra la escala para el parámetro B y
41 en la figura 1B la escala para el parámetro R.
42
43
44
45
46
47
48
49
50
51
52
53

54 **Análisis Perceptivo de la Voz (GRABS).**

- 55
56
57 1. **Calificación de Referencia:** Mediante la reproducción sonora de cada muestra vocal se
58 calificó conjuntamente y sin ayuda visual por dos profesionales con experiencia en
59 patología vocal utilizando el método GRABS puntuando cada uno de los siguientes
60
61
62
63
64
65

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

ítems de 0 a 3 (0 = normal, 1 = leve , 2 = moderado, 3: severo): G (Grado), el grado global de afectación vocal. R (Roughness, Aspereza), la calidad de la voz relacionada con la impresión de pulsos glóticos irregulares, de un componente de ruido de baja frecuencia, de aspereza o vocal fry. B (Breathiness, voz aérea), la voz relacionada con el ruido originado por las turbulencias creadas por una glotis insuficiente. A (Astenia), la impresión auditiva de debilidad en la fonación espontánea, Voz hipocinética o hipofuncional. S (Strain, tensión vocal), la impresión auditiva de excesivo esfuerzo, de tensión asociada con la fonación espontánea. Este análisis se repitió de forma independiente por los dos observadores 6 meses después de la primera valoración, también sin ayuda visual, con el fin de comparar la conjunta inicial con las dos valoraciones independientes posteriores.

2. **Calificación por parte de los profesionales no entrenados en el sistema GRABS:**

Participaron 4 profesionales no entrenados específicamente en el sistema GRABS, aunque con experiencia en la atención de trastornos vocales (1 especialista en ORL, 1 fonoaudióloga y 2 residentes de otorrinolaringología de 3er año).

- a. Puntuación sin ayuda visual: Escucha de las voces por medio de su reproducción en el ordenador de cada observador sin ninguna pista visual utilizando auriculares.
- b. Puntuación observando el espectrograma: Para utilizar el espectrograma como ayuda visual se eligieron voces patrón para los parámetros R (aspereza), A (astenia), B (aire) y S (tensión). Cada parámetro se ilustra con espectrogramas de banda estrecha de voces con severidad mínima y con severidad máxima, creando una escala analógico visual que ayude a calificar después las voces a estudio (Figura 1A y 1B). Los trazados espectrográficos que se relacionan con cada ítem del GRABS son: la escala de Yanagihara para la puntuación de la voz aérea (B), la presencia de subarmónicos como referencia de presencia de aspereza (R), la falta de señal (no armónicos, no ruido) para calificar la astenia (A) y la señal prominente tanto de armónicos, como de ruido y subarmónicos para calificar la tensión (S). Esta puntuación con la ayuda visual del espectrograma se llevó a cabo 6 meses después de la primera calificación sin espectrograma.

3. **Variables y análisis estadístico:** Las variables continuas se describen mediante media y desviación estándar. Las variables categóricas se describen mediante frecuencias relativas. Para estudiar la diferencia entre las puntuaciones crudas (pertenecientes a

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

cada individuo y la puntuación de referencia) se utiliza el coeficiente de correlación de Spearman y el índice Kappa (permiten ver el grado de concordancia entre las medidas). Para contrastar la igualdad en los resultados antes y después se utilizó el test de McNemar. El coeficiente de correlación de Spearman se utilizó como medida de correlación (la asociación o interdependencia) entre las calificaciones de los profesionales no expertos con la calificación de referencia (si es +1 hay una correlación perfecta positiva, si es -1 hay una correlación perfecta negativa, si es 0 indica ausencia de correlación). El índice Kappa se utilizó para descartar coincidencias por azar (si es +1 hay total acuerdo, si es -1 hay total desacuerdo, si es 0 indica total independencia).

El análisis estadístico se llevó a cabo mediante el programa SPSS versión 15.0 para Windows (SPSS Inc., Chicago, IL).

RESULTADOS

1. **Calificación de Referencia.** Se comparan las puntuaciones GRABS obtenidas de forma conjunta por los dos observadores con las puntuaciones GRABS obtenidas de forma independiente por cada observador 6 meses después. Se alcanza significación estadística ($p < 0.05$) en las pruebas de Ji-cuadrado de Pearson, Razón de verosimilitud, Coeficiente Phi, V de Cramer y Coeficiente de Contingencia en todas las comparaciones. En la tabla I se resumen los valores de p de Ji-cuadrado.
2. **Calificación por profesionales no entrenados:** Los resultados de las calificaciones sin ayuda del espectrograma y con la ayuda visual aportada por el espectrograma se muestran en la tabla II y en la figura 2.

Los parámetros grado de disfonía (G), aspereza (R), astenia (A) y tensión (S) obtuvieron valores promedios inferiores tras la utilización del espectrograma mientras que el parámetro de la voz aérea (B) aumentó. (Figura 2C).
3. **Comparación de los resultados en los profesionales no entrenados con la calificación de referencia:** Se compararon las puntuaciones de los profesionales no entrenados con las calificaciones de referencia utilizando el porcentaje de aciertos totales, el coeficiente κ , y el coeficiente de correlación de Spearman haciendo un promedio de la valoración total de los cuatro evaluadores sin ayuda del espectrograma y tras la utilización del mismo. (Tabla III, Figura 3).
 - a. Los valores que presentaban una mayor correlación con la calificación de referencia sin la utilización del espectrograma fueron la G y la R con valores de Spearman de 0,52 y 0,47 seguidos de la B y la A con valores de 0,39 y 0,37. El parámetro que presentó una correlación más baja fue la S con un valor de 0,34.

1
2
3
4
5
6
7
8
9
Tras la utilización del espectrograma los parámetros G y R permanecieron sin cambios importantes disminuyendo únicamente 0,01 en los valores de correlación de Spearman, la A presentó una discreta mejoría (aumento de 0,03) siendo la mejoría importante en la B en donde se pasó de un valor de 0,39 a un valor de 0,50. Al evaluar la S tras la utilización del espectrograma disminuye de forma moderada la correlación pasando de 0,34 a 0,27.

- 10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
b. El porcentaje de aciertos sin la utilización del espectrograma fue mayor al evaluar el grado de disfonía (G) con un 55,2%, seguido de la astenia (A) con un 48,1%, la aspereza (R) con un 43,7% y el parámetro de tensión (S) con un 39%. El parámetro de la voz aérea (B) fue el que menor número de aciertos presentó con un 38,2%. Tras la utilización del espectrograma el porcentaje de aciertos mejoró en los parámetros G, A, B, S. Esta mejoría fue discreta en los parámetros G y A existiendo una marcada mejoría en los parámetros de la voz aérea (6,5%) y de la tensión (6,8%). Los parámetros en los que hubo una mayor concordancia con la valoración de referencia sin la utilización del espectrograma fueron la G y la R con valores Kappa de 0,28 y 0,22 respectivamente seguidos de la A y la B con valores de 0,18 y 0,14. El parámetro S fue el más diferente con respecto a la valoración de referencia con un valor Kappa de 0,08. Tras la utilización del espectrograma la concordancia mejoró de forma importante en los parámetro de la voz aérea (κ sin espectrograma 0,14 a 0,21 tras la utilización del espectrograma) y en el de tensión (κ sin espectrograma 0,08 a 0,16 tras la utilización del espectrograma). Para saber si estas mejorías eran estadísticamente significativas se realizó un test de Mc Nemar comparando la evaluación sin espectrograma con la calificación de referencia y la valoración tras la utilización del espectrograma con la calificación de referencia obteniendo diferencias estadísticamente significativas tanto para la B ($p=0,001$) como para la S ($p=0,007$).

47 **DISCUSION**

48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
La sensación psicoacústica que provoca escuchar una voz disfónica sigue siendo la principal herramienta para juzgar los trastornos vocales, tanto para el paciente que la sufre como para el clínico que ha de tratarla. Las calificaciones perceptuales de la calidad vocal definen la presencia o ausencia de un trastorno vocal clínicamente, por tanto, existe la necesidad de establecer una forma de llevar a cabo tales análisis perceptuales que tenga un significado clínico y que sea consistente^{7,8}.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

En la actualidad no existe ningún método de calificación de la calidad percibida de la voz que sea claramente mejor que los demás. Todos los esfuerzos encaminados a describir una herramienta de calificación perceptual de la disfonía han chocado con un cúmulo de problemas relacionados con su fiabilidad, utilidad y validez. No obstante, el desarrollo de protocolos-guía y el diseño de métodos para documentar las características auditivo-perceptuales de una calidad vocal anormal han ayudado enormemente en estas tareas al mejorar la comunicación y la consistencia entre los profesionales. En este contexto se han recomendado una serie de principios orientativos para los sistemas de análisis perceptual de la voz: 1) Las dimensiones perceptuales deberían reflejar un conjunto mínimo de parámetros perceptuales de la voz clínicamente relevantes. 2) Los procedimientos y resultados deberían ser obtenidos convenientemente. 3) Los procedimientos y los resultados deberían ser aplicables a un amplio rango de patologías vocales y contextos clínicos. 4) Las calificaciones deberían demostrar que optimizan la fiabilidad intra e interobservador en estudios de validación y 5) Se deben utilizar muestras de voz para utilizarlas como referencia para las calificaciones y para el entrenamiento⁷.

Uno de los métodos más difundidos para la calificación perceptual de la voz como es el método GRABS carece de un protocolo estándar para seguir, de referencias para calibrar las puntuaciones y de un método de entrenamiento con que mejorar la consistencia de la valoración clínica. Los examinadores con experiencia en la calificación perceptual de la voz utilizan por lo general referencias propias para orientarse, sin embargo se ha comprobado que el uso de referencias externas mejora la fiabilidad de las puntuaciones^{9,10}. En este trabajo se analiza la influencia de la referencia visual espectrográfica combinada con la referencia acústica (voces patrón) como método de entrenamiento, con el fin de mejorar la consistencia de la calificación perceptual de la voz utilizando el sistema GRABS.

El principal hallazgo de este estudio es que a excepción del parámetro R (voz áspera), todos los ítems del sistema GRABS mejoraron en la valoración de la coincidencia de las puntuaciones de observadores no entrenados con la de referencia. Los parámetros B y S del sistema GRABS son los que más mejoraron al ayudarse de la visión del espectrograma para su puntuación. Esta mejoría en la coincidencia puede explicarse por las características específicas observadas en los espectrogramas. Estas características son la presencia de ruido que es indicativa de la voz aérea y la presencia de una energía acústica fuerte en la región de las altas frecuencias del espectro asociada a la voz tensa¹.

1 Otra importante observación es el empeoramiento de la coincidencia del parámetro R con la
2 presencia de subarmónicos en el espectrograma, de hecho la calificación sin la ayuda del
3 espectrograma mostraba los valores más altos de coincidencia. Este hecho tiene que ver con
4 que los subarmónicos no son las únicas características del espectrograma que se asocian con
5 una voz áspera, hecho ya observado por varios autores^{1,11,12}. Este hallazgo confirma la noción
6 de que la valoración perceptual no puede ser reemplazada por parámetros acústicos ni por
7 parámetros espectrográficos, al menos en lo que se refiere a la cualidad áspera de una
8 disfonía. En el estudio de Martens¹³ se emplea el espectrograma de banda ancha para
9 determinar si existen irregularidades en los pulsos glóticos que determinen la percepción de
10 una voz áspera. Estas irregularidades son independientes de los valores de perturbación a
11 corto plazo y no se describe una forma objetiva de medirlos.

12 En el presente trabajo se partió de la hipótesis de que los subarmónicos se relacionan con la
13 voz áspera basándose en los estudios previos sobre el tema, en los que se establece que las
14 configuraciones con subarmónicos se relacionan con perturbaciones en la frecuencia que se
15 perciben como voces quebradas y ásperas¹⁴. La percepción de la aspereza depende de
16 características combinadas de perturbación, en las que si los parámetros de perturbación
17 originan una voz con dos o más frecuencias se percibe como diplofonía, pero si se introduce un
18 elemento pulsátil se percibe como una voz quebrada o vocal fry¹⁵. En estudios fisiológicos y
19 anatómicos utilizando grabaciones de imágenes de alta velocidad se observó que existen
20 diferentes frecuencias vibratorias entre ambas cuerdas o diferencias de fase que se repiten
21 cada 6 o 9 ciclos en pacientes con voces descritas como ásperas sin que sea imprescindible la
22 presencia de asimetrías en la rigidez o en la masa de las cuerdas¹⁶. Smith¹⁷ Creó un modelo en
23 el que confirma que la diferente frecuencia vibratoria entre ambas cuerdas que da origen a los
24 subarmónicos depende de unas condiciones muy específicas como cierta presión subglótica,
25 un área glótica suficientemente amplia y una rigidez concreta en las cuerdas. Otros autores
26 confirman también que los subarmónicos dan lugar a la aspereza vocal^{18,19}.

27 Respecto al uso de una vocal sostenida en comparación con el habla normal como muestras
28 grabadas para su reproducción y posterior calificación perceptual es un tema controvertido. El
29 clínico en su relación con el paciente disfónico tiene una constante exposición a su voz hablada
30 y esto sin duda va a influir en la puntuación que otorgará al GRABS, el protocolo básico de
31 evaluación de la voz de la European Laryngological Society (ELS) recomienda realizar la
32 clasificación perceptual de la disfonía durante la anamnesis del paciente²⁰. Se ha demostrado
33 tanto que las vocales sostenidas permiten realizar una clasificación más precisa que si se
34 emplean grabaciones de habla continua²¹, como lo contrario²², aunque sin grandes ventajas

1 del habla sobre la vocal sostenida. Lo que sí parece claro es que una vocal sostenida es un
2 sonido no influenciado por los órganos articuladores y, por tanto, representa de forma más
3 pura el producto de la vibración vocal.
4

5 **CONCLUSIONES**

6 Las referencias acústicas (voces patrón) combinadas con las referencias visuales
7 (espectrograma de banda estrecha) ayudan a mejorar la consistencia del sistema de
8 clasificación perceptual GRABS en todos sus parámetros excepto en la aspereza (R).
9

10 La cualidad aérea de la voz disfónica es el parámetro que mejor se predice por la observación
11 de la presencia de ruido y la desaparición de armónicos en el espectrograma.
12

13 La cualidad áspera de la voz disfónica no se determina únicamente por la presencia de
14 subarmónicos en el espectro.
15

16 El uso de los patrones espectrográficos junto con la percepción de la correspondiente voz que
17 se han empleado en este estudio es una herramienta válida para el entrenamiento en la
18 calificación perceptual de la disfonía mediante el sistema de puntuación GRABS.
19

20 **BIBLIOGRAFIA**

- 21 1. Núñez Batalla F, Corte Santos P, Sequeiros Santiago G, Señaris González B, Suárez
22 Nieto C. Evaluación Perceptual de la Disfonía: Correlación con los Parámetros
23 Acústicos y Fiabilidad. Acta Otorrinolaringológica Española 2004; 55: 282-287.
24
- 25 2. Titze IR. National Center for Voice and Speech. Workshop on Acoustic Voice Analysis.
26 Summary Statement, Denver, 1994.
27
- 28 3. Paul Boersma and David Weenink. Phonetic Sciences, University of Amsterdam,
29 Holanda, <http://www.fon.hum.uva.nl/praat/>
30
- 31 4. Yanagihara N. Significance of harmonic changes and noise components in hoarseness,
32 Journal of Speech and Hearing Research 1967; 30: 531-541.
33
- 34 5. Cavalli L, Hirson A. Diplophonia reappraised. Journal of Voice 1999; 13: 542-556.
35
- 36 6. F. Núñez, P Corte, B Señaris, N Rodriguez C Suárez. Evaluación espectral cuantitativa de
37 la hipofunción vocal. Acta Otorrinolaringológica Española 2004; 55: 327-333.
38
- 39 7. Kempster GB, Gerratt BR, Verdolini Abbott K, Barkmeier-Kraemer J, Hillman RE.
40 Consensus auditory-perceptual evaluation of voice. Am J Speech Lang Pathol. 2009;
41 18:124-32.
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
8. Carding PN, Carlson E, Epstein R, Mathieson L, Shewell C. Formal perceptual evaluation of voice quality in the United Kingdom. *Logopedics, Phoniatrics, Vocology* 2000; 25: 133-138.
9. Chan KM, Yiu EM. The effect of anchors and training on the reliability of perceptual voice evaluation. *J Speech Lang Hear Res* 2002; 45: 111-26.
10. Yiu EM, Chan KM, Mok RS. Reliability and confidence in using a paired comparison paradigm in perceptual voice quality evaluation. *Clin Linguist Phon* 2007; 21: 129-45.
11. Núñez Batalla F, Suárez Nieto C, Muñoz Pinto C, Baragaño Rio L, Alvarez Zapico MJ. Estudio sonográfico de la disfonía: Subarmónicos. *Acta Otorrinolaringológica Española* 2000; 51: 23-27.
12. Omori K, Kojima H, Kakani R, Slavik DH, Blaugrund SM. Acoustic Characteristics of Rough Voice: Subharmonics. *J Voice* 1997; 11: 40-47.
13. Martens J, Versnel H, Dejonckere PH. The effect of visible speech in the perceptual rating of pathological voices. *Arch Otolaryngol Head Neck Surg* 2007; 133: 178-185.
14. Schereibweiss-Merin D, Terrio L. Acoustic analysis of diplophonia: A case study. *Percept Mot Skills* 1986; 63: 755-765.
15. Imaizumi S, Gauffin J. Acoustic and perceptual characteristics of pathological voices: rough, creak, fry and diplophonia. *Ann Bull Res Inst Logop Phoniatr* 1991; 25: 109-119.
16. Kiritani S, Hirose H, Imagawa H. High-speed digital image analysis of vocal fold vibration in diplophonia. *Speech Commun* 1993; 139: 23-32.
17. Smith M, Berke G, Gerratt B, Kreiman J. Laryngeal paralysis: Theoretical considerations and effects on laryngeal vibration. *J Speech Hear Res* 1992; 35: 545-554.
18. Imaizumi S. Acoustic measures of roughness in pathological voice. *J Phonet* 1986; 14: 457-462.
19. Dejonckere P, Obbens C, de Moor GM, Wieneke GH. Perceptual evaluation of dysphonia: reliability and relevance. *Folia Phoniatr (Basel)* 1993; 45: 76-83.
20. Dejonckere PH, Bradley P, Clemente P, Cornut G, Crevier-Buchman L, Friedrich G, et al. A basic protocol for functional assessment of voice pathology, especially for investigating the efficacy of (phonosurgical) treatments and evaluating new assessment techniques. Guideline elaborated by the Committee on Phoniatrics of the European Laryngological Society (ELS). *Eur Arch Otorhinolaryngol* 2001;258:77-82.
21. Parsa V, Jamieson DG. Acoustic discrimination of pathological voice: Sustained vowels versus continuous speech. *Journal of Speech, Language and Hearing Research* 2001; 44: 327-339.
22. Bele IV. Reliability in perceptual analysis of voice quality. *Journal of Voice* 2005; 19: 555-573.

TABLAS

	Observador 1	Observador 2
G conjunto vs G individual	0.000	0.000
R conjunto vs R individual	0.000	0.000
A conjunto vs A individual	0.000	0.000
B conjunto vs B individual	0.000	0.000
S conjunto vs S individual	0.002	0.003

Tabla I. Valores de p (Ji-cuadrado de Pearson) en el estudio que compara las puntuaciones de ambos observadores de forma conjunta con las puntuaciones de los observadores de forma individual seis meses después.

Parámetro	Sin espectrograma	Con espectrograma
G	1,98 (0,73)	1,89(0,75)
R	1,42(0,90)	1,18(1,02)
A	0,69(0,93)	0,40(0,72)
B	1,00(0,87)	1,35(0,93)
S	1,09(0,86)	0,78(0,81)

Tabla II. Influencia del espectrograma en la evaluación perceptual de voces patológicas según el método GRABS. Valores reportados como la media de las puntuaciones de cuatro profesionales no entrenados (Desviación típica).

	G			R			A			B			S		
	%	S	K	%	S	K	%	S	K	%	S	K	%	S	K
Evaluación Pre	55.2	0.52	0.28	43,7	0.47	0.22	48.1	0.37	0.18	38.2	0.39	0.14	39.0	0.34	0.08
Evaluación Post	56,4	0.51	0.30	42.3	0.46	0.21	49,3	0.40	0.14	44.7	0.50	0.21	45,8	0.27	0.16

Tabla III. Influencia del espectrograma en la evaluación perceptual de voces patológicas según el método GRABS. Valores reportados como los valores promedio de la evaluación de cuatro profesionales no expertos en cuanto al porcentaje de aciertos (%), Coeficiente de correlación de Spearman (S) y Coeficiente Kappa (K).

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Figura 1 A. Escala analógico-visual para el parámetro B en mujeres, con tres grados de severidad, donde se pueden oír las correspondientes muestras vocales y observar simultáneamente los espectrogramas de banda estrecha.

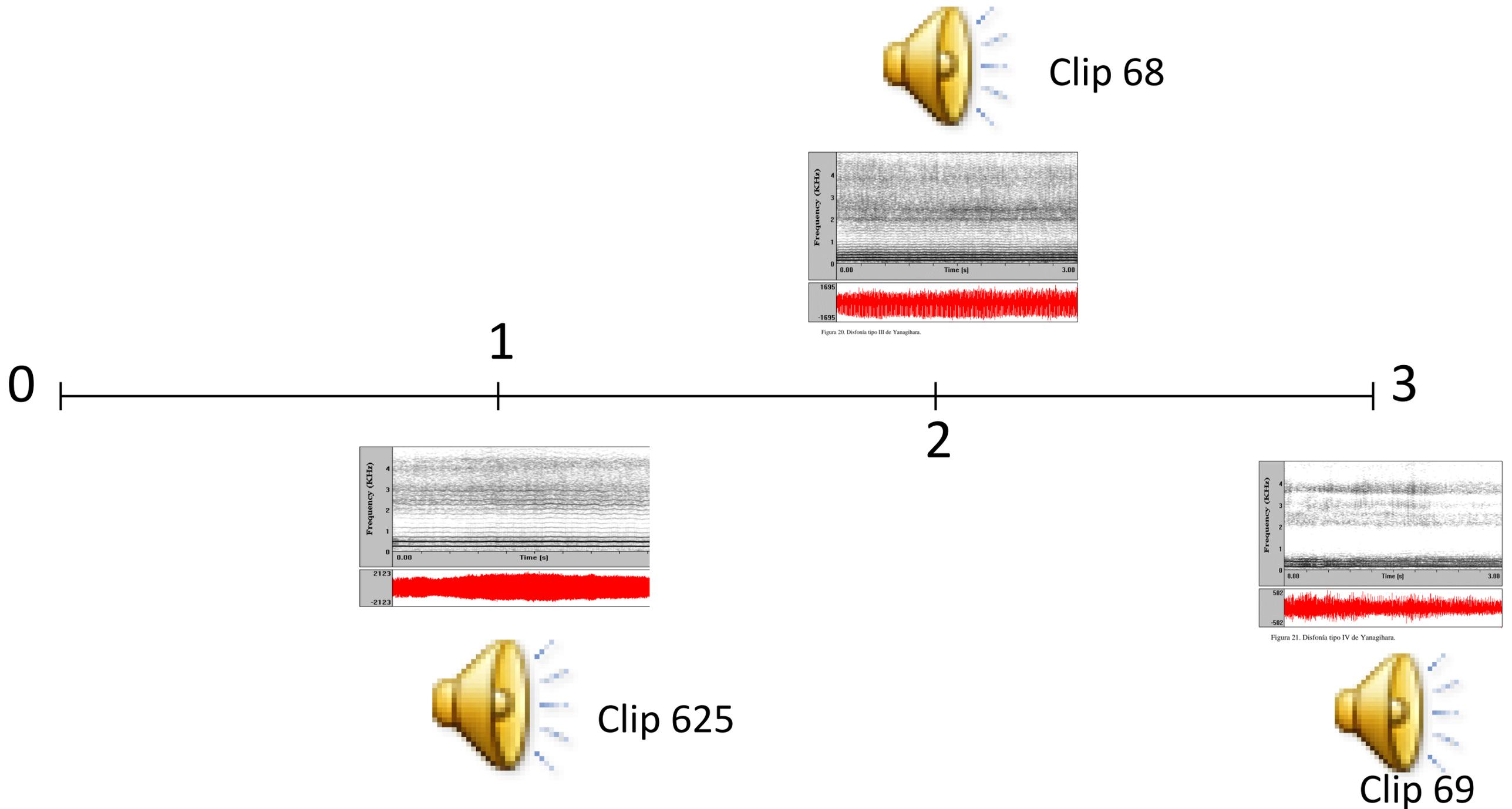
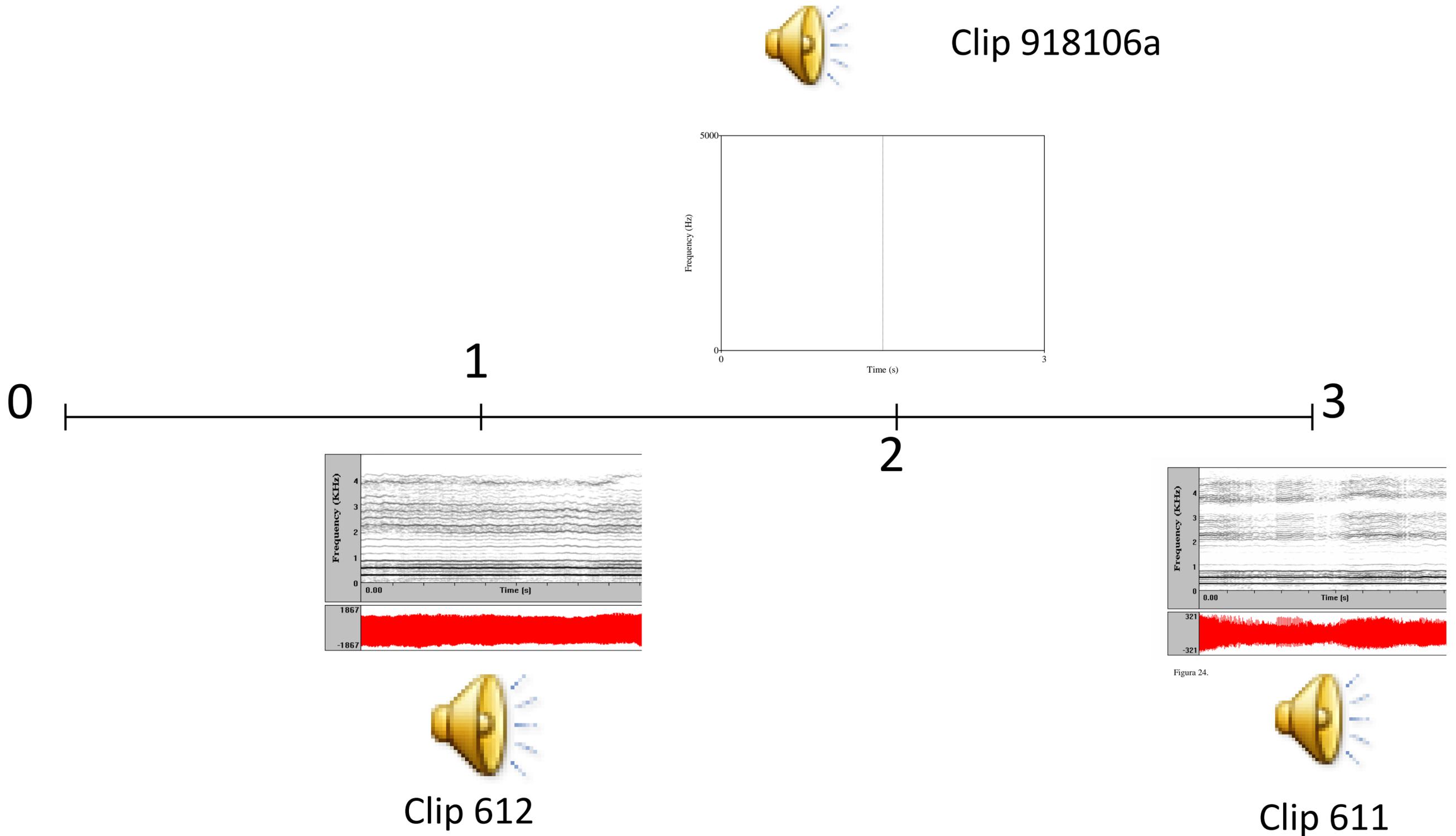


Figura 1 B. Escala analógico-visual para el parámetro R en mujeres, con tres grados de severidad, donde se pueden oír las correspondientes muestras vocales y observar simultáneamente los espectrogramas de banda estrecha.



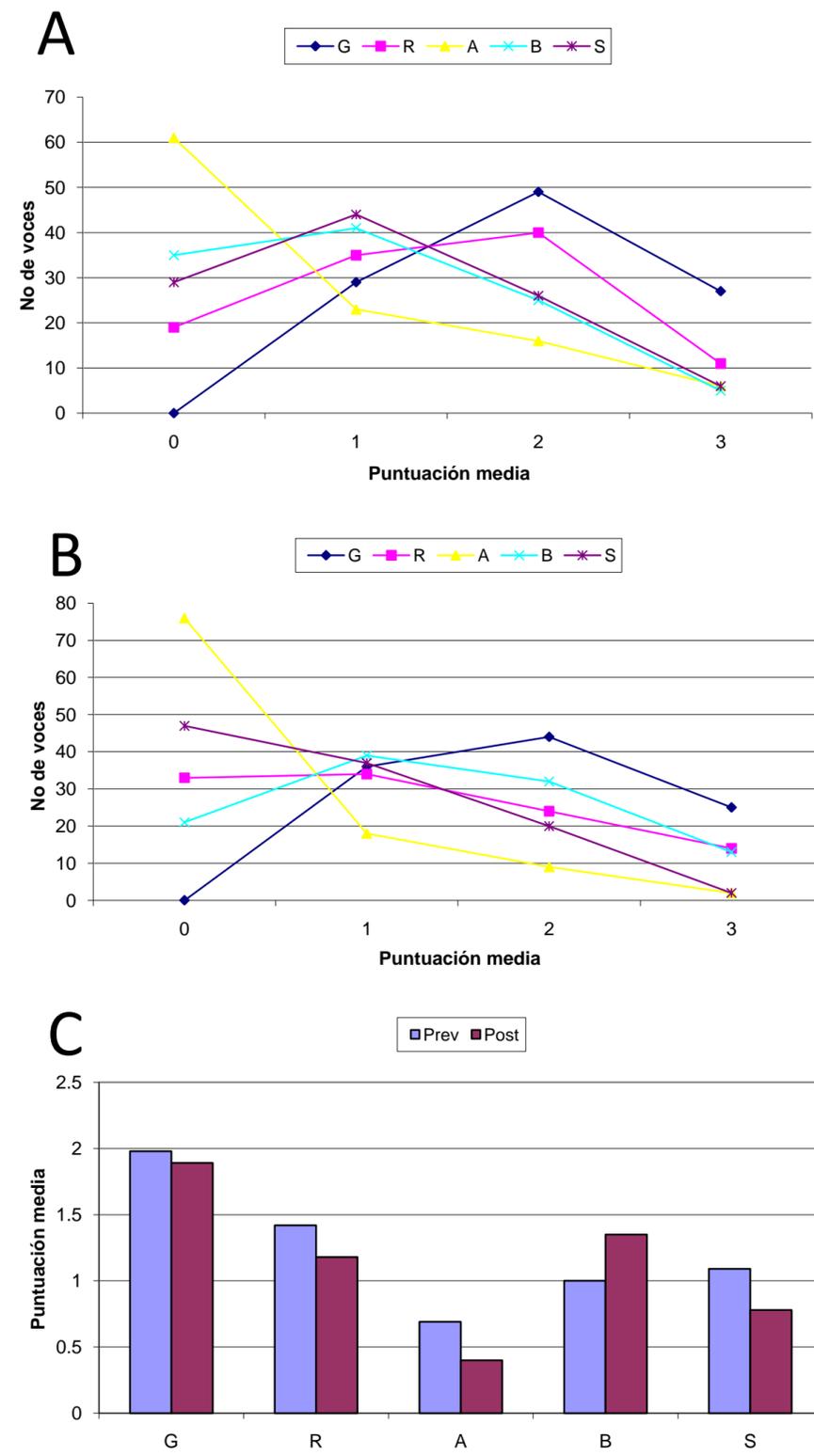


Figura 2: Distribución de las puntuaciones promedio del GRABS (grado, aspereza, astenia, voz aérea, y tensión) sin la ayuda del espectrograma (A) y con la ayuda del espectrograma (B) en las 107 voces evaluadas por 4 profesionales no expertos. Representación de las diferencias encontradas en la evaluación de las voces con y sin ayuda del espectrograma (C).

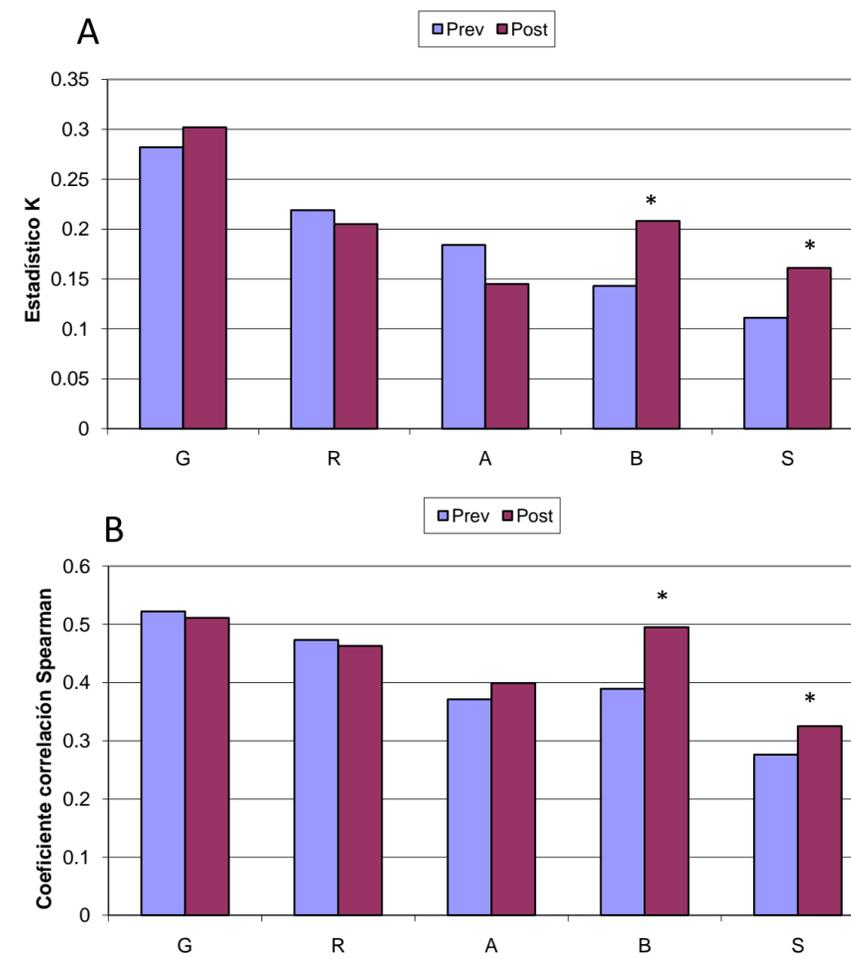


Figura 3: Representación gráfica de las comparaciones entre la valoración por profesionales no expertos y la valoración de referencia con y sin ayuda del espectrograma. A: Representación de los resultados del valor Kappa. B: Representación de los resultados del coeficiente de correlación de Spearman. Marcado con asterisco los resultados que alcanzaron la significancia estadística mediante la aplicación del test de Mc Nemar.