

----- TABLAS DE PRUEBAS NPAR TESTS Y CUADRADO-----  
 ESTUDIO DE LAS DISTRIBUCIONES DE FRECUENCIA DE LOS 2.800 A LOS 10.000 Hz.  
 (Análisis realizado a partir de "Bandas Críticas")

Comparación de las D. de Fr. de cada locutor respecto a la equiprobabilidad.

-----  
 L2/Equip L4/Equip L3/Equip L5/Equip

0,008 0,051 0,005 0,000

Comparación de las D. de Fr. entre locutores distintos.

-----  
 L2/L4 L2/L3 L2/L5 L4/L3 L4/L5 L3/L5

0,000 0,000 0,000 0,000 0,000 0,000

Comparación de las D. de Fr. entre distintas versiones de un mismo locutor.

-----  
 V12/V22 V14/V24 V13/V23 V15/V25

0,007 0,108 0,164 0,001

Comparación de las D. de Fr. entre distintos subgrupos de una misma versión.  
 (Pruebas alternativas a V12/V22 y V15/V25).

-----  
 V12(1)/V12(2): 0,214 V13(1)/V13(2): 0,967  
 V22(1)/V22(2): 0,285 V23(1)/V23(2): 0,900  
 V14(1)/V14(2): 0,082 V15(1)/V15(2): 0,224  
 V24(1)/V24(2): 0,987 V25(1)/V25(2): 0,859

Comparación de las D. de Fr. entre versiones diferentes de distintos locutores.

-----  
 V14/V12:0,000 V24/V12:0,000 V12/V13:0,007 V22/V15:0,021  
 V14/V22:0,035 V24/V22:0,008 V12/V23:0,035 V22/V25:0,021  
 V14/V13:0,001 V24/V13:0,000 V12/V15:0,000 V13/V15:0,000  
 V14/V23:0,210 V24/V23:0,005 V12/V25:0,000 V13/V25:0,000  
 V14/V15:0,000 V24/V15:0,007 V22/V13:0,000 V23/V15:0,000  
 V14/V25:0,000 V24/V25:0,007 V22/V23:0,001 V23/V25:0,000

V.13(2), ... etcétera, fue la misma que explicamos ya más arriba en el apartado 5.4.1.1.

Se realizó pues una nueva serie de tests JI CUADRADO, esta vez comparando entre sí las distribuciones de formantes de cada pareja de subgrupos obtenida al fragmentar las versiones anteriores. Exponemos a continuación los resultados de esa prueba alternativa:

V.12(1)/V.12(2) : 0,214	V.13(1)/V.13(2) : 0,967
V.22(1)/V.22(2) : 0,285	V.23(1)/V.23(2) : 0,900
V.14(1)/V.14(2) : 0,082	V.15(1)/V.15(2) : 0,224
V.24(1)/V.24(2) : 0,987	V.25(1)/V.25(2) : 0,859

Como puede comprobar el lector el cambio ha sido sustancial, ahora todas las pruebas indican que las muestras comparadas no tienen una distribución de formantes significativamente distinta y, por tanto, podemos afirmar que pertenecen a la misma población; o lo que es lo mismo, parece ser que la distribución de los formantes altos de los sonidos vocálicos no solo se organiza en función de la especificidad tímbrica individual de cada voz, sino que también responden a las modificaciones sonoras que es capaz de imprimir un locutor en su voz cuando interpreta acústicamente personajes diferentes.

Como el lector puede comprobar, todos los tests excepto uno indican que las muestras de sonidos vocálicos que se han comparado tienen una distribución de los formantes significativamente distinta entre sí. O, lo que es lo mismo, que en el 95,8% de las pruebas realizadas se comprueba que la distribución de los formantes, efectivamente, depende también de las modificaciones sonoras que es capaz de imprimir un locutor en su voz cuando interpreta acústicamente personajes diferentes.

#### PRUEBA D.

Apoyándonos en el análisis estadístico que acabamos de exponer podemos afirmar que para la las voces estudiadas se cumple la segunda parte de la hipótesis. Así, puesto que ya hemos comprobado que la distribución estadística de los formantes en la Zona Alta del espectro es un instrumento capaz de discriminar unas voces de otras, podemos proceder a al análisis de la primera parte de nuestra hipótesis. Es decir, a estudiar si los rasgos acústicos lingüísticos realmente no influyen en la Zona Alta del espectro.

Si las distintas posiciones del resonador bucal para la construcción de sonidos vocálicos no tuviese ninguna influencia en la ubicación de los formantes de la Zona Alta, no deberíamos encontrar diferencias significativas al comparar distintos sonidos vocálicos de una misma versión del texto.

Esta comprobación se realizó comparando mediante tests JI CUADRADO la distribución de los formantes altos de cada una de las vocales analizadas de una determinada versión con la distribución de formantes del resto de las vocales de esa misma versión. Y esa comparación se hizo, uno a uno, para todos los sonidos vocálicos. Así, tras seleccionar los datos de una versión concreta, cada prueba comparaba la distribución resultante de los datos de una vocal (por ejemplo la "A") con la distribución resultante de los datos de todas las vocales restantes de esa versión (con la distr. conjunta de "E", "I", "O" y "U").

Finalmente, si es cierto que los rasgos lingüísticos no influyen en la distribución de los formantes altos, debería cumplirse también que la distribución de formantes de la "A" de una versión determinada del locutor "X" fuese significativamente distinta de la distribución de formantes de las "A" de cualquier versión de cualquier otro locutor "Y". Obviamente, si la hipótesis es cierta, esta condición la cumplirían, uno a uno, todos los sonidos vocálicos de las dos versiones de cada locutor. Resumiendo, si los rasgos fonéticos determinan el espectro de la Zona Alta, la distribución de formantes de las "A" de V.12 no tendría que ser significativamente distinta de la distribución de las "A" del Loc.24, ni del Loc.14, ni del Loc.13, ni del Loc.23, ...etcétera; y viceversa.

Para verificar si se cumplía esta última condición se comparó la distribución de formantes de cada uno de los fonemas de las dos versiones de cada locutor, con la distribución de formantes que resultaba de todos los datos sobre ese mismo fonema para cada uno de los locutores restantes, mediante los correspondientes tests JI CUADRADO.

En las tablas de la página siguiente se exponen los resultados obtenidos al realizar esta dos últimas series de tests JI CUADRADO.

## ----- TABLAS DE PRUEBAS NPAR TESTS JI CUADRADO-----

ESTUDIO DE LAS DISTRIBUCIONES DE FRECUENCIA DE 2.900 A 10.000 Hz. CON LOS DATOS FRAGMENTADOS EN FUNCION DE LOS FONEMAS.

(Análisis realizado a partir de "Bandas Críticas")

Comparación de la D. de Fr. de cada uno de los fonemas de una determinada versión, con todos los datos restantes de esa misma versión.

	a/V-(a)	e/V-(e)	i/V-(i)	o/V-(o)	u/V-(u)
V12	0,374	0,019	0,929	-	-
V22	0,026	0,262	0,091	0,196	-
V14	0,414	0,1	0,214	0,119	-
V24	0,482	0,004	0,216	0,652	0,173
V13	0,000	0,364	0,837	0,429	-
V23	0,467	0,07	0,000	0,005	-
V15	0,000	-	-	0,000	0,000
V25	-	0,001	0,005	0,019	-

Comparación de la D. de Fr. de cada uno de los fonemas de las dos versiones de determinado locutor, con el total de datos de ese mismo fonema para los restantes locutores.

	A	E	I	O	U
V12/L4	0,000	0,006	0,028	-	-
V12/L3	0,000	0,000	0,011	-	-
V12/L5	0,000	0,000	0,011	-	-
V22/L4	0,003	0,071	0,203	0,199	-
V22/L3	0,000	0,104	0,181	0,001	-
V22/L5	0,000	0,033	0,979	0,006	-
V14/L2	0,01	0,457	0,071	0,112	-
V14/L3	0,021	0,256	0,658	0,314	-
V14/L5	0,000	0,000	0,008	0,004	-
V24/L2	0,001	0,002	0,129	0,016	-
V24/L3	0,003	0,063	0,449	0,1	-
V24/L5	0,02	0,000	0,141	0,055	0,065
V13/L2	0,002	0,004	0,396	0,002	-
V13/L4	0,000	0,029	0,137	0,004	-
V13/L5	0,000	0,000	0,027	0,000	-
V23/L2	0,001	0,709	0,003	0,013	-
V23/L4	0,116	0,27	0,817	0,017	-
V23/L5	0,000	0,001	0,003	0,03	-
V15/L2	0,000	-	-	0,000	-
V15/L3	0,000	-	-	0,000	-
V15/L4	0,000	-	-	0,000	-
V25/L2	-	0,003	0,39	0,001	-
V25/L3	-	0,000	0,546	0,001	-
V25/L4	-	0,000	0,073	0,000	-
	95,2%	71,4%	38,1%	80,9%	-

En la tabla de la parte superior se muestran los resultados de comparar entre sí distintos sonidos vocálicos de una misma versión. En la parte izquierda de la tabla (eje de ordenadas) se indica con los datos de qué versión se ha hecho la comparación, y en la parte superior (eje de abscisas) cada una de las anotaciones indica qué dos conjuntos de datos de esa versión han sido comparados entre sí. Así, a/V-(a), por ejemplo, expresa que en esa columna están comparados sistemáticamente los datos de la "A" con los todos los demás datos, excluyendo los de "A", de la versión correspondiente. Los guiones representan a las pruebas que el procedimiento NPAR-TESTS no ha podido realizar por insuficiencia de datos.

Si el lector revisa los resultados de la primera tabla podrá comprobar que alrededor del 50% de las pruebas se ajustan a nuestra hipótesis y el restante 50% no. Lo que significa que sólo en la mitad de los casos que hemos comparado sonidos vocálicos diferente estos han mostrado distribuciones de formantes similares. Ya que con estos resultados la variabilidad acústica podría tanto justificar la hipótesis propuesta como la hipótesis alternativa, tendrá que ser la siguiente serie de pruebas la que nos oriente en la toma de una decisión al respecto.

En la segunda tabla de la hoja anterior se presentan los resultados de comparar uno a uno los sonidos vocálicos de las dos versiones de cada locutor con cada uno de los

sonidos vocálicos del resto de los locutores. En la parte superior de la tabla (eje de abscisas) se indica cual es la vocal cuyos datos han sido seleccionados para realizar los tests, y en la parte izquierda de la tabla (eje de ordenadas) se indican los grupos de datos sobre esa vocal que se han comparado entre sí.

Como observará el lector, al comparar la distribución de formantes las "A" de V.12 con la distribución de todas las "A" del locutor 3, luego con las "A" del locutor 4, después con las "A" del locutor 5,...,etc. los tests indican que estas distribuciones son significativamente distintas entre sí en el 71% de las pruebas realizadas. Es decir, que la distribución de los formantes altos de las "A" de la primera versión del locutor 2 no tiene nada que ver con la distribución de los formantes de las "A" del locutor 3, ni con la de los formantes de las "A" del locutor 4, ni con la de los "A" del locutor 5,...,etc.

Así pues, el resultado de esta última serie de pruebas indica con claridad que la posición del resonador bucal al construir los fonemas vocálicos no influye en la distribución de los formantes en la Zona Alta del espectro, y, por tanto, que se cumple la primera parte de nuestra hipótesis, en la que se afirma que, a partir del tercero, los formantes del espectro frecuencial de los sonidos vocálicos no se organizan en función de los rasgos acústicos lingüísticos, sino que dependen fundamentalmente de



características sonoras individuales del locutor. No obstante, la distribución de los resultados indica también que nuestra hipótesis no se cumple para todos los sonidos vocálicos en el mismo grado. En la parte inferior de la segunda tabla de la página anterior se exponen el porcentaje de pruebas en las que cada uno de los sonidos estudiados cumple la hipótesis. Si el lector observa estos porcentajes, notará que hay unas diferencias sustanciales entre unos y otros, las diferencias reflejan que unas vocales se ajustan mucho más que otras a la hipótesis.

Intentemos ahora buscar razones que expliquen por qué una unas vocales se ajustan mejor que otras a la hipótesis.

Teniendo en cuenta que la posición del resonador bucal para construir el fonema "A" es la más abierta y la que menos obstrucciones opone al sonido de la laringe, y, contrariamente, que para construir la "I" la posición del resonador bucal es mucho más obstructiva que para la "A", la "E", o la "O", y observando como el porcentaje de cumplimiento de nuestra hipótesis se ajusta al grado de obstrucción que impone a la voz la posición de la boca, no resulta nada descabellado pensar que, si bien la posición del resonador bucal al construir los fonemas no determina la posición de los formantes altos, sí que actúa a modo de filtro atenuador que destruye más información cuanto más cerrado está.

Aun es posible definir mejor esta teoría. Si ordenamos de forma decreciente los cuatro sonidos vocálicos sobre los que tenemos información según el grado de cumplimiento de nuestra hipótesis:

1) A: 95,2%

2) O: 80,9%

3) E: 71,4%

4) I: 38,1%

podemos comprobar como este orden se ajusta perfectamente al orden que obtendríamos organizando estas cuatro vocales según la distancia que separa el dorso de la lengua del paladar medio en su realización articulatoria, de modo que, ya de una forma más concreta, parece coherente formular la siguiente teoría:

la especificidad acústica del timbre individual de la voz resulta tanto más clara en un análisis espectral de los sonidos vocálicos: a) cuanto más abierto esté el resonador bucal, b) cuanto mayor sea la distancia entre el dorso de la lengua y el paladar medio.

Esta teoría que hemos esbozado podría explicar el diferente grado de cumplimiento de la hipótesis según el fonema pronunciado por el locutor; no obstante, la conclusión que nos parece más interesante sobre esa localización de distintos grados de cumplimiento de la

hipótesis para cada vocal, a nuestro modo de ver, es que puede ser un instrumento de orientación muy eficaz en la elección de muestras para cualquier investigación que trabaje sobre la información acústica de los formantes altos.

#### 6.4.2.4. Conclusiones.

Obviamente, nuestra primera y esencial conclusión de es el cumplimiento de las dos partes de la hipótesis propuesta. No obstante, esta confirmación no solo implica la validez de dos afirmaciones puntuales sino que comporta la posibilidad de desarrollar los primeros pasos de una teoría sobre el análisis del timbre personal de cada locutor desde una base suficientemente sólida.

Intentaremos ordenar los resultados puntuales a los que hemos llegado en esta parte de la investigación de modo que esbocen por sí mismos la teoría que propondremos más adelante.

La confirmación de nuestra hipótesis implica que los dos primeros formantes de los sonidos vocálicos se sitúan en el espectro acústico siguiendo una lógica distinta de la del resto de los formantes. Mientras los dos primeros se organizan en función de la construcción de las formas

acústicas vocálicas, a partir del tercero se organizan dependiendo del timbre individual de cada locutor, es decir a partir de algunas de las formas acústicas esenciales que nos hacen distinguir auditivamente a un locutor de otro.

La asunción de que los dos primeros formantes de los sonidos vocálicos transportan una información acústica distinta al resto de los formantes supone la necesidad de dividir metodológicamente el espectro acústico en dos partes que deben ser estudiadas de forma distinta. Esta división, establecida en función altura máxima en Hz que suele alcanzar el segundo formante vocálico, es la siguiente:

ZONA BAJA: de 0 a 2.800 Hz.

ZONA ALTA: de 2.800 a 10.000 Hz

(Que hayamos limitado la Zona Alta a los 10.000 Hz, cuando la percepción del oído humano alcanza a percibir frecuencias entre los 15.000 y los 20.000 Hz, responde solo a la razón de que esta zona ha sido investigada en nuestro estudio y el resto -de 10.000 a 20.000- no)

La confirmación de la hipótesis, indica también que la situación de los formantes en la Zona Alta del espectro puede y debe ser estudiada estadísticamente con métodos que se ajusten a las leyes de la percepción auditiva. Y, concretamente, demuestra que la utilización del método de la

distribución estadística de formantes en las bandas críticas es adecuado y eficaz para este estudio.

No obstante, el tipo de vocal que el locutor pronuncia en cada momento determina la eficacia del análisis de la parte alta del espectro. El resonador bucal parece actuar en la articulación de ciertas vocales como un leve filtro atenuador de toda la zona alta que disminuye tanto más su influencia:

- a) Cuanto más abierto esté el resonador bucal.
- b) Cuanto mayor sea la distancia entre el paladar medio y el dorso de la lengua.

Finalmente, el proceso de comprobación de la hipótesis nos ha llevado también a la conclusión de que el locutor es capaz de influir sobre la distribución de los formantes de su voz en la parte alta del espectro, y esto ocurre cuando intenta construir una personalidad ficticia alterando el sonido de su voz.

**7. CONFRONTACION ENTRE LOS RESULTADOS DE LOS TESTS DE  
PERCEPCION Y LOS PARAMETROS ACUSTICOS DISCRIMINANTES.**

---

## 7.1. LOGICA DE LA LOCALIZACION DE RELACIONES ENTRE EL ANALISIS ELECTROACUSTICO Y LOS TESTS DE PERCEPCION.

Los gráficos de medias de rangos que habíamos obtenido en la primera parte del experimento eran el punto de partida para encontrar modelos de relación entre la imagen que genera una voz en la mente del radioyente y el espectro acústico de dicha voz. Si conseguíamos encontrar relaciones claras entre la variación de las formas espectrales y la variación de los gráficos de medias de rangos, significaría que habíamos localizado algunos de los rasgos acústicos desencadenantes de la imagen auditiva generada por el sonido de una voz.

Para poder comparar los resultados de los tests de percepción con los resultados del análisis electroacústico de las voces, necesitábamos hacer un nuevo tratamiento de los datos obtenidos en el análisis sonoro que nos permitiese extraer configuraciones espectrales simplificadas, representativas de cada una de las voces estudiadas y, a la vez, manejables con comodidad. La forma ideal de trabajar conjuntamente con todos los datos del análisis sería elaborando gráficos que reconstruyesen por aproximación y de una forma simplificada el espectro medio de cada una de las voces estudiadas.

Así, el primer problema que se nos planteaba en esta última etapa del experimento era el tratamiento de los datos

de forma que nos permitieran construir gráficos representativos de cada voz.

La forma lógica de tratar los datos del análisis espectral era calculando medias de cada uno de los parámetros que, en la segunda parte del experimento, habíamos descubierto que actuaban de forma discriminante. Naturalmente, ajustándonos en este nuevo tratamiento de los datos a la teoría que habíamos propuesto en nuestras últimas conclusiones. Se trataba, entonces, de:

a) extraer medias de las distancias entre el primer y el segundo formante ( $F_2 - F_1$ ), vocal a vocal, para cada versión estudiada;

b) decidir, a partir del análisis de la distribución de los formantes en las bandas críticas, cuales eran las bandas definitorias de cada voz y calcular los valores medios de estas bandas para poder configurarlas como formantes representativos.

También se decidió introducir en los gráficos como parámetros acústicos discriminantes la  $F_0$  (frecuencia fundamental) de cada una de las voces y las intensidades de  $F_0$ ,  $F_1$  y  $F_2$ , así como la intensidad media de las bandas críticas seleccionadas como representativas. La razón que justifica introducir la frecuencia fundamental y las intensidades en nuestra análisis comparativo sin haber



realizado previamente una investigación sobre ellas es, obviamente, la existencia de estudios previos que demuestran y asumen su valor discriminante en el timbre individual de la voz (CASACUBERTA,1987); (GIMENEZ,1986); (LAVER,1980); (QUILIS,1981); (MARTI,1984).

#### 7.1.1. Extracción de medias de los parámetros discriminantes.

Para calcular las medias de todos los parámetros que nos interesaban se utilizó el procedimiento FRECUENCIAS VARIABLES del paquete estadístico "SPSS", pidiéndole al programa los estadísticos: MEDIA, y DESVIACION ESTANDAR de una forma sistemática.

En las dos páginas siguientes se presentan los valores medios la desviación estandar de: la frecuencia fundamental (FO), la frecuencia del primer formante (F1) para cada vocal, y la frecuencia y la intensidad conjunta (de todas las vocales conjuntamente) del primer y el segundo formantes (F1) y (F2). Como puede ver el lector en las tablas, estos datos se han ordenado en ejes de coordenadas, colocando en la columna de la izquierda (eje de ordenadas) la versión sonora a la que corresponde cada fila de datos, y en la fila superior (eje de abscisas) el parámetro al que corresponde cada columna de datos (frecuencia en Hz, Intensidad en Db o