

Análisis acústico y articulatorio del vocalismo valenciano

Luis J. Faus y Amando García
Laboratorio de Acústica
Departamento de Física Aplicada
Universidad de Valencia

Resumen

En el presente trabajo se han analizado las vocales del valenciano desde el punto de vista acústico y articulatorio. Desde la perspectiva de la fonética acústica, hemos evaluado los parámetros físicos que definen la parte estacionaria de los siete sonidos vocálicos del valenciano (a, e, è, i, o, ò, u). Concretamente, se ha medido la frecuencia fundamental con que vibran los pliegues vocálicos cuando se articulan dichas vocales, así como la frecuencia central, el ancho de banda y la intensidad relativa de las cuatro primeras resonancias del tracto vocálico para las diferentes vocales. Por lo que respecta a la descripción articulatoria, hemos obtenido las configuraciones anatómicas con que se articulan cada una de las vocales estudiadas.

1. Introducción

La rápida evolución que ha experimentado la electrónica digital a lo largo de estos últimos años ha permitido avances espectaculares en los procesos de síntesis y reconocimiento de la voz humana por parte de las máquinas. Sin embargo, la realización de estos procesos es imposible sin la existencia de un estudio acústico previo muy detallado del idioma en cuestión. La necesidad de este requisito ha motivado un aumento muy notable de los estudios fonéticos tanto en los

apartados acústicos como fisiológicos de los idiomas más importantes. A lo largo de estas últimas décadas se han realizado estudios muy completos del inglés, francés, ruso, japonés, etc. (1) (2). En particular, por lo que se refiere a la acústica del catalán, cabe destacar los estudios llevados a cabo por R. Cerdá (3) y J. Martí (4).

En los estudios pioneros de R. Cerdá se registraron las frecuencias centrales de los dos primeros formantes de las vocales catalanas pronunciadas por dos locutores diferentes. El trabajo realizado por J. Martí se basó en el algoritmo de la Transformada Rápida de Fourier (FFT) y estaba orientado a la síntesis de este idioma por ordenador.

Nuestra intención es repetir este último estudio, pero con un enfoque mucho más general y refiriéndolo a las vocales del valenciano. En el presente trabajo presentamos los resultados más importantes obtenidos en dicha investigación.

2. El proceso de fonación

La teoría acústica de la producción del habla describe a este proceso como la respuesta de un sistema de filtros sobre una fuente de sonido. Simbólicamente, si consideramos que F representa la fuente sonora (cuerdas vocales) y T la función de transferencia (tracto vocal), el producto F.T representará el sonido resultante.

La perturbación sonora está ori-

ginada por la acción de una estructura articulatoria que, según las condiciones, estrecha el paso de una corriente de aire que procede de los pulmones generando un sonido más o menos estacionario, frena momentáneamente dicho paso y produce un sonido transitorio cuando se libera la corriente, o, en el caso de sonidos periódicos, modula de forma cuasi periódica el flujo de aire en cuestión.

Por su parte, el tracto vocal, como sistema de filtros o cavidades resonantes, actúa sobre el sonido producido por la fuente mencionada, confiriéndole una determinada cualidad fonética.

En síntesis, los sonidos del habla son el resultado de la excitación acústica del tracto vocal por la acción de una o más fuentes. Recordemos que nuestro sistema fonatorio incluye los pulmones, la tráquea, la laringe, la faringe y las cavidades oral y nasal. Dentro de la laringe se encuentran los pliegues o cuerdas vocales que regulan el paso de la citada corriente de aire mediante la abertura y cierre de la glotis. La zona comprendida entre la laringe (glotis) y los labios constituye el tracto vocal propiamente dicho.

3. Parámetros característicos de las vocales

El proceso de fonación de una vocal en cualquier idioma se produce

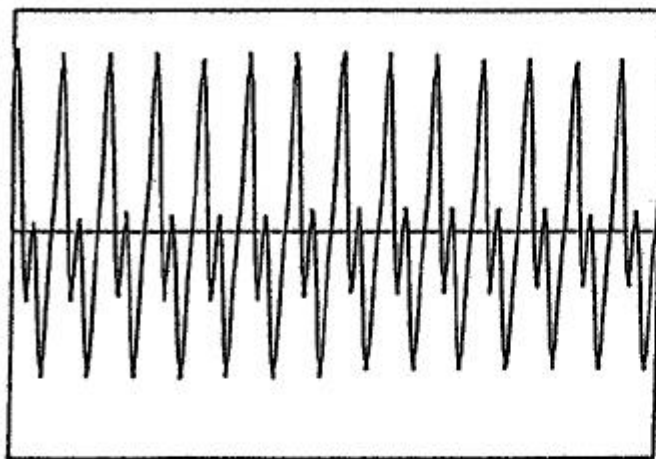


Figura 1. Dominio temporal del tono glotal de una vocal <i> normal

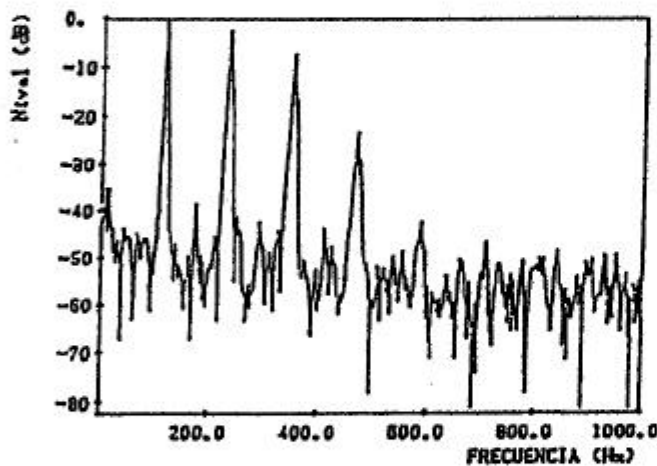


Figura 2. Dominio espectral del tono glotal de una vocal <i> normal.

a través de dos etapas independientes. La primera etapa tiene lugar en los pliegues vocálicos: el aire que sale de los pulmones excita los pliegues vocálicos y produce una sucesión muy rápida de aperturas y cierras de los mismos. El valor de la frecuencia F_0 de esta vibración de los pliegues vocálicos (tono glotal) presenta pocas variaciones para las diferentes vocales, siendo más propio de la persona que de la vocal (éste es precisamente uno de los parámetros que nos permiten diferenciar entre

voces masculinas y femeninas). A título de ejemplo, en la Figura 1 hemos representado el dominio temporal de este tono glotal en el caso de una <i> normal. En la Figura 2 representamos el dominio espectral de esta misma señal periódica: la correspondiente señal viene definida completamente por una serie de máximos que aparecen en la región de bajas frecuencias.

Tal como hemos indicado anteriormente, el sonido glotal sufre un proceso de modulación en su ra-

diación desde los pliegues vocálicos hasta los labios. Los resultados de este proceso de modulación son diferentes para cada vocal, dado que la configuración del tracto vocálico es diferente en cada caso. Como consecuencia, el sonido producido en los pliegues vocálicos, prácticamente idéntico para todas las vocales, resulta amplificado o atenuado de forma diferente para unas frecuencias u otras en función de la vocal que se está articulando y, por lo tanto, el sonido que sale de los labios resulta diferente para cada vocal.

Desde el punto de vista acústico, las diferentes vocales se pueden caracterizar por los valores de las frecuencias, los anchos de banda y las intensidades de sus formantes (resonancias propias del tracto vocálico). En la Figura 3 hemos representado la función de transferencia (frecuencias amplificadas y atenuadas por el tracto vocálico) para una vocal <i> normal.

Desde el punto de vista fisiológico, dado que el tracto vocálico se puede considerar constituido por un conjunto de tubos cilíndricos de igual longitud y área variable, que actúa como una caja de resonancia del sonido producido inicialmente por los pliegues vocálicos, las diferentes vocales vienen definidas por las respectivas funciones de área de dicho tracto vocálico. Esta caracterización permite obtener una "radiografía" del tracto en el momento en que se articula una determinada vocal y conocer la posición de los oportunos órganos fonadores en ese momento.

4. Tratamiento digital de la señal

Como es sabido, la llegada de la voz humana a un micrófono provoca una señal analógica continua $f(t)$ cuyo valor en cada instante corresponde a la tensión producida en los bornes del micrófono. Dado que con este tipo de señales no se pueden utilizar los algoritmos implementables en un ordenador, para llevar a cabo nuestro trabajo es necesario convertir dicha señal analógica en una señal digital $s(n)$ constituida por una secuencia dis-

creta de valores de la variable independiente n .

El anterior proceso se lleva a cabo fácilmente mediante un conversor analógico-digital (AD). Estos conversores poseen dos características fundamentales, que determinan la calidad del proceso de conversión: la frecuencia de muestreo de la señal analógica y los bits de resolución de la conversión.

La frecuencia de muestreo de dicho conversor debe ser ajustada de forma tal que en el proceso de conversión citado anteriormente no se pierda información relevante. En nuestra investigación, dicha frecuencia de muestreo ha sido fijada en 12.8 kHz, lo que nos ha permitido analizar espectralmente las voces humanas hasta un fondo de escala de 6 kHz (el rango de frecuencias cubierto por este trabajo está comprendido entre 12.5 Hz y 5.0 kHz, un intervalo más que suficiente para registrar completamente los cuatro primeros formantes de todas las vocales).

Por otra parte, hemos optado por utilizar un conversor de 12 bits. Este conversor transforma la señal analógica en una señal digital cuyas muestras pueden tomar un total de 40960 valores diferentes, lo cual representa un nivel de resolución completamente satisfactorio para nuestros objetivos.

5. Análisis espectral

La deducción de los valores de los parámetros vocálicos mencionados anteriormente se ha llevado a cabo utilizando dos de los algoritmos más corrientes en el tratamiento digital de señales, la Transformada Rápida de Fourier y la Predicción Lineal (6) (7) (8).

En particular, la Transformada Rápida de Fourier (FFT) se ha utilizado en el algoritmo que se conoce con el nombre de "Cepstrum" (espectro del logaritmo de un espectro), para evaluar la frecuencia fundamental con que vibran los pliegues vocálicos en el proceso de fonación de las vocales (9). La aplicación de este algoritmo a un tramo concreto de sonido vocálico proporciona una gráfica a partir

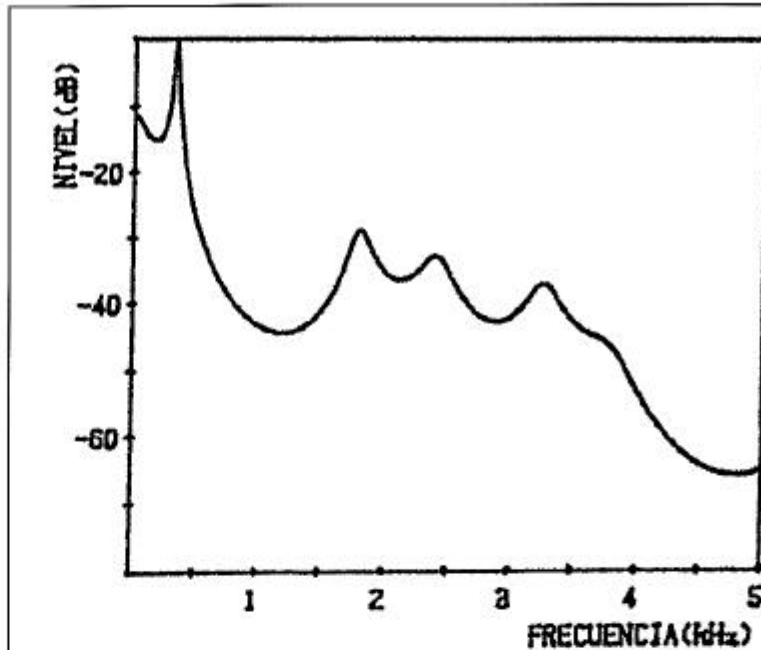


Figura 3. Función de transferencia de una vocal <i>normal</i>

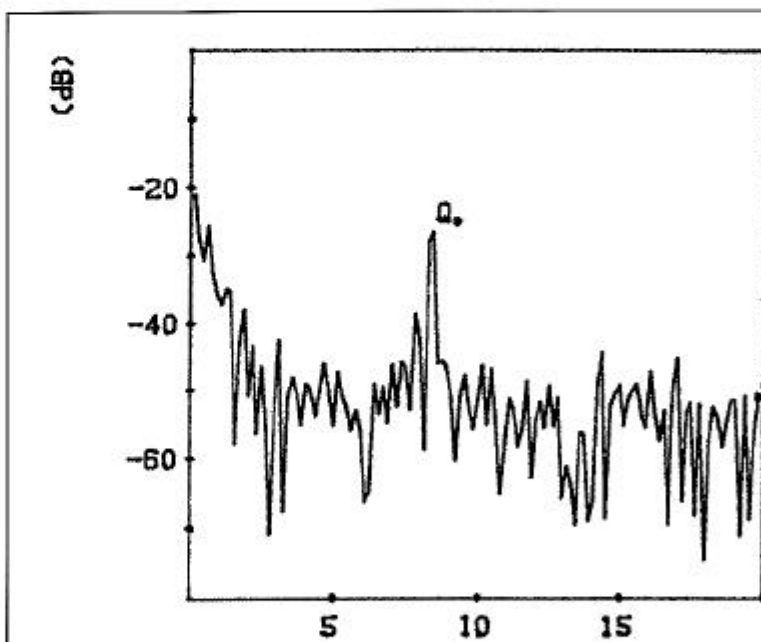


Figura 4. Ejemplo de Cepstrum vocálico

de la cual se obtiene muy fácilmente la frecuencia fundamental de dicho sonido. En la Figura 4 hemos representado un ejemplo de Cepstrum vocálico. Dado que Q_0 es la Quefrecuencia en ms (periodo temporal de la vibración de los pliegues vo-

cálicos), la frecuencia fundamental F_0 buscada vendrá dada simplemente por la inversa de Q_0 .

Los algoritmos de Predicción Lineal (LPC) están basados en la suposición de que la voz es una señal acústica linealmente predecible. Esto

Tabla 1. Lista de palabras utilizadas en las grabaciones

	a	e	è	i	o	ò	u
pam	pésol	pèl	pipa	jove	poble	judici	
mánec	metge	mèl	bisbe	soroll	got	dur	
gat	fes	fèmur	vida	dona	vol	lucid	
dada	cera	cèntim	ciri	baró	sol	casull	
nas	demà	dècim	dit	tonell	dona	ruc	

Tabla 2. Valores medios y desviaciones típicas de la frecuencia fundamental F0 (en Hz) para las diferentes vocales valencianas.

Vocal	a	e	è	i	o	ò	u
Val. med.	135	132	137	136	132	128	138
Des. típ.	19	19	22	28	19	22	30

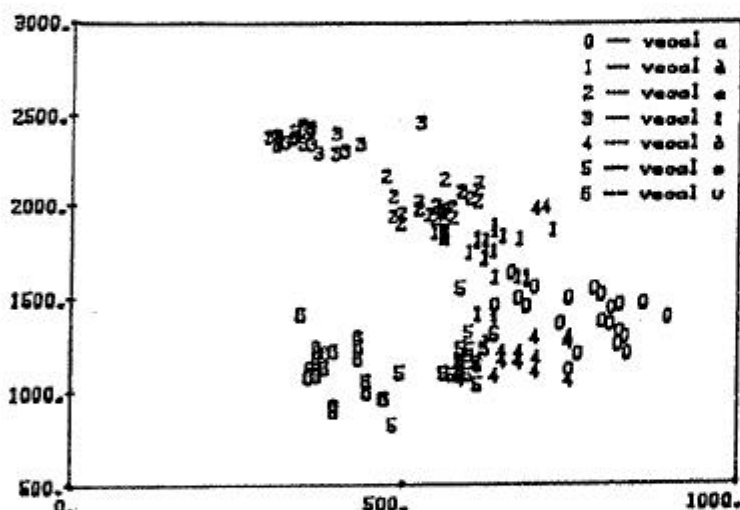


Figura 5. Variabilidad de la voz humana. Plano de los dos primeros formantes vocálicos F1/F2 para un único informante.

significa que, a partir de una serie finita de muestras de dicha voz $\langle x(t_1), x(t_2), \dots, x(t_n) \rangle$, se puede predecir el valor de la muestra que corresponderá al instante $t(n+1)$. Esta muestra vendrá dada por una combinación lineal de las muestras anteriores. Si la serie de muestras sucesivas que se toman pertenecen a un sonido vocálico, los parámetros de la correspondiente combinación lineal permiten obtener los formantes de la vocal articulada, así como la anatomía del tracto con que se ha articulado esa vocal en particular.

Este método se puede implementar fácilmente en un programa de ordenador que puede considerar como datos iniciales las muestras digitalizadas de la palabra en la que se encuentra la vocal a estudiar, así como el instante temporal en que se encuentra la parte estacionaria de dicha vocal. A partir de estos datos, el programa preparado por nosotros calcula la frecuencia fundamental, las frecuencias centrales, los anchos de banda y el nivel relativo de los cuatro primeros formantes de la vocal correspondiente, así como la anatomía del tracto con que se ha articulado esa vocal.

6. Informantes y palabras utilizadas

La voz humana presenta una gran variabilidad para personas diferentes e incluso para una misma persona en circunstancias diferentes. En este trabajo se han analizado un total de 25 informantes varones, con edades comprendidas entre los 20 y 50 años, procedentes de cinco zonas dialectales diferentes de la Comunidad Valenciana (L' Alcoià/Comtat, Marina Baixa, La Safor, L'Horta y Plana Baixa). El número total de informantes considerado en nuestra investigación es del mismo orden que el utilizado en otros trabajos similares a éste (5) y, en cualquier caso, resulta adecuado para un estudio general de caracterización vocálica.

Cada uno de estos informantes ha leído una lista de 35 palabras diferentes; esta lectura se ha realizado siempre de forma pausada y con la pronunciación más natural posible. Para facilitar el proceso, se han elegido palabras cortas y de uso común en el lenguaje cotidiano. En la Tabla 1 hemos reproducido la lista de palabras utilizadas. Obsérvese que las diferentes vocales se presentan en cinco contextos consonante-vocal (CV) diferentes; estimamos que este número es suficiente para nuestros objetivos.

7. Realización y procesado de las grabaciones

Todos los informantes han llevado a cabo la lectura de la anterior lista de palabras en una cámara insonorizada, en la que el nivel de ruido de fondo ha sido siempre inferior a 35 dBA. Nuestra cadena de grabación estaba constituida por un micrófono Brüel & Kjaer 4155 de respuesta plana entre 3 Hz y 20 kHz, un sonómetro BK2230 y una registradora de cinta UHER 4200 provistas de cintas de cromo BASF.

El análisis de las grabaciones se basa fundamentalmente en la utilización de un conversor AD (encargado de convertir la señal analógica procedente de la registradora

Tabla 3. Valores medios y desviaciones típicas de las frecuencias centrales (en Hz) de los cuatro primeros formantes de las vocales valencianas.

Vocal	a	e	è	i	o	ò	u
1. Val. med.	761	511	620	326	535	661	393
Des. típ.	85	70	54	34	60	51	47
2. Val. med.	1242	1910	1742	2203	1088	1060	1018
Des. típ.	108	151	147	181	120	89	166
3. Val. med.	2464	2583	2533	2831	2325	2380	2316
Des. típ.	247	190	200	210	273	282	273
4. Val. med.	3481	3626	3687	3508	3304	3387	3307
Des. típ.	347	296	339	280	339	441	345

Tabla 4. Valores medios y desviaciones típicas de los anchos de banda (en Hz) de los cuatro primeros formantes de las vocales valencianas.

Vocal	a	e	è	i	o	ò	u
1. Val. med.	133	108	89	112	121	135	153
Des. típ.	70	65	57	75	72	79	97
2. Val. med.	163	166	163	123	112	149	170
Des. típ.	114	122	167	70	72	94	100
3. Val. med.	365	270	353	252	294	321	308
Des. típ.	225	160	200	128	144	197	205
4. Val. med.	395	302	377	243	352	358	380
Des. típ.	277	209	209	177	194	207	282

en una señal digital) y un ordenador AT (encargado de procesar mediante determinados algoritmos estas muestras digitales). De hecho, el convertor AD se encuentra instalado en un analizador de espectros en tiempo real Brüel & Kjaer 2033 conectado al citado ordenador mediante un interface GP-IB.

El convertor AD utilizado posee una resolución de 12 bits, con una frecuencia de muestreo de 12.8 kHz. Este convertor permite cubrir un intervalo de señal digitalizada de 800 milisegundos, con un ancho de banda de hasta 5 kHz, lo cual permite registrar todas las palabras completas consideradas en el presente trabajo. Las 10240 muestras por palabra que están almacenadas en ciertas direcciones de memoria del analizador espectral BK2033 se transfieren posteriormente a un ordenador Tandon AT para su correspondiente procesamiento.

Todos los programas utilizados en este proceso han sido desarrollados en nuestro Laboratorio. Un primer programa transforma el fi-

chero ASCII creado por el analizador BK2033 en un fichero binario codificado de forma correcta. A partir de este fichero, mediante dos nuevos programas, se calculan los parámetros acústicos y articulatorios que definen a cada una de las vocales consideradas. El paquete incluye además una serie de programas de utilidades preparados para realizar un tratamiento estadístico y gráfico de todos los resultados almacenados en la correspondiente base de datos.

Tabla 5. Valores medios y desviaciones típicas de los niveles relativos al máximo (en dB) de los cuatro primeros formantes de las vocales valencianas.

Vocal	a	e	è	i	o	ò	u
1. Val. med.	-6	-3	-2	-6	-3	-1	-1
Des. típ.	4	5	3	6	4	3	3
2. Val. med.	-3	-5	-5	-6	-4	-3	-7
Des. típ.	4	6	6	7	5	3	6
3. Val. med.	-16	-10	-13	-10	-20	-18	-19
Des. típ.	7	6	8	9	10	9	10
4. Val. med.	-21	-17	-21	-14	-24	-23	-22
Des. típ.	9	9	11	9	12	11	15

8. Resultados y discusión

8.1. Variabilidad en la voz humana

El término variabilidad de la voz humana se utiliza aquí para referirnos al hecho de que una misma persona, cuando articula diferentes veces un mismo sonido, no obtiene nunca resultados absolutamente idénticos, sino que en cada caso el sonido producido es diferente. Las diferencias entre estos sonidos (voz) dependen del nivel de dificultad que entraña la producción del sonido en cuestión, el estado de los órganos fonadores y las condiciones específicas del locutor.

El estudio de la variabilidad de la voz humana puede ser importante en muchas aplicaciones: reconocimiento de la voz, estudio de las diferencias dialectales de una lengua, etc.

En el presente trabajo, hemos estudiado la variabilidad de los parámetros que definen a las 7 vocales valencianas para tres informantes diferentes. Cada uno de estos informantes ha leído cinco veces seguidas la misma lista de palabras.

Nuestros resultados confirman que la variabilidad de estos parámetros (referidos tanto a las frecuencias fundamentales como a los formantes de las vocales) es muy importante: los valores medios obtenidos como resultado de estos análisis para los datos de cada uno de los informantes presentan una dispersión del mismo orden de magnitud que la obtenida cuando promediamos los datos de todos los informantes. Este hecho

viene ilustrado en la Figura 5, en la que se representan los valores de los dos primeros formantes (F1 y F2) para uno de los informantes estudiados.

Este hecho nos impide determinar, por ejemplo, si existen diferencias dialectales entre los diferentes locutores: en el caso de que existieran realmente tales diferencias, la propia variabilidad de la voz de una misma persona impediría que se pusieran de manifiesto. Evidentemente, esta consideración se refiere estrictamente al contexto estudiado en este trabajo, es decir, tales diferencias serían más difíciles de localizar en el estudio de ciertos parámetros no estacionarios, tales como transiciones de formantes en contextos consonante-vocal o en la evolución de la frecuencia fundamental con el tiempo. La comprobación de todas estas hipótesis será objeto de futuros trabajos.

8.2. Frecuencia fundamental

Como ya hemos indicado anteriormente, la frecuencia fundamental F0 es la frecuencia de vibración de los pliegues vocálicos. Este parámetro es más característico del locutor y de su anatomía que de la propia vocal y su valor representa la frecuencia de vibración de los pliegues vocálicos, siendo función de factores tan diversos como el tamaño físico de los pliegues y el flujo de aire que sale de los pulmones cuando se articula la vocal. A mayor tamaño de los pliegues menor es la frecuencia de vibración; esto explica el hecho de que las mujeres presenten frecuencias fundamentales mayores que los hombres (voces más agudas). A mayor flujo de aire pasando por los pliegues mayor es la frecuencia en la vibración de éstos.

Por lo tanto, los valores de las frecuencias fundamentales de cada vocal que hemos obtenido en el presente trabajo deben entenderse como valores medios para una población de varones adultos. En la Tabla 2 recogemos dichos valores medios y las correspondientes desviaciones típicas. Como puede observarse, los valores de F0 para las siete vocales son prác-

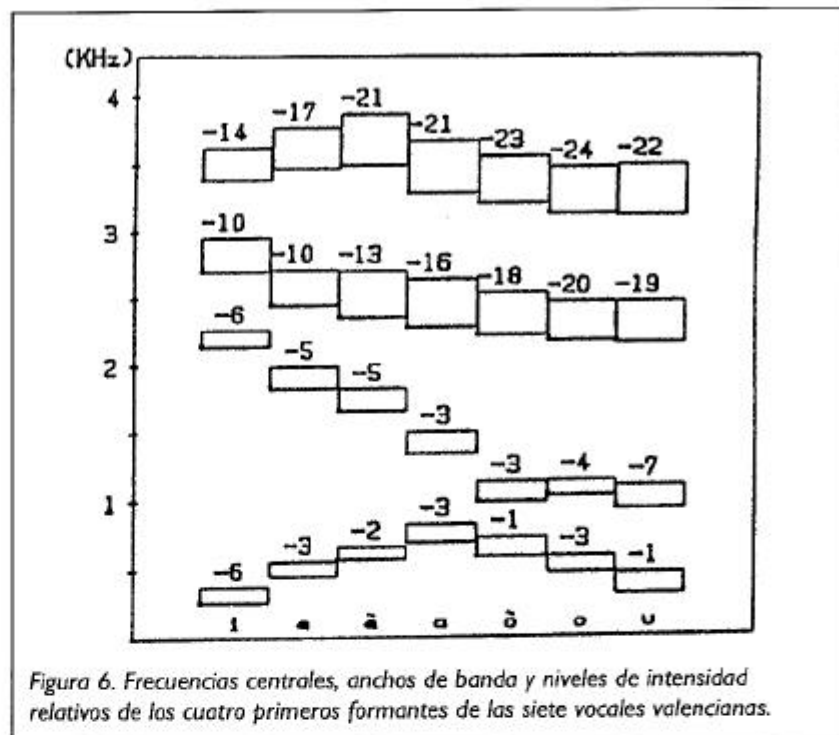


Figura 6. Frecuencias centrales, anchos de banda y niveles de intensidad relativos de las cuatro primeras formantes de las siete vocales valencianas.

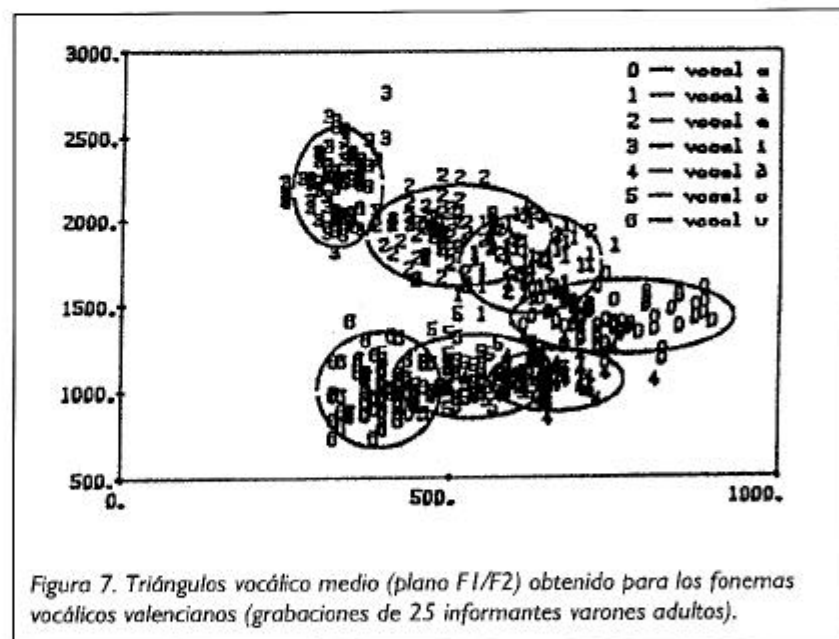


Figura 7. Triángulo vocálico medio (plano F1/F2) obtenido para los fonemas vocálicos valencianos (grabaciones de 25 informantes varones adultos).

ticamente coincidentes: la anatomía de los pliegues vocálicos de los locutores obliga a éstos a articular todas las vocales con frecuencias fundamentales muy parecidas.

8.3. Formantes

En la Tabla 3, 4 y 5 recogemos los valores medios de las frecuencias cen-

trales, los anchos de banda y los niveles de intensidad relativos de los cuatro primeros formantes de las vocales valencianas. La frecuencia central es la frecuencia en que se localiza la máxima amplificación del formante. El ancho de banda indica el intervalo de frecuencias para el que la amplificación del formante ha disminuido en 3 dB. Finalmente, el nivel de in-

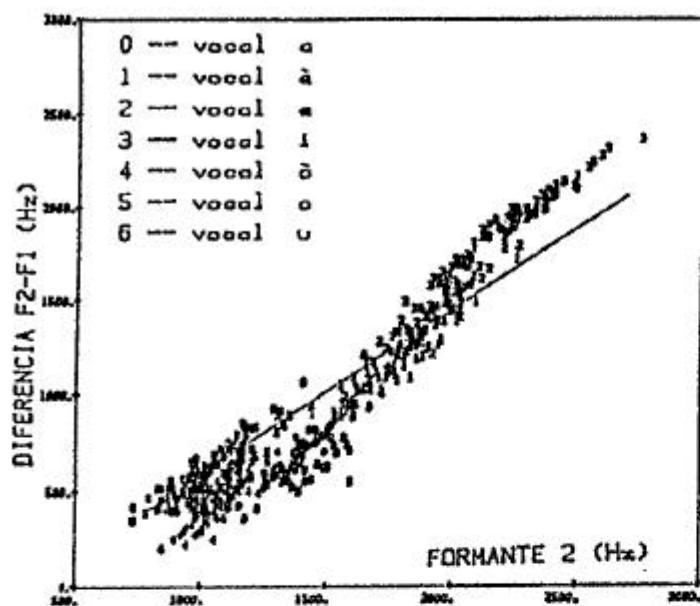


Figura 8. Correlación entre el valor de la diferencia de las frecuencias centrales de los dos primeros formantes (F_2-F_1) y el valor de la frecuencia del segundo formante (F_2).

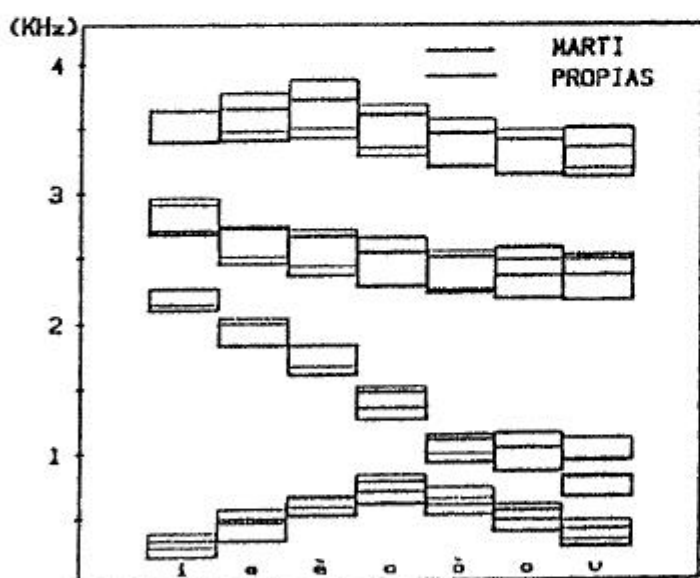


Figura 9. Comparación entre los valores de las frecuencias centrales y anchos de banda de las vocales encontrados en el estudio de J. Martí (catalán central) y en el presente trabajo (valenciano).

tensidad relativo expresa la diferencia entre la intensidad de dicho formante y la del formante más intenso.

Así como la frecuencia fundamental es más característica de la

anatomía peculiar de los locutores que de las propias vocales, los formantes son función de las vocales y no tanto de los locutores. La información que nos permite di-

ferenciar a las diferentes vocales se concentra fundamentalmente en el primer y segundo formante; el hecho de que estos formantes sean los más intensos facilita la discriminación.

Como resumen de este apartado, en la Figura 6 reproducimos una gráfica en la que los formantes de cada vocal vienen representados en la frecuencia central del formante, mientras que su altura expresa el ancho de banda. Los números que aparecen sobre cada uno de dichos rectángulos indican los valores de las intensidades de cada formante.

8.4. El plano F_1/F_2

El llamado plano F_1/F_2 constituye la forma más habitual de caracterizar los fonemas vocálicos de un idioma. Este plano es la presentación cartesiana de las frecuencias centrales de los dos primeros formantes de cada vocal. Debido a la gran variabilidad de la voz humana (véase el apartado 8.1), no podemos esperar que las coordenadas de las diferentes vocales presenten unos valores perfectamente definidos para un grupo de informantes diferentes.

En la Figura 7 hemos reproducido el plano F_1/F_2 para los fonemas vocálicos valencianos estudiados en este trabajo. Los ejes de las elipses que delimitan cada una de las siete zonas que definen a estas vocales valencianas tiene una longitud equivalente al doble de las correspondientes desviaciones típicas. La vocal más abierta es la <a>, seguida por la <è> y <ò>. Las vocales más cerradas son la <i> y la <u>, por este mismo orden.

8.5. La recta F_2-F_1 en función de F_2

Con el fin de reducir al mínimo la información necesaria para reconocer o sintetizar vocales, resulta muy interesante encontrar correlaciones entre los diferentes parámetros que caracterizan dichas vocales. El análisis de todos los datos obtenidos en este trabajo nos ha permitido determinar la relación existente entre la diferencia entre los valores centrales de

las frecuencias de los dos primeros formantes (F2-F1) y la del segundo formante (F2). Las correlaciones buscadas responden a las siguientes rectas de regresión:

Vocal a:	$(F2-F1) = 0.9047 F2 - 617$	$(r = 0.777)$
Vocal e:	$(F2-F1) = 1.1140 F2 - 727$	$(r = 0.926)$
Vocal è:	$(F2-F1) = 1.0600 F2 - 724$	$(r = 0.947)$
Vocal i:	$(F2-F1) = 0.9754 F2 - 272$	$(r = 0.982)$
Vocal o:	$(F2-F1) = 0.9665 F2 - 500$	$(r = 0.880)$
Vocal ò:	$(F2-F1) = 0.9808 F2 - 635$	$(r = 0.863)$
Vocal u:	$(F2-F1) = 0.5739 F2 - 354$	$(r = 0.960)$

En la Figura 8 hemos representado la totalidad de los datos disponibles (todas las vocales juntas). La recta de regresión obtenida en este caso responde a la ecuación:

$$(F2-F1) = 1.0990 F2 - 690$$

$$(r = 0.962)$$

Obsérvese que los valores de los coeficientes de correlación son bastante elevados. Resulta también interesante el hecho de que los datos incluidos en la Figura 8 parecen sugerir la existencia de dos pendientes diferentes, una válida para las vocales en las que F2 es muy diferente de F1 y otra para las vocales en que las diferencias entre F2 y F1 son menores.

8.6 Comparación con otros trabajos

Todas las vocales presentes en el vocalismo valenciano están presentes en el vocalismo del catalán central estudiado previamente por J. Martí (4). Dado que el método de análisis utilizado por este autor es diferente al nuestro, sólo cabe comparar los valores de las frecuencias centrales y anchos de banda de los respectivos formantes.

En la Figura 9 hemos superpuesto los rectángulos relativos a los cuatro primeros formantes de cada vocal encontrados en ambos trabajos. En general, la coincidencia es muy buena; la única diferencia importante corresponde al segundo formante de la

vocal <u>. Aunque es evidente que existen diferencias muy notables entre ambas zonas dialectales (concretamente, en el vocalismo valenciano sólo hay 7 vocales, en tanto

que en el catalán central hay 10 vocales), estos resultados ponen de manifiesto que tales diferencias son poco significativas en lo que respecta a la parte estacionaria de las siete vocales consideradas en el presente trabajo.

8.7 Articulación de las vocales valencianas

Como ya hemos indicado anteriormente, la acústica de las vocales tiene su origen en el proceso de vibración de los pliegues vocálicos y su posterior modulación en la cavidad de resonancia en que se transforma el tracto vocálico. En otras palabras, la acústica es el efecto, en tanto que la causa hay que buscarla en el funcionamiento de los

pliegues y en la anatomía particular del tracto vocálico.

En particular, el estudio de la anatomía de dicho tracto resulta relativamente simple usando técnicas de predicción lineal (10) (11) (12). Este método de trabajo considera que el tracto se comporta como un conjunto de tubos cilíndricos de igual longitud y diferente área. En nuestro caso particular, hemos considerado un total de 13 tubos, lo cual nos permite una resolución más que aceptable en un tiempo de cálculo no demasiado grande.

En la Tabla 6 presentamos los valores de las áreas de estos 13 tubos obtenidos en dichos cálculos para cada una de las siete vocales del valenciano; el tramo "1" corresponde a los labios y el tramo "13" a los pliegues vocálicos. Aunque todos estos valores son relativos (referidos al área máxima), hemos comprobado que la fonación viene condicionada por la forma del tracto y no por los valores absolutos de las áreas del mismo: cuando se utilizan tractos simulados con la misma forma pero con áreas diferentes, el valor de los formantes no experimenta cambios significativos.

La posición de los principales órganos fonatorios viene ilustrada gráficamente en la Figura 10, obtenida a partir de los datos recogidos en la Tabla anterior. En estas gráficas, las funciones de área del tracto vocálico configuran en cierto modo una es-

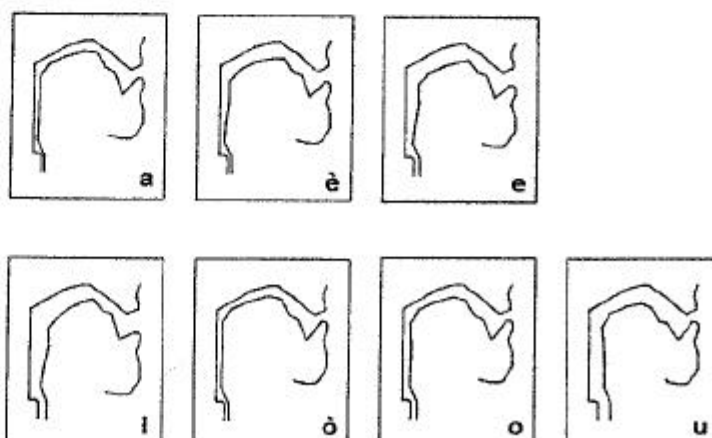


Figura 10. Tractos vocálicos calculados para cada una de las siete vocales valencianas (análisis articulatorio).

pecie de "radiografías" para cada una de las vocales consideradas en este trabajo.

Al igual de lo que sucedía con el plano F1/F2, esta figura pone de manifiesto la existencia de tres tractos diferentes y cuatro tractos intermedios. Los tractos extremos corresponden a la <a>, <i> y <u>. Los tractos con posiciones intermedias son los de la <e>, <è>, <o> y <ò>. En el caso de la vocal <a>, la posición de todos los órganos es muy relajada, es decir, no se aprecia ninguna constricción. En el caso de la vocal <i>, las posiciones de los órganos fonatorios varían: la lengua se adelanta subiendo hasta la parte anterior del paladar, apreciándose una ligera constricción. El caso de la vocal <u> es totalmente opuesto al de la <i>: la lengua se desplaza hacia atrás,

acercándose a la parte posterior del velo del paladar y produciendo allí una ligera constricción. El resto de las vocales adoptan posiciones intermedias a las de éstas. Así, los tractos con que se articulan las vocales <e> y <è> tienen formas intermedias a los de la <a> e <i>; el tracto de la <e> es similar al de la <i>, en tanto que el de la <è> se parece más al de la <a>. Por lo que respecta a las vocales <o> y <ò>, sucede algo similar a lo anterior, pero ahora en relación con las vocales <u> y <a>.

Como conclusión, podemos afirmar que las posiciones de los órganos fonadores en el tracto vocálico (causa) no presentan diferencias significativas para las diferentes vocales. Como consecuencia de este hecho, los formantes de todas estas vocales

(efecto) tampoco presentan grandes diferencias entre sí. Este hecho, unido a la considerable variabilidad de la voz humana, explica el motivo por el que en el plano F1/F2 no podemos separar claramente las diferentes vocales, sino tan sólo delimitar aproximadamente las zonas que corresponden a cada una de ellas.

Agradecimientos

El presente trabajo ha sido financiado por la Institució Valenciana d'Estudis i Investigació (IVEI), en el marco del proyecto CPE-032. Los autores desean agradecer también la amable colaboración prestada por todos los locutores en la realización de las grabaciones.

Referencias

- (1) G.Fant. "Acoustic theory of speech production". Mouton. The Hague (1960).
- (2) J.L.Flanagan. "Speech analysis, synthesis and perception". Springer-Verlag. Berlín (1972).
- (3) R.Cerdá. "El timbre vocálico en catalán". G.S.I.C. Madrid (1972).
- (4) J.Martí. "Paràmetres vocàlics del català". Folia Phonetica 1, 23 (1984).
- (5) J. Martí. "Estudi acústic del català i síntesi automàtica per ordinador". Tesi Doctoral. Universitat de València (1986).
- (6) L.R.Rabiner y R.W.Schafer. "Digital processing of speech signals". Prentice Hall. New York (1978).
- (7) J. Markhoul. "Spectral linear prediction: properties and applications". IEEE Transactions A.S.S.P. 23, 283 (1975).
- (8) A.J.Markel y A.H.Gray. "Linear prediction of speech". Springer-Verlag. Berlín (1980).
- (9) A.M.Noll. "Cepstrum pitch determination". Journal Acoust. Soc. Amer. 41, 293 (1967).
- (10) M.M.Sondhi. "Estimation of vocal tract areas: the need for acoustical measurement". IEEE Transactions A.S.S.P. 27, 268 (1979).
- (11) H.Wakita. "Estimation of vocal tract shapes from acoustical analysis of the speech wave: the state of the art". IEEE Transactions A.S.S.P. 27, 281 (1979).
- (12) F.Lonchamp, J.P.Lefevre y J.P.Zerling. "Comparaison de deux méthodes d'estimation des fonctions d'aires vocaliques". XI Congrès International d'Acoustique, 4, 23 (1983).