

PROPUESTA DE EXPERIMENTO SOBRE LA NOCIÓN DE CAMPO DE DISPERSIÓN FONEMÁTICA

RAMON CERDÀ (1), MIREIA FARRÚS (2)
JAVIER HERNANDO (2), MONTSERRAT VEYRAT (3)

(1) *Universitat de Barcelona*
(2) *Universitat Politècnica de Catalunya*
(3) *Universitat de Valencia*

HIGGINS. [...] Tired of listening to sounds?

PICKERING. Yes. It's a fearful strain. I rather fancied myself because I can pronounce twenty-four distinct vowel sounds; but your hundred and thirty beat me. I can't hear a bit of difference between most of them.

HIGGINS. [...] Oh, that comes with practice. You hear no difference at first; but you keep on listening, and presently you find they're all as different as A from B.

(BERNARD SHAW, *Pygmalion*, Act II.)

1. PREÁMBULO

La presente comunicación no ofrece todavía resultados de ningún experimento ya ejecutado, sino, como su título indica, una serie de comentarios en torno a la noción de 'campo de dispersión fonemática' (o 'fonológica') o 'nube de realización vocálica' destinados a presentar las bases preliminares para una experimentación múltiple que de paso pretende reabrir una oportunidad de cooperación efectiva entre perspectivas lingüísticas e ingenieriles orientada hacia la identificación individualizada de la voz. Más en concreto, presenta el

marco conceptual de la propuesta –desde la perspectiva de la fonética lingüística– así como los indicios de sus posibilidades de éxito a partir sobre todo de referencias remotamente extraídas de Cerdà 1972.

2. PRESENTACIÓN

2.1. A juzgar por los logros conseguidos, todo el mundo en nuestro ámbito profesional ha tenido la oportunidad de ver y quizá de comprobar –y, en consecuencia, debe saber– que, hoy por hoy, extraer información sobre lo gramatical de una emisión fónica es considerablemente más fácil que extraer información sobre lo que tiene de peculiar o individual. Acepten una vez más un breve resumen sobre el asunto, sazonado, en lo posible, con algunos comentarios no triviales que nos llevarán a la propuesta experimental.

Toda emisión fónica proferida en condiciones habituales contiene datos que permiten averiguar su contenido lingüístico y datos que envuelven ese contenido y suelen ofrecer, a su vez, la posibilidad de obtener referencias sobre el sexo, la edad, el estado de ánimo y otras informaciones acerca de quien la ha emitido, a menudo incluso sobre su identidad. Los llamaremos respectivamente ‘datos α ’ y ‘datos β ’. Tanto para el hablante como para el lingüista, el contenido lingüístico de toda emisión, los datos α , constituye su núcleo de atención, alrededor y dentro del cual se inscriben los demás datos β , que pueden o no ofrecer algún interés para ellos (cf. Cerdà 2000). Para ser exactos, la moderna fonética experimental en el ámbito de la lingüística nació precisamente a partir del descubrimiento de cuáles son los datos α , esto es, los datos con información gramatical, que ofrece la estructura acústica de una emisión y, secundariamente, de cuál es su correlación con los procesos articulatorios que los promueven.¹

Visto así, parece como que el problema lo encuentran quienes pretenden indagar sobre los datos β , para los cuales no se han encontrado correlaciones precisas –para ser precisos, ni siquiera imprecisas–, entre otras razones porque la información que se pretende extraer de ellos aparece totalmente

¹ Esto sólo fue posible gracias a la correlación metodológica entre la espectrografía y la síntesis del lenguaje. Así sucedió, por ejemplo, cuando Delattre (1951) estableció que la frecuencia del primer formante vocálico es inversamente proporcional a la elevación de la lengua.

difuminada en el conjunto de la emisión, envolviéndolo todo, hasta los datos α que contienen la información gramatical. Ante esta situación, si los investigadores de los datos β envolventes pueden, no obstante, conformarse con indicios aproximados, explotan al máximo dichos indicios echando mano de recursos intuitivos, tal como sucede, pongamos por caso, en el mundo del canto y la declamación.² Pero si, por el contrario, son instados a suministrar resultados precisos e inequívocos –tal como sucede en nuestro ámbito forense–, entonces deben recurrir a exploraciones analíticas específicas y típicamente muy distintas, al menos en la práctica, de las que se emplean en fonética lingüística (Cerdà 2005).

A pesar de ello, la cooperación en acústica forense entre lingüistas e ingenieros no tiene por qué ser, al menos desde el punto de vista metodológico, completamente inconexa o al menos no tiene por qué tender a serlo. Es cierto que el vasto conjunto de los datos observables, α y β , invita a distinguir distintas perspectivas, simplificadas en dos niveles, uno alto y otro bajo, cuyos extremos propenden a ser respectivamente privativos de lingüistas e ingenieros: el nivel alto, que comprende en distintos estratos o subniveles las variables lingüísticas (rasgos dialectales, expletivos recurrentes...), y el nivel bajo que comprende estimaciones físicas sobre las variables acústicas (perfiles de energía, coeficientes cepstrales...). Lo importante es que existen zonas en el nivel bajo que comprenden variables, como la entonación, la intensidad o la duración, en cuya estructura física se solapan interpretaciones lingüísticas efectivas o, dicho de otro modo, donde el fonetista lingüista también tiene algo que decir por la sencilla razón de que son variables que contienen datos α .³ En cualquier caso, los campos de dispersión fonemática se encuentran en el nivel más bajo al que el fonetista lingüista tradicionalmente suele acceder.

2.2 Antes de profundizar en nuestra propuesta, creemos que merece la pena citar, siquiera sea de paso, ciertas falsas apariencias sobre la deriva de la fonética lingüística especialmente en este último nivel. En efecto, aunque parece

² En rigor, recurren a las estrategias que utilizan como hablantes tratando, eso sí, de profundizar al máximo sobre ellas.

³ Lo cierto es que la información que ofrecen aquí los lingüistas es radicalmente más pobre que la que ofrecen sobre los datos fonemáticos debido a que en este espacio trabajan con elementos analógicos que no se dejan digitalizar como las vocales y las consonantes, y porque en ellos confluyen distintos estratos de análisis lingüístico (fonológicos, morfológicos, gramaticales, pragmáticos...) que los lingüistas tienden también a tratar por separado.

que desde tiempo atrás los lingüistas tienen más o menos cubiertas sus necesidades indagadoras, en realidad no existe un acuerdo unánime sobre las interpretaciones que extraen de sus experimentos, en especial sobre el análisis fonemático. Nos referiremos tan solo a un par de casos que sin duda han obstaculizado la cooperación con los ingenieros.

En primer lugar, no todos los fonetistas lingüistas parecen entender que la separación entre los datos α gramaticales, o interpretables gramaticalmente, y los datos β peculiares de toda emisión no constituyen conjuntos complementarios, disjuntos –como se desprende de muchas estimaciones interpretativas, que cabe visualizar en la figura 1–, sino que estos datos peculiares inundan todo el espacio de la emisión y, por tanto, incluyen el conjunto de los mismos datos gramaticales, como se desprende de la figura 2.

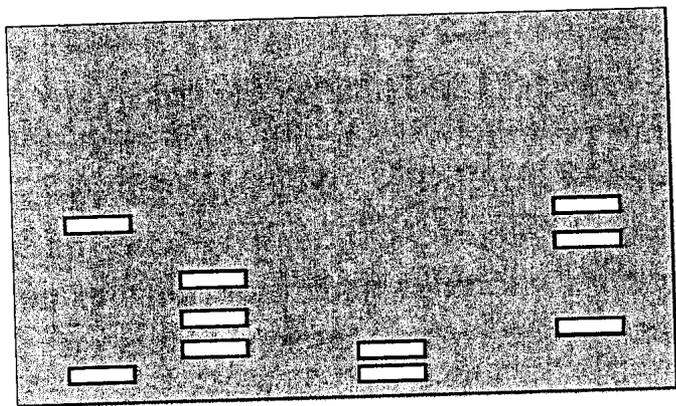


Figura 1

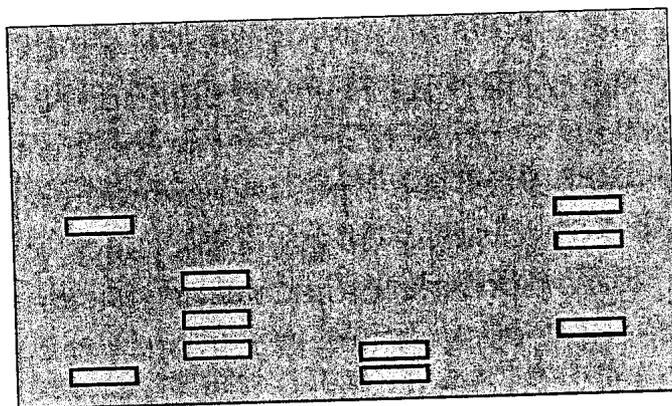


Figura 2

Como es sabido, en términos comunicativos, esto implica que dos hablantes (normales) de un mismo dialecto utilizarán, pongamos, un mismo esquema fonológico con iguales elementos relativamente identificados o distinguidos entre sí por una pronunciación clara y suficientemente diferenciada. Sus voces, sin embargo, serán diferentes en múltiples aspectos derivados de sus respectivas estructuras anatómicas y del dinamismo enormemente variado que imprimen a la articulación factores tanto habituales como esporádicos. Esto se percibe especialmente bien en el espectrograma, el tradicional medio de análisis acústico para los lingüistas. Así, si comparamos, por ejemplo, la manifestación espectrográfica de sus respectivos sistemas vocálicos, comprobaremos que presentará una coincidencia en el número y en la disposición abstracta relativa –triangular, cuadrangular...– de los fonemas (datos α), pero una disposición absoluta diferente, que se traducirá en un marco geométrico

más o menos grande o deformado respecto de otros (datos β). Veámoslo en un ejemplo concreto.

3. ANTECEDENTES

3.1 De Cerdà 1972 extraemos dos representaciones del sistema vocálico del catalán central por medio de sendos triángulos inscritos en las llamadas cartas logarítmicas de frecuencias, formadas, como se sabe, por un sistema de coordenadas cuya abscisa refleja la frecuencia del primer formante y la ordenada la frecuencia del segundo, de modo que los puntos vocálicos se definen por la intersección de ambas frecuencias. El de la figura 3 corresponde al informante B, en torno al cual gira el trabajo entero, y el de la figura 4 al informante C, obtenido tan solo a partir de la emisión aislada y correlativa de las vocales: [i], [e], [ɛ], [ə], [a], [ɔ], [o], [u].⁴ Para ser precisos, el triángulo de la figura 3 presenta los puntos geoméricamente centrales de las nubes de realización vocálica o, en términos tradicionales, campos de dispersión fonemática, o fonológica (figura 5), es decir, el lugar marcado por la intersección del eje vertical y del eje horizontal trazados ambos sobre el espacio que sobre la carta de frecuencias ocupa el conjunto de incidencias vocálicas (señaladas con puntos en la figura) analizadas a partir de la emisión de una colección de frases compuesta para que aparecieran (más o menos) todos los contextos fonológicos posibles para cada elemento fonemático.⁵

Conviene subrayar que tanto el triángulo de la figura 3, obtenido a partir de una pronunciación más o menos espontánea, como el segundo de la figura 4, obtenido a partir de una pronunciación aislada –y, por tanto, lingüísticamente abstracta– pueden considerarse canónicamente aceptables para cubrir su cometido experimental. El primero porque constituye una abstracción a partir de un número considerable de frases leídas *ad hoc* y, por tanto, perfectamente inteligibles y claramente pronunciadas. Ello no impide, desde luego, la posibilidad de que ninguna de las incidencias vocálicas realmente

⁴ Omitimos los detalles, irrelevantes aquí, sobre el complejo comportamiento fonológico del conjunto. Es necesario advertir también que la obra en cuestión, que nos servirá de referencia principal, no contiene ninguna prueba directa sobre el experimento que se propone en la comunicación, pero ofrece, sobre todo en sus conclusiones, indicios inequívocos a su favor.

analizadas y ubicadas en la carta coincide con el punto geométrico que se asignó a lo que cabe llamar cada 'fonema ideal'. En cuanto al segundo triángulo, también resulta metodológicamente aceptable por la validación auditiva directa que diversos hablantes nativos otorgaron a la pronunciación de los sonidos aislados como identificación igualmente ideal de los respectivos fonemas.⁶

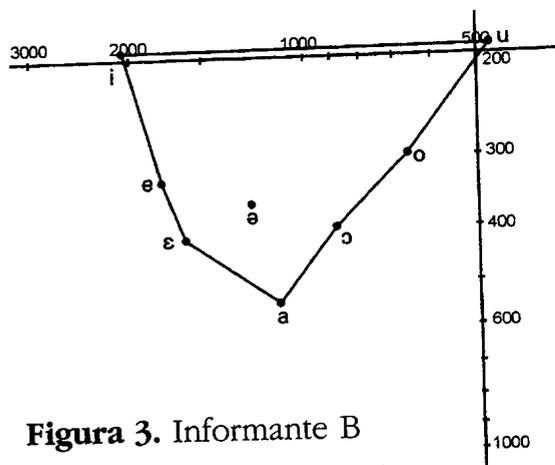


Figura 3. Informante B

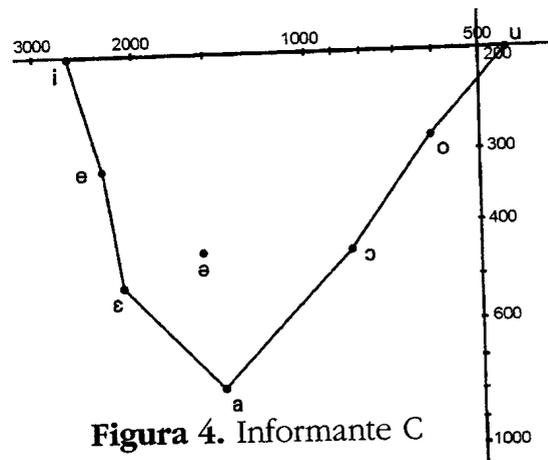


Figura 4. Informante C

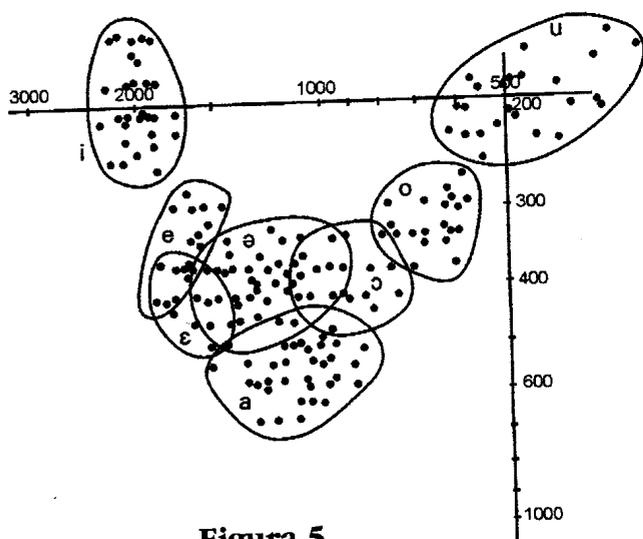


Figura 5

En suma, ambas figuras coinciden en presentar sendos triángulos con ocho elementos relativamente situados de un modo idéntico (dato α) y discrepan en el tamaño del conjunto y en la ubicación precisa de cada elemento (dato β). De ahí que la superposición de ambos triángulos (figura 6), presenta, como decía antes (2.2), diferencias evidentes que, en la obra de referencia, se atribuyen correctamente a factores individuales:

- 5 En el marco teórico de allí, estos elementos son alófonos de fonemas y de archifonemas. Para el español, véase, por ejemplo, Ramón 1979, Monroy 1980, Romero 1988, Fernández 1993, Martínez Celdrán 1995.
- 6 En realidad, la experimentación sobre pronunciaciones modélicas constituye un problema que sigue lejos de encontrar una solución satisfactoria porque los resultados que arroja ofrecen una nitidez excesiva que no se corresponde con la realidad cotidiana. Pero lo malo es que las pronunciaciones espontáneas de la realidad cotidiana aparecen dentro de un ámbito comunicativo –técnicamente, pragmlingüístico– cuyas numerosas variables trascienden, como decíamos en la nota 3, el marco estrictamente fonético.

Poco importa que el campo de dispersión fonológica de [i] de un individuo tenga, en la carta logarítmica de frecuencias, una localización alta, baja o desplazada a cualquier lado. Lo propio es que otro individuo de la misma comunidad lingüística (y mucho más si es de distinto sexo y de distinta edad) realice sus [i] en otro sitio parcial o totalmente separado (Cerdà, 1972, 149).⁷

Pues bien, no faltan lingüistas –sobre todo fonetistas y dialectólogos– que elevan a la condición de categoría absoluta,

como datos α escuetos, mediciones obtenidas con precisión milimétrica a partir de un solo informante y que hasta interpretan diferencias espectrográficas del tipo que examinamos aquí, y a veces aún más pequeñas, como un criterio inequívoco para establecer diferencias de dialecto e incluso de lengua.

3.2 Se trata, sin embargo, de un error que no es epistemológicamente baladí y que ofrece un considerable interés para nosotros. El error en cuestión consiste en no percatarse de que la noción de fonema es una idealización fonética a la que, evidentemente, tiende a acomodarse la fonación de cada hablante formando un conjunto de correlaciones articulatorias y auditivas relativamente estable, de donde derivan, como es lógico, unas propiedades acústicas relativamente estables también (datos α). Pero por ello mismo, volvemos a repetir, estas propiedades acústicas contienen al mismo tiempo las características no solo de cada fonema sino también del hablante que lo promueve (datos β). En el caso del informante B, esta mezcla se manifiesta en la gran superficie que suman en conjunto las intersecciones de las nubes de realiza-

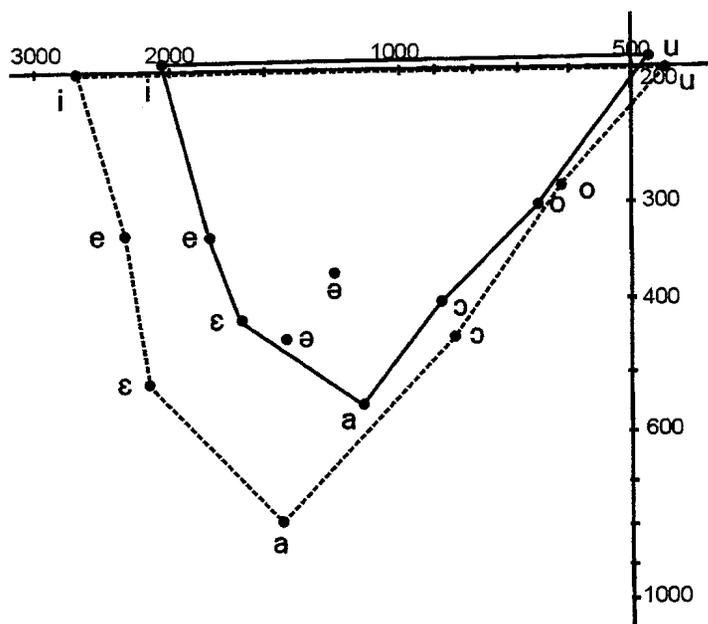


Figura 6. Informantes B y C

⁷ Cabe observar, de paso, que la serie palatal presenta mucha más diferenciación que la velar seguramente porque la zona anterior de las cavidades bucales ofrece más distinciones articulatorias (y, en consecuencia, acústicas), como se ha observado experimentalmente desde hace mucho. No solo eso; no cabe duda de que si el informante C hubiese reunido sus datos a partir de las mismas frases que el informante B, el resultado sería distinto al de la figura 4, pero en todo caso seguiría siendo distinto del de la figura 3.

ción vocálica de la figura 5 de antes. Salvo en los casos de las vocales cerradas [i] y [u], no faltan espacios frecuenciales siquiera mínimos donde concurren –al menos, a la vista de los datos β – hasta tres realizaciones, junto con otros mayores con dos; y donde la vocal central [ə] extiende su espacio por los de sus vecinas más inmediatas. ¿Qué conclusiones se desprenden de ello?

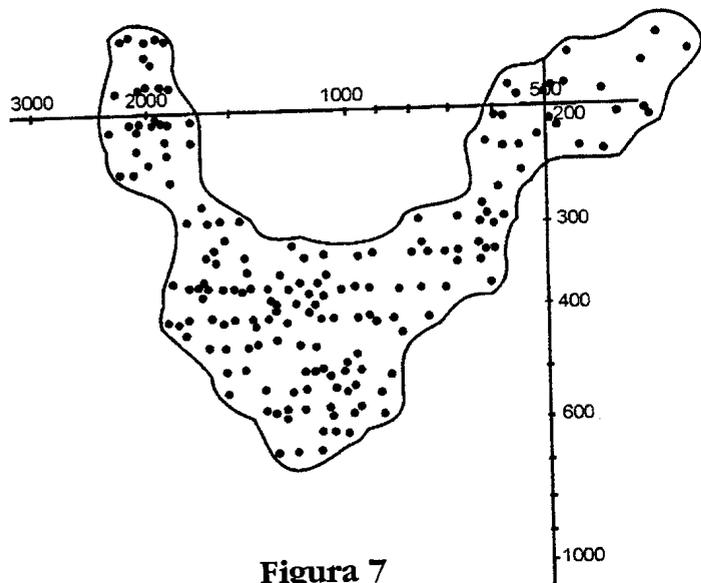


Figura 7

Para recabar una visión más completa examinemos el conjunto de incidencias espectrográficas que la obra de referencia presenta sobre los datos del informante B pero ahora sin señalar sus respectivos límites de interpretación fonológica (figura 7). La primera impresión es que el sistema vocálico en su conjunto (la constelación de datos β) ocupa un espacio acústico de un modo básicamente compacto, lo que se traduce en la imposibilidad de averiguar cuántos fonemas –y lógicamente cuáles– lo constituyen

(esto es, qué datos α contiene). Si se sintetizara su voz, aparecerían como alófonos capaces de ocupar desde posiciones extremas hasta otras supuestamente centrales pasando por secuencias intermedias con diferencias casi imperceptibles. Tanto es así, que sin duda no podría identificarse el sistema que contiene. Dicho al revés, no es descabellado pensar que un castellano hablante, un italiano o un árabe podrían reconocer entre aquellos alófonos su propio sistema vocálico, pero no así un francés, un inglés o un ruso, por la sencilla razón de que echarían de menos ciertos espacios frecuenciales en las zonas medias y altas del centro. De ello se deduce también que las denominaciones de ‘campo de dispersión fonemática’ y ‘nubes de realización vocálica’, que hasta aquí hemos considerado meras alternativas terminológicas, no lo son estrictamente hablando, pues la primera (propia de la fonética estructuralista) fue acuñada bajo el supuesto de que los datos acústicos reales reflejarían fielmente tales campos dejando incluso espacios intermedios que serían a su vez denominados ‘zonas de seguridad fonemática’ (o fonológica) sin prever, desde luego, la gran cantidad de intersecciones que se producen en realidad. Dicho de otro modo, se pensaba que los datos α eran independientes de los datos β .

En ocasiones todo ello ha inducido a pensar en la necesidad de contar no con dos sino con los tres primeros formantes vocálicos y delimitar las nubes tridimensionalmente (como en la figura 8, sobre un hipotético vocalismo del español). De este modo éstas adquieren un volumen que ocupa un espacio donde la altura puede reducir y hasta suprimir las intersecciones. Al

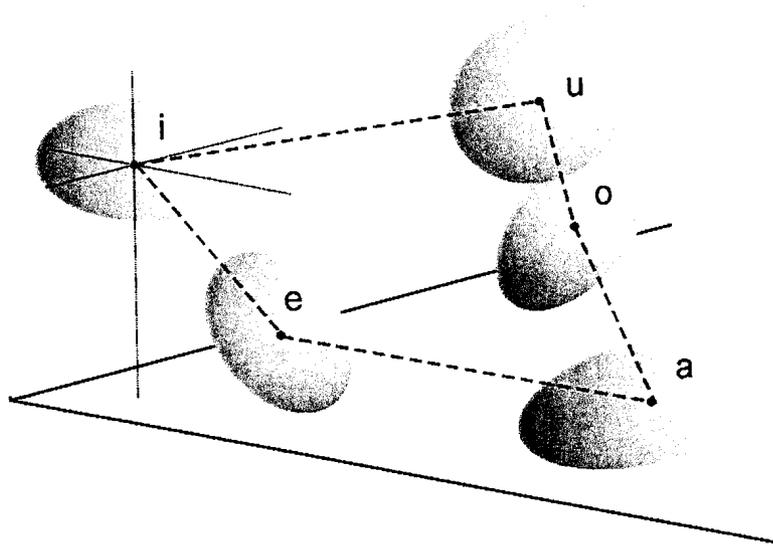


Figura 8

margen de la leve dificultad de presentar los resultados en tres dimensiones y de que a veces los terceros formantes son poco perceptibles, especialmente en la serie velar, dejamos abierta esta posibilidad para la propuesta final.⁸

Ahora bien, volviendo a la caracterización individual de la voz, es evidente que esta agrupación de incidencias alofónicas ha de contener una multitud de datos peculiares y, en consecuencia, identificadores. A la vista de las diferencias que exhibe la comparación entre los triángulos en abstracto de la figura 6, no cabe duda de que la ubicación precisa en la carta de frecuencias de un solo alófono pronunciado en un determinado contexto ofrece, al menos sobre el papel, una dosis suficiente de peculiaridad de un informante frente a otro u otros. La cuestión consiste en encontrar un cauce metodológico que satisfaga y optimice al mismo tiempo los requisitos esenciales de la acústica forense: la fiabilidad de los resultados y la operatividad experimental.

4. LA PROPUESTA

4.1 Hemos de reconocer que, a partir de lo que se desprende de Cerdà 1972, la propuesta experimental que anunciamos está inevitablemente mediatizada tanto por la certeza como por la presunción. Y es que todavía quedan un par

⁸ Entre los pocos trabajos con más de dos formantes que hemos podido localizar destacaremos tal vez Llamas 1997, Stevens 1998 y Jassem 2004.

de supuestos indiciarios en aquella obra de referencia que merecen al menos una breve atención añadida. En un momento (págs. 179-183), se da cuenta de diversos experimentos realizados por el autor sobre las características espectrográficas de vocales pronunciadas distorsionando la posición habitual de los órganos articulatorios, algo así como pronunciar [i] o [u] con gran abertura bucal, [a] con labialización, etcétera. Aun sin llegar a establecer con precisión las condiciones del experimento, los resultados espectrográficos mostraban una sorprendente regularidad acústica frente a la deformación articulatoria. Esto nos induce a pensar, sin duda un tanto apriorísticamente, en la posibilidad de que los campos de dispersión alofónica ofrezcan un cierto grado de independencia frente a factores distorsionantes tanto involuntarios como voluntarios.

El otro supuesto es como un negativo de nuestra perspectiva actual. Y es que allí se optó por intentar reproducir una réplica 'ideal' del sistema vocálico del catalán central eligiendo, como hemos visto más arriba, el punto geométrico central de cada nube de realización (los datos α por excelencia). No cabe duda de que era lo correcto teniendo en cuenta los supuestos fonológicos del trabajo, que no versaba precisamente sobre la pronunciación del informante B sino al contrario, haciendo en lo posible abstracción de sus peculiaridades. Pero si desde la perspectiva de la acústica forense lo que importa es precisamente la pronunciación del informante B, la situación se invierte. Para distinguirlo del informante C o de cualquier otro, seguramente lo más fiable, como decíamos un poco más arriba, sería comparar una gran constelación espectrográfica de sus pronunciaciones alofónicas si no fuese porque semejante cometido exigiría unas condiciones y un tiempo de experimentación casuística que raras veces se encuentran a mano.⁹ Si, por el lado opuesto, se simplifica y se extraen, como en la figura 3, los mismos puntos geométricos centrales de cada nube de realización (más o menos rica de datos; eso lo veremos luego), la indagación puede llegar a simplificarse hasta el punto de reducir en exceso la fiabilidad del resultado e invalidar la prueba.

⁹ La observación debería ampliarse, por ejemplo, a las frecuencias formánticas de alófonos en determinados contextos, como se ha hecho, por ejemplo en catalán, en Martínez Celdrán 1986, Solà (1988), Poch 1992, Andreu 1994 o Carrera 2001. Cf. también Casacuberta 1991.

4.2 Existe una alternativa intermedia entre ambos extremos. Consiste sencillamente en establecer el punto de la nube donde confluye la media de los dos o tres formantes considerados a fin de identificar el lugar preciso donde tienden a confluir las pronunciaciones alofónicas de cada vocal. No cabe duda de que se trata de una información que acentúa considerablemente la extracción de una peculiaridad acústica del informante y que puede obtenerse a partir de una serie no demasiado extensa de datos registrados en unos minutos de grabación. Incluso parece una prueba capaz de pasar por alto la diferencia –a menudo trascendental por lo negativa que resulta, especialmente para el analista lingüista– entre una dicción espontánea y una lectura forzada y realizada, encima, delante de fiscales y magistrados (cf. Quilis 1999). No hay que descartar que para alcanzar un punto óptimo de economía en la ejecución y de fiabilidad y robustez en los resultados habrá que tener en cuenta distintos requisitos de comprobación en función de las características de cada informante. Sin entrar en ello, desde luego, no es posible precisar nada aún sobre aquellos requisitos y esas características.

Se trata, en suma, de un espacio de nivel bajo y de común acceso tanto a lingüistas como ingenieros, que puede ser abordado al mismo tiempo desde distintas perspectivas –formánticas o cepstrales– por ambos a la vez. Intuitivamente al menos, y sin lugar a dudas prematuramente, estamos convencidos de que la propuesta de trabajar también en esta dirección permitirá aportar un recurso complementario metodológica y técnicamente sencillo de ejecutar en cooperación otros procedimientos que puedan practicarse (por ejemplo, en la línea de Andrews 2002, Jin 2003, Klusacek 2003, etc.).

Bibliografía

- ANDREU ANDRÉS, M^a Á. Y GARCÍA CASAS, M. (1994) “Sobre la relación entre el uso de los fonemas vocálicos y factores acústico-articulatorios: primeros resultados”, *Estudis de Fonètica Experimental*, 6, 113-132.
- ANDREWS, W.; KOHLER, M.; CAMPBELL, J.; GODFREY, J. y FERNÁNDEZ-CORDERO, J. (2002) “Gender-dependent phonetic refraction for speaker recognition”, *Proceedings of the ICASSP*, vol. I, 149-152.
- CARRERA I SABATÉ, J. (2001) “Algunes consideracions generals sobre l’anàlisi acústica de [e] i de [a] àtones”, *Estudis de Fonètica Experimental*, 11, 67-87.
- CASACUBERTA, F. et al. (1991) “Articulatory-acoustic correlations in coarticulatory process”, *Procesamiento del Lenguaje Natural*, 11.
- CERDÀ, R. (1972) *El timbre vocálico en catalán*, Madrid, Consejo Superior de Investigaciones Científicas, Instituto Miguel de Cervantes.
- y VEYRAT, M. (2000) “Reflexiones y paradojas en torno a la unicidad de la voz”, *Actas del I Congreso de la Sociedad Española de Acústica Forense*, Madrid.
- FARRÚS, M. y HERNANDO, J. (2005) “Hacia una sinergia metodológica en la identificación de locutores” *Filología y Lingüística. Estudios ofrecidos a Antonio Quilis*, Madrid, CSIC, 1505-1518.
- DELATTRE, P. (1951) “Physiological interpretation of sound spectrograms”, *Publications of the Modern Language Association of America*, 66, págs. 864-876.
- FERNÁNDEZ PLANAS, A. M. (1993) Estudio del campo de dispersión de las vocales castellanas, *Estudis de Fonètica Experimental*, 5, 129-162.
- JASSEM, W. y WALDEMAR, G. (2004) “Off-line classification of Polish vowel spectra using artificial neural networks”, *Journal of the International Phonetic Association*, 34, 1, 37-52.
- JIN, Q.; NAVRATIL, D. A.; REYNOLDS, J. P.; CAMPBELL, J. P.; ANDREWS, J. S. y ABRAMSON, J. S. (2003) “Combining cross-stream and the time dimensions in phonetic speaker recognition”, *Proceedings of the ICASSP*.
- KLUSACEK, D.; NAVRATIL, J.; REYNOLDS, D. A. y CAMPBELL, J. P. (2003) “Conditional pronunciation modelling in speaker recognition”, *Proceedings of the ICASSP*.

- LLAMAS BELLO, C. y V. CARDEÑOSO PAYO (1997). *Reconocimiento automático del habla. Técnicas y aplicación*, Valladolid, Universidad de Valladolid
- MARTÍNEZ CELDRÁN, E. (1986) “La velarización de la vocal neutra del catalán central causada por la proximidad de l”, *Estudis de Fonètica Experimental*, 2, 111-149.
- (1995) “En torno a las vocales del español: análisis y reconocimiento”, *Estudios de Fonética Experimental*, 7, págs. 27-44.
- MONROY, R. (1980) *Aspectos fonéticos de las vocales españolas*, Madrid, SGEL.
- POCH, D. et al. (1992) “Caracterización acústica de la variación fonética en distintos estilos”, *Procesamiento de Lenguaje Natural*, 12.
- PUIG I RIERA, J. y FREIXA I AIMERICH, J. (1990) “El camp de dispersió de les vocals catalanes des del punt de vista de la percepció”, *Estudis de Fonètica Experimental* 4, 123-146.
- QUILIS, A. (1999) “El reconocimiento de la voz en la investigación judicial. La experiencia del lingüista”, *Lectura y discurso. Estudios dedicados a Vidal Lamíquiz*, Madrid, Arco Libros.
- RAMÓN, J. L.; SPREKELSEN, C.; MARÍN, M. L.; GONZÁLEZ-ORTÍN, M. y RUBIO, M. V. (1979) “Método de estudio de las frecuencias fundamentales y análisis espectrográfico de los dos primeros formantes de las vocales castellanas emitidas por cien sujetos normales”, *Acta Otorrinolaringoscopia Española*, 30, 399-414.
- ROMERO, J. (1988) “Campos de dispersión auditivos de las vocales del castellano. Percepción de las vocales”, *Estudis de Fonètica Experimental*, 3, 181-205.
- SOLÀ I PUJOLS, J. (1988) “Efectes co-articulatoris de les consonants palatals sobre la vocal neutra del català central”, *Estudis de Fonètica Experimental*, 3, 135-156.
- STEVENS, K. (1998). *Acoustic Phonetics*. Massachusetts: MIT Press.