

CREUS QUE LA TEVA VEU ÉS ÚNICA?

Ramon Cerdà Massó

Departament de Lingüística General
Universitat de Barcelona

1. PRESENTACIÓ

1.1. *Fonologia i fonètica, com sempre*

Per assegurar la diversitat temàtica, i evitar de pas l'avorriment, enfocaré la qüestió de la fonètica forense aplicada a la identificació del locutor a partir d'una perspectiva foneticofonològica, és a dir, epistèmica, que diferirà —espero— de la que utilitzaran els meus col·legues, molt més especialitzats, que tractaran el mateix tema. I és que, de fet, la interpretació o identificació fonològica és lògicament essencial des del punt de vista comunicatiu encara que es pugui discutir sobre les modalitats i els límits del que això comporta. Es tracta d'una qüestió fonamental, que analitzarem tot seguit dintre d'una breu referència històrica especialment adreçada a fonetistes i fonòlegs clàssics en el sentit més escaritat possible.

Començaré per la citació, publicada a *La Vanguardia* el 12 de novembre del 2010, d'un especialista dotat d'una sensibilitat auditiva extrema i que es mou en un àmbit artístic ben diferent del que freqüenten almenys els fonetistes i fonòlegs, com dic, escaritats. L'entrevistat, Mark Berger, supervisor i muntador cinematogràfic de mescles sonores —amb quatre Oscars de Hollywood per *Apocalypse now*, *The right stuff*, *Amadens* i *The English patient*— afirmava: «Jo puc veure una persona i no recordar-ne el nom o la cara, però quan li sento la veu i tanco els ulls la reconec. La veu té per mi moltes més imat-

ges que no pas la pròpia imatge, i em dona molta més informació que les paraules.»

1.2. *Intuïcions i experiències*

Abans d'arribar tan lluny, considerem l'experiència habitual de qualsevol. Per descomptat, en condicions normals gairebé tothom identifica la veu no només pròpia, sinó també de molts coneguts, tant de l'entorn proper com de personatges més o menys populars de la política, l'esport, el cant o similars, sovint sobretot mitjançant canals mediàtics, com ara, en especial, la televisió. Justament per això no falten imitadors que poden arribar a reproduir amb versions entre gairebé calcades i caricaturesques la imatge característica d'alguns personatges des del timbre de veu, la prosòdia, la caracterització facial, la gesticulació... fins als discursos temàtics que els són habituals. En aquest cas, cal precisar que quan es tracta d'una caricatura, tothom aprecia l'exageració de les característiques més notables del model, sigui quin sigui l'àmbit de la imitació: fisiognomònic, tímbric, discursiu, gestual... Això comporta un supòsit amb dues manifestacions que s'impliquen mútuament. La segona és el reconeixement de la diferència entre el model i la reproducció, cosa que pressuposa necessàriament la primera, la identificació dels trets que individualitzen el model. De fet, en general la identificació, a partir de la qual cada observador experimenta sobre si mateix, és la base de la certitud,

encertada o errònia, del coneixement i de l'orientació vital.

Més encara, en condicions normals no sols s'identifica la veu d'un conegut que parla sinó també el que diu i com ho diu; per exemple, cridant, parlant amb veu baixa, rient, dubitant... Tant és així, que qualsevol d'aquests dos factors (el que es diu i com es diu) pot ser fins i tot prioritari a l'hora d'interpretar el que realment es vol dir. I és que, en efecte, podem arribar a entendre una expressió fins i tot en un sentit contrari (per exemple, «Joan és un geni») si, posem per cas, comprenem que s'ha dit amb ironia perquè l'emissor ens ha fet l'ullet en el moment de parlar.¹ En principi, però, cap d'aquestes suposicions arriba a alterar per si mateixa la capacitat d'identificar la veu de cada parlant (Cerdà i Veyrat 2000, Moure 1996).²

1.3. *Inevitables variacions*

Ara bé, res d'això impedeix que es puguin produir importants alteracions susceptibles de posar més o menys en entredit, fins a la totalitat, la certesa de la identificació. Els experts en fonètica forense solen aduir una relació considerablement llarga d'inconvenients des de molt diversos punts de vista (Andrews et al. 2002).³ A part d'alteracions fisiològiques, com una afonia, cal tenir en compte la producció de veus fàcilment forçades per caricaturitzar-ne d'altres (com en el cas dels mateixos imitadors còmics) o simplement pel desig de canviar-la fent-la

¹ És clar que si els receptors sabem que el Joan de referència és estúpid, no caldrà que l'emissor faci l'ullet perquè entenguem que la frase és irònica i que, per tant, cal interpretar-ne el sentit antonímicament.

² Cal remarcar, en tot cas, que la veu d'un mateix no es percep exactament igual quan se sent de fora estant, en una gravació.

³ En especial es refereixen a les condicions ambientals i tècniques en què es produeixen els registres i les gravacions, factors importantíssims que no interessen a l'enfocament de la qüestió que plantejo aquí.

més greu o més aguda, amb un ronc o un falset. Ben mirat —o escoltat—, no sol coincidir del tot la veu en el moment de llevar-se, quan es menja, quan s'està cansat o quan es llegeix amb veu alta, com saben prou bé aquells especialistes.

Per tot el que diem, la veu tendeix a ser característica de cadascú o, dit d'una altra manera, és un factor biomètric, una mesura biològica característica de cada individu, com l'ADN, les empremtes digitals, l'iris dels ulls, la forma de les orelles, la manera de caminar i uns quants altres factors més o menys fàcils de reconèixer amb tècniques adequades. La diferència, però —pel que ja hem dit—, és que la veu encara no es deixa identificar amb prou certitud d'una manera absolutament precisa, i sobretot inequívoca, al costat de moltes de les altres mesures biomètriques, com ara l'ADN (Rose 2002).

2. LA FONÈTICA DELS LINGÜISTES

2.1. *Digital i analògic*

La fonètica ofereix una multitud d'exploracions i aplicacions, des de la imposició de la veu per a actors, locutors, cantants... fins al tractament d'alteracions otorinolaringològiques passant per la correcció logopèdica de disfèmies, la identificació de la veu per raons de seguretat en l'accés a àmbits restringits o d'atribució de responsabilitats judicials. Aplicada a la lingüística, s'orienta sobretot a l'aprenentatge d'una llengua diferent de la d'origen o a la presentació de models de pronúncia estandarditzats o de variants dialectals, socials, històriques...

Pel que fa a les estratègies conceptuals en lingüística, fins fa poc, almenys les més clàssiques i habituals solien limitar-se a identificar els elements fònics bàsics, els fonemes, que permeten establir l'estructura gramatical i semàntica dels enunciats. La més recent ja s'endinsa en l'exploració en àmbits prosòdics

o suprasegmentals, a la recerca d'elements i paràmetres distingibles, més o menys digitalitzables.

2.2. Les primeres experiències històriques

Si fem un recorregut sumari de tot això des de l'aparició de la fonètica experimental, podem compondre una perspectiva històrica ben interessant per plantejar les estratègies i els requeriments de la fonètica forense.

Arribar a resultats més o menys fiables, tanmateix, va costar molts i molts esforços. Un dels principals punts de referència va ser l'aparició de la primitiva fonètica experimental a començaments del segle XIX i el seu desenvolupament cap a mitjan XX, amb tècniques com la quimografia, la palatografia, l'experimentació articulatòria, entre d'altres, i, més tard, l'oscil·lografia i la fonografia, en el camp de l'acústica.

D'una manera espontània, i en certa manera lògica, els principals impulsors de la fonètica experimental van prendre com a suposició indiscutible que els aparells que exploraven les característiques dels sons oferien resultats objectius, molt més precisos que no pas l'observació directa i que, en conseqüència, eren absolutament correctes. I era cert. El mal és que —curiosament, per no dir paradoxalment— els elements tradicionals i bàsics (les vocals, les consonants i les seves combinacions en síl·labes) no es deixaven identificar perquè ni sortien clarament diferenciades ni presentaven propietats prou estables en la pronúncia de diferents locutors, ni tan sols en pronúncies successives d'un mateix locutor. Com més precises eren les mesures —descobrint encara, per exemple, l'estructura de les ones complexes gràcies al teorema de Fourier—, més difícil es feia distingir-les.⁴ Tant era així que va

⁴ Bàsicament, el teorema de Fourier, presentat el 1829, permet analitzar una ona composta en la fonamental i els seus harmònics; però això sol no va per-

costar molt desfer aquesta aparent paradoxa, com demostraré més endavant.

2.3. El límit entre igual i diferent

Cal precisar que en aquells mateixos moments s'estava iniciant un auge gairebé generalitzat de l'estructuralisme lingüístic a Europa i als EUA. Tot i que presentava diverses modalitats conceptuals i metodològiques, fins i tot importants, l'estructuralisme tenia un tret distintiu en la concepció del fonema basada fonamentalment no només en la diferència fonètica entre dos sons, sinó en la seva commutabilitat en parells mínims. Així, per exemple, la diferència fonètica entre [e] (*e* tancada) i [ɛ] (*e* oberta) és commutable en català central en un context com ara /s_k/, ja que *cec* [sek] no és el mateix que *sec* [sek]. Però no ho és mai en castellà, tot i que també hi és, per exemple, entre *cien* [θjen] i *miel* [mjel].⁵ Això fa que /e/ i /ɛ/ siguin fonemes en català, però només variants fonètiques, al·lòfons, en castellà. Així es va sistematitzar la diferència entre la fonètica, per a la caracterització quantitativa o analògica dels sons, i la fonologia, per a la identificació de la transcendència dels sons en la significació dels mots, és a dir, dels elements digitalment identificables, els fonemes.⁶

L'aparició de l'espectrografia, cap als anys 50 del segle passat, va suposar una millora substancial almenys en la recerca del senyal acústic de la parla. Amb una presentació diferenciada d'elements, oferia, ni més ni

metre a l'antiga fonètica experimental del XIX i mitjan XX obtenir cap conclusió interpretativa sobre la identificació inequívoca dels sons de les llengües.

⁵ Com dic, els exemples en català són de la varietat central. Pel que fa als de castellà, no es poden commutar en un mateix context articuladori perquè la diferència entre les variants tancada i oberta del fonema /e/ només estan condicionades pel context.

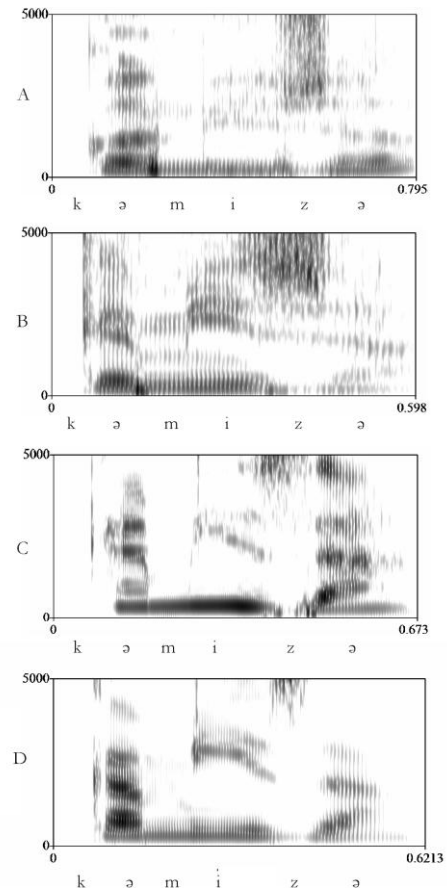
⁶ Cal no perdre de vista que, des de la perspectiva epistèmica, conceptual, la fonologia és la que controla la fonètica.

menys, la possibilitat d'aïllar automàticament la durada i la freqüència dels components acústics sense haver de recórrer al teorema de Fourier i d'obtenir-ne una imatge nítida que invitava a explorar-ne la interpretació lingüística, el repte fonamental sense resoldre que prevalia des de l'inici de la fonètica experimental. El supòsit que això era no sols possible sinó, fins i tot, fàcilment possible va fer pensar que l'espectrògraf seria, a més a més, l'eina ideal perquè els sords poguessin comunicar-se sense haver de recórrer a llenguatges de signes o a alternatives no assequibles als interlocutors.

Aquest supòsit, però, des del principi va xocar almenys amb tres problemes, dels quals el més important era de nou la interpretació de les dades. I és que els espectrogrames no identifiquen per si sols el contingut lingüístic que representen. Per obtenir aquesta identificació, l'usuari ha de fer un ensinistrament teòric i pràctic considerablement complex, entre altres raons perquè, a més a més, un mateix contingut sol presentar aspectes poc o molt diferenciats —sovint molt— a partir de la parla d'un o altre locutor (v. Ramón et al. 1979, Cerdà et al. 2003, Lucena 2007).⁷

Així es pot apreciar, posem per cas, comparant els quatre espectrogrames de la figura 1, que representen el mot *camisa* pronunciat successivament les tres primeres vegades per un home amb veu natural (A), forçada greu (B) i forçada aguda (C), i la quarta per una dona amb veu natural (D).

Figura 1. Diversos resultats espectrogràfics a partir de la paraula *camisa*



L'observació detallada d'aquests exemples ens mostra precisament fins a quin punt de diferenciació, tan espontània com analítica, es pot arribar a produir en la pronúncia d'una sola expressió, en especial tenint en compte que els tres primers espectrogrames, tan diferents, corresponen a un mateix parlant, mentre que els dos que s'assemblen més, el tercer i el quart, pertanyen a parlants diferents.

2.4. Per fi, la identificació

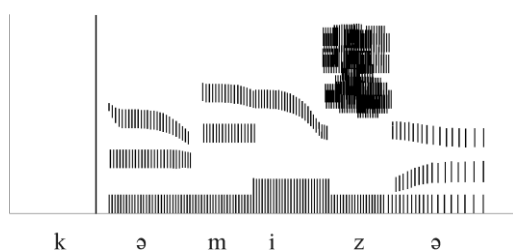
Aquest problema no es va poder resoldre en certa manera fins a l'aparició, cap a finals

⁷ Els altres problemes eren el volum i el pes de l'espectrògraf, que no feien possible transportar-lo, a part de la necessitat de mantenir-lo connectat al corrent elèctric; i el temps que es precisava per compondre un espectrograma (en condicions normals, per processar una frase de deu a dotze síl·labes es podia trigar uns deu minuts entre la gravació i l'obtenció del resultat).

dels 50, de l'anomenat *synthesizer de llenguatge*, tecnològicament un anti-espectrògraf, un aparell que processava un espectrograma i n'extreia la pronúncia original. Finalment s'obria la clau amb una perspectiva metodològica ben senzilla. Es tractava tot just de comprovar quins efectes acústico-perceptius tenia la supressió de diferents sectors dels espectrogrames per tal de localitzar, sobretot, els resultats mínims que permetien la identificació del contingut lingüístic de cada espectrograma, és a dir, entendre'l. Així es van arribar a aïllar els elements espectrogràfics que eren imprescindibles per obtenir aquella identificació, encara que el resultat de l'experiment perdés altres aspectes acústics, com ara la riquesa qualitativa de la veu resultant, convertida en una emissió com a metàl·lica i deshumanitzada, segons reconeixien els autors de les comprovacions.

Gràcies a tot això es van identificar els dos primers formants per emissions vocàliques, els espais en blanc amb barres d'explosió i transicions formàntiques per a consonants oclusives, la barra de sonoritat per als sons sonors, els sorolls inharmonics per als estridents... i així successivament; tot dintre d'unes coordenades de freqüència i de durada que permetien la distinció entre els elements de cada classe (/i/, /e/, /ε/... entre les vocals; /p/, /t/, /k/... entre les oclusives sordes...) (Delattre 1951, Jakobson et al. 1955, Rodman et al. 2002). A la figura 2 hi ha una caracterització així, en abstracte, de la pronúncia «ideal» de *camisa*.

Figura 2. Conjunt d'elements espectrogràfics bàsics que conformen la paraula *camisa*



© Ramon Cerdà (2011)

Llengua, Societat i Comunicació, 9, pàgines 32-40
<http://www.ub.edu/cusc>

2.5. Del particular al general

Fixem-nos, però, en el comentari anterior sobre les veus metàl·liques o deshumanitzades dels espectrogrames reduïts als elements bàsics. De fet, el que succeïa és que quan se n'eliminaven els elements no bàsics, segons aquella metodologia, dels resultats produïts amb veu humana es perdien d'una manera proporcional les característiques identificatives d'aquella veu. I no sols de la veu humana en abstracte sinó, encara més, les que eren específiques del locutor a qui s'havia extret l'espectrograma objecte de l'experiment (Schmidt-Nielsen i Crystal 2000).

Tanmateix, això no va representar cap inconvenient, perquè l'únic que se cercava llavors era la identificació dels elements lingüísticament bàsics per a la identificació fonològica dels sons. I com que les manifestacions concretes d'aquests elements bàsics estaven de tota manera condicionades per la veu de cada emissió, només calia reunir prou manifestacions estadísticament vàlides o acceptables en nombre i varietat per poder extraure'n els valors ideals o abstractes, de referència. Al marge, per descomptat, que sonessin deshumanitzats.

2.6. La cerca del particular, però quin?

D'acord amb el que acabo d'exposar, el problema que pretén solucionar la fonètica forense, almenys en principi, és justament el contrari al mètode exposat a 2.5: ometre la identificació del contingut lingüístic d'una elocució i trobar els indicis que permeten identificar precisament quin individu humà en particular l'ha emès. O, més en general, verificar i quantificar en quina mesura dues o més elocucions pertanyen a un mateix locutor. No cal dir que la suggestió més immediata s'orienta cap a la identificació dels trets alternatius als trets bàsics o fonològics. I de fet, això es té efectivament en compte, tot i que l'operació comporta un important in-

convenient conceptual: no s'han pogut aïllar, identificar (digitalitzar) i, per tant, establir quants són i sobretot quins són els trets, o el conjunt de trets, que individualitzen inequívocament una elocució.

Es podria pensar que aquesta actuació és correcta precisament perquè, en comptes de cercar els trets que deshumanitzen l'emissió (els estrictament fonològics), caldria eliminar-los si no fos perquè els trets fonològics contenen igualment les característiques de la veu de cada parlant. Abans ja ho hem donat a entendre quan hem dit que els trets d'identificació fonològica requereixen un tractament estadístic per determinar-ne uns valors més o menys ideals. I és que, en realitat, els trets individuals es troben per tot arreu en les emissions, i les característiques individuals depenen del conjunt o, en tot cas, de conjunts que encara no s'han pogut aïllar i identificar d'una manera prou segura com les característiques fonològiques.

2.7. Una hipòtesi més

Tenint en compte aquest marc conceptual des del punt de vista de l'anàlisi espectrogràfica, hi ha una estratègia que, almenys en teoria, sembla que ha de funcionar per a la identificació del locutor. Consisteix en el següent: partir dels valors mitjans que caracteritzen, si no tots, alguns dels trets acústics que hem indicat a 2.4. i a la figura espectrogràfica «ideal» de *camisa*, i comprovar les desviacions o alteracions (per descomptat mínimes, microscòpiques o no tant) que caracteritzen la realització de cada parlant, per determinar estadísticament, amb la màxima precisió possible, les peculiaritats quantitatives o qualitatives d'aquelles desviacions. Ara bé, per dur a terme això cal tenir en compte que les minidesviacions obeeixen a les característiques particulars del tracte vocal del locutor, però també, i molt en especial, a les nombrosíssimes condicions de coar-

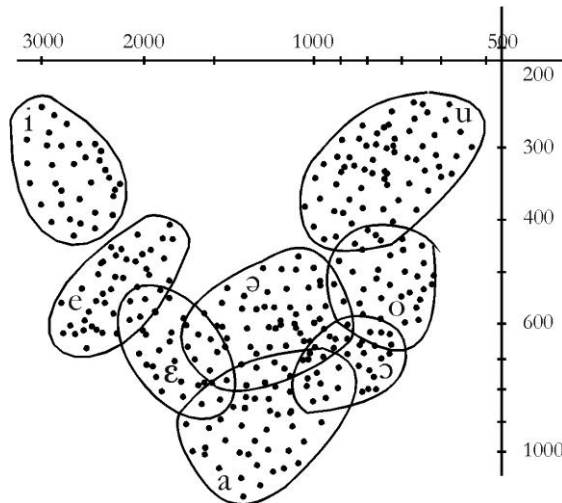
ticulació en què es troben totes les realitzacions fonemàtiques en la parla espontània, les quals, al seu torn, estan igualment condicionades pel tracte vocal del locutor. I com que no es poden aïllar ni abstraure ambdós condicionants, almenys a primera vista, la solució només pot ser estadística, és a dir, un cop comprovat un nombre significatiu de realitzacions.

A Cerdà et al. 2007 i especialment a Cerdà 2009 es posen alguns fonaments metodològics d'aquest tipus de plantejament a partir d'una estimació estadística del mesurament freqüencial dels dos primers formants vocàlics en la parla de diversos locutors. L'espai freqüencial ocupat per cada vocal en un conjunt de realitzacions vocàliques és el que es coneix com a *camp de dispersió* (a la figura 3 s'hi representen els camps a què donen lloc unes realitzacions vocàliques d'un locutor imaginari, en català oriental, similars a les adduïdes a l'article). Almenys en teoria (és a dir, quan se suposa que la quantitat de comprovacions és prou representativa des del punt de vista estadístic), cal determinar la forma i l'extensió de cada camp de dispersió, les interseccions i/o el punt central, geomètric o freqüencial (ponderat o no) de les incidències vocàliques per tal d'obtenir una característica biomètrica prou fiable del locutor. Tot plegat, almenys, amb algunes condicions: que la pronúncia sigui natural, espontània i gens alterada per cap raó orgànica o voluntària.

I encara més, els especialistes en diverses modalitats lingüístiques mai passen per alt la possibilitat de col·leccionar el màxim de propietats expressives, com ara manifestacions expletives (pauses, expressions repetitives com *No? Oi que m'entén?* i similars) o dialectals, socials, culturals... que solen reduir les característiques identificatives d'un locu-

tor encara que molt —molt— rarament solen individualitzar-lo (Moure 1996).⁸

Figura 3. Conjunt ideal de realitzacions vocàliques delimitades en els seus respectius camps de dispersió (Cerdà 1972)



3. CONCLUSIÓ

3.1. Massa disjuncions encara

Fa cosa de set o vuit anys recordo que ja es parlava d'un sistema d'exploració de veu que, a partir d'una senzilla emissió sense cap contingut lingüístic necessari, permetria reproduir gràficament i d'una manera precisa les característiques anatòmiques de les cavitats glòtiques i supraglòtiques de l'emissor. Segur que es tractaria d'una tècnica biomètrica eficient, i qui sap si definitiva.

No em consta, però, que s'hagi implementat. Mentrestant, els especialistes que intervenen en processos forenses encara solen distribuir-se en grups massa disjunts, en el sentit que actuen a partir de tècniques i fites no tan convergents com seria lògic i desitjable (Cerdà et al. 2005, Cerdà 2008):

⁸ I és que no serveix de gaire per si sola la precisió de fixar la parla d'algú com a pertanyent a un determinat àmbit dialectal, sociolingüístic o ambdós, atès que qualsevol d'aquests àmbits, per petit que sigui, caracteritza molts parlants.

a) Els enginyers del senyal. Són els que practiquen les metodologies més fiables, o almenys més notòries, des de perspectives basades sobretot en les propietats físiques de les emissions i en l'avaluació estadística de les alternatives resultants.

b) Els lingüistes. En un primer nivell, exploren totes les manifestacions que poden per tal de situar amb la màxima precisió els locutors en un àmbit dialectal, sociològic i psicològic. En un segon nivell, examinen quantitativament els trets acústics, amb estimacions estadístiques o no.

c) Els agents policials i similars. Solen combinar les metodologies anteriors donant, en tot cas, una absoluta prioritat als enginyers.

d) Els metges. Sovint es limiten a fornir informació sobre les alteracions que la veu pot experimentar a causa d'una munió de possibles disfuncions especialment orgàniques.

e) Els advocats. Solen queixar-se per la manca de laboratoris particulars on es puguin obtenir fàcilment i lliurement comprovacions alternatives a les designades pels jutges.

f) Els magistrats. Insisteixen a declarar que la certesa d'una conclusió pericial, per molt alta que sigui, no pot donar lloc per ella mateixa a una sentència judicial.

Urgeix que els membres d'aquells col·lectius no sols uneixin les seves aportacions, sinó que n'afegeixin encara d'altres, com, per exemple, les de la psicologia.

O ens mentalitzem i ens entrenem fins a atènyer la perícia del senyor Mark Berger, que cito a 1.1, o esperem que aviat hi hagi almenys algun canvi significatiu, des de perspectives inevitablement multi-disciplinàries (Broeders 2002).

3.2. I una altra paradoxa

Només hi veig un problema paradoxal. I és que, com més s'afinin l'anàlisi i la caracterització de la individualitat de la veu, més oberta quedarà la porta a simulacions i falsificacions.

No és difícil d'imaginar un sistema que a partir d'un enregistrament capturi les característiques fòniques d'un parlant i que a continuació es teclegi una frase (com ara, *Ja tinc un testafarro per al diner negre*) que soni d'una manera aparent, o molt i molt aparent, amb la mateixa veu. Fa anys que s'experimenta.

AGRAÏMENTS

Per acabar, només em falta agrair ben de cor les oportunes observacions que m'han fet Mireia Farrús, Ana Maria Fernández Planas i Gavina Freixa.

4. REFERÈNCIES BIBLIOGRÀFIQUES

- ANDREWS, W.; KOHLER, M.; CAMPBELL, J.; GODFREY, J.; FERNÁNDEZ-CORDERO, J. (2002). «Gender-dependent phonetic refraction for speaker recognition». *Proceedings of the ICASSP*, 1, 149-152.
- BROEDERS, A. P. A. (2002). «Forensic speech and audio analysis: the state of the art in 2000 AD». *Acústica Forense*, SEAF, 1, 33-42.
- CERDÀ, R. (1972). *El timbre vocàlic en catalán*. Collectanea Phonetica, Madrid: CSIC.
- CERDÀ, R.; VEYRAT, M. (2000). «Reflexiones y paradojas teóricas en torno a la unicidad de la voz». *Libro de Actas del I Congreso de la Sociedad Española de Lingüística Forense*. Madrid: Universidad Politécnica, 135-144.
- CERDÀ, R.; MUÑIZ, C.; VEYRAT, M. (2003). «Lo esencial y lo accesorio en la composición individual de la voz. Propuesta para su identificación acústica». *Actas del II Congreso de la Sociedad Española de Acústica Forense*. Barcelona: Universitat Politècnica de Catalunya, 71-78.
- CERDÀ, R.; FARRÚS, M.; HERNANDO, J. (2005). «Hacia una sinergia metodológica en la identificación de locutores». *Filología y Lingüística. Estudios ofrecidos a Antonio Quilis*. Madrid: CSIC, 1505-1518.
- CERDÀ, R.; FARRÚS, M.; HERNANDO, J.; VEYRAT, M. (2007). «Propuesta de experimento sobre la noción de campo de dispersión fonemática». *Actas do III Congreso da Sociedade Espanhola de Acústica Forense*, 127-139.
- CERDÀ, R. (2008). «Sobre alguns aspectes contraposats en fonètica forense». *Estudios de Fonética Experimental*, XVII, 45-64.
- CERDÀ, R. (2009). «Camps de dispersió vocàlica en imitacions de veu: primers indicis d'un experiment sobre identificació de locutor». *Estudios de Fonética Experimental*, XVIII, 65-88.
- DELATTRE, P. (1951). «Physiological interpretation of sound spectrogram». *Publications of the Modern Language Association of America*, 66, 864-876.
- JAKOBSON, R.; FANT, G.; HALLE, M. (1955). *Preliminaries to Speech Analysis. The Distinctive Features and their Correlates*. Cambridge, Massachusetts: MIT Press.
- LUCENA MOLINA, J. J. (2007). «Errores de interpretación en la valoración de la prueba científica». *III Congreso da Sociedade Espanhola de Acústica Forense*. Xunta de Galicia, 83-111.
- MOURE, T. (1996). *La alternativa no-discreta en lingüística. Una perspectiva histórica y metodológica*. Santiago de Compostela: Universidade de Santiago de Compostela.
- RAMÓN, J. L.; SPREKELSEN, C.; MARÍN, M. L.; GONZÁLEZ-ORTÍN, M.; 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.
- RODMAN, R. et al. (2002). «Forensic speaker identifications based on spectral moments». *Forensic Linguistics*, 1, 9.

ROSE, P. (2002). *Forensic Speaker Identification*, UK and N America, Taylor and Francis.
Ressenya de JESSEN, M. (2003). *Forensic Linguistics*, 1, 10, 138-151.

SCHMIDT-NIELSEN, A.; CRYSTAL, T. H. (2000). «Speaker verification by human listeners: experiments comparing human and machine performance using the NIST 1998 Speaker Evaluation Data». *Digital Signal Processing*, 10, 1-3, 249-266.