

**MÈTODES DE NORMALITZACIÓ I DE REPRESENTACIÓ DE  
DADES ACÚSTIQUES I ARTICULATÒRIES**

DANIEL RECASENS VIVES  
*Universitat Autònoma de Barcelona*  
Daniel.Recasens@uab.cat

### **RESUM**

Aquest treball revisa alguns mètodes de normalització i de representació de dades acústiques i articulatòries. Pel que fa als mètodes de normalització de dades acústiques, el mètode Lobanov redueix de forma més satisfactòria que el mètode CILH les dades F1 x F2 de cinc, sis i set vocals de dialectes majors i menors del català. També s'apunten alguns mètodes de normalització i de representació de dades de contacte linguopalatal i de moviment lingual obtingudes mitjançant l'electropalatografia i l'articulometria midsagital electromagnètica.

Paraules clau: *mètodes de normalització, formants, dialectes del català, EPG, EMA.*

### **ABSTRACT**

The present study reviews several normalization methods of acoustic and articulatory data. It is shown that the Lobanov normalization procedure performs better than CILH when F1 x F2 frequency data for five, six and seven vowel systems of major and minor Catalan dialects are taken into consideration. Several ways of normalizing and representing linguopalatal contact and tongue movement data obtained by means of electropalatography and electromagnetic midsagittal articulometry are also discussed.

Keywords: *normalization methods, formants, Catalan dialects, EPG, EMA.*

## 1. DADES ACÚSTIQUES

El procés de normalització de dades acústiques de freqüència resulta d'aplicació indispensable quan es tracta de comparar valors absoluts de parlants de procedència dialectal diversa. Això és així per tal com les dimensions del tracte vocal de parlants diferents, siguin o no del mateix sexe, difereixen entre si. Així doncs, en el cas d'una comparació dialectal, si no portéssim a terme aquest procés de normalització barrejaríem les característiques dels parlants individuals amb les característiques pròpies dels dialectes.

Segons Adank (2003), els mètodes de normalització de freqüències dels formants vocàlics més satisfactoris són el mètode Lobanov de transformació en puntuació típica o *z-score* i el mètode de Nearey anomenat CILH o *Constant Log Interval Hypothesis*. L'aplicació de la transformació de Lobanov, a saber,  $(F_i - \mu_i) / \sigma_i$ , comporta restar les freqüències de F1 o de F2 d'una vocal determinada de la freqüència del mateix formant amitjanada a través de totes les vocals del sistema, i dividir el resultat obtingut per la desviació típica corresponent a aquesta mitjana global. La transformació mitjançant el mètode CILH, i.e.,  $F_{N[V]s}^* = G_{N[V]s} - G_{N[,]s}$ , calcula el valor normalitzat del formant *N* de la vocal *V* del parlant *s* restant el seu logaritme natural de la mitjana dels logaritmes naturals de les freqüències del mateix formant corresponents a totes les vocals del sistema.

Per tal d'estudiar les possibles diferències entre els dos mètodes de normalització, comparem els valors no normalitzats i normalitzats de F1 i F2 de les vocals dels dialectes catalans Rossellonès i Oriental o Central i dels subdialectes Gironí i Sitgetà del dialecte Oriental. Els valors normalitzats han estat obtinguts amb aplicació dels dos mètodes CILH i Lobanov. Els quatre parlars catalans difereixen pel que fa al nombre de parelles de vocals mitjanes neutralitzades, a saber, dues parelles en els sistemes de cinc vocals del Sitgetà i del Rossellonès (/i, E, a, O, u/), una parella en el sistema de sis vocals del Gironí (/i, e, ε, a, O, u/) i cap parella en el sistema de set vocals del dialecte Oriental (/i, e, ε, a, o, u/). La mostra de dades analitzades correspon a les set vocals /i, e, ε, a, o, u/ llegides per quatre o cinc parlants masculins de cada parlar inserides en mots i en frases breus amb sentit. Les parelles de vocals que experimenten neutralització en un parlar determinat, e.g., /e/ i /ε/ en Sitgetà, foren llegides també pels seus parlants.

Els valors de F1 i F2 no normalitzats i normalitzats de les vocals són presentats a la figura 1.

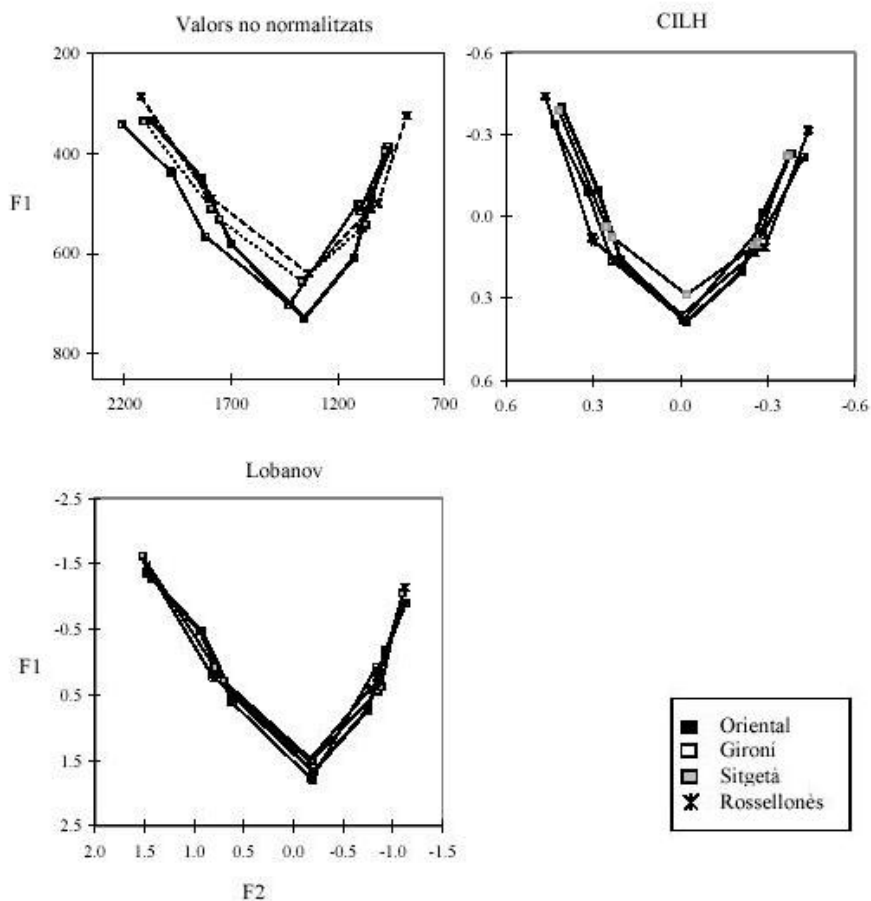


Figura 1. *Espais vocàlics corresponents als parlars català Oriental, Gironí, Sitgetà i Rossellonès, amb valors de F1 i F2 no normalitzats en Hz i normalitzats segons els mètodes CILH i Lobanov.*

Segons la figura 1, els valors dels membres de les parelles de vocals mitjanes neutralitzades es troben prop un de l'altre en l'espai acústic si bé no se superposen

necessàriament. Aquest és el cas, per exemple, de /e/ i /ɛ/ i de /o/ i /ɔ/ en Sitgetà. Per altra banda, els valors dels formants d'una mateixa vocal en els diferents parlars s'aproximen entre si en major mesura si han estat normalitzats que si no ho han estat. Així, les vocals anteriors del Gironí resten més allunyades de les dels altres parlars en el gràfic superior esquerre de la figura que en els altres dos gràfics. A més a més, el grau d'aproximació és superior si la normalització ha estat portada a terme amb el mètode Lobanov que si ho ha estat amb el mètode CILH. Per altra banda, l'espai vocàlic obtingut amb el mètode Lobanov és més similar al de les vocals no normalitzades i, per tant, més natural que l'obtingut amb el mètode CILH. Aquesta diferència es deu al fet que la fórmula corresponent al mètode CILH opera amb logaritmes.

## 2. DADES ARTICULATÒRIES

Ocupem-nos ara de dades articulatòries i, més concretament, de dades de contacte linguopalatal obtingudes amb el sistema d'electropalatografia de Reading. Cal indicar que les fórmules de reducció de dades d'emplaçament o de grau de contacte (e.g., com ara les corresponents al mètode d'índex de contacte; Fontdevila et al., 1994) operen ja sobre valors normalitzats. Efectivament, malgrat que els paladars artificials dels parlants individuals puguin diferir en forma i tamany, aquelles fórmules s'apliquen a les files i columnes d'elèctrodes, que es troben emplaçades en un mateix indret en el cas de tots els parlants. Així, per exemple, la fila 4 d'elèctrodes resta sempre localitzada a la zona postalveolar independentment que les dimensions de la zona alveolar siguin menors o majors. En conseqüència, atès que les distàncies entre files i entre columnes són proporcionalment equivalents, els valors dels índex resulten comparables entre parlants.

El problema en aquest cas no està en els valors numèrics dels índex, sinó en la representació gràfica dels palatogrames per tal com llurs impressions ocupen el mateix espai sobre paper amb independència del tamany real dels paladars. Per tal de solucionar aquest problema portàrem a terme a Recasens (1991) representacions bidimensionals de patrons de contacte linguopalatal com les de la figura 2 tot prenent en consideració les distàncies reals dels paladars artificials dels parlants. Aquestes distàncies s'adiuen, sobre paper, amb la llargada dels costats dels quadrilàters corresponents als elèctrodes, a saber, amb llur llargada vertical que correspon a la distància real entre una fila d'elèctrodes i la fila immediatament anterior o posterior, i amb llur llargada horitzontal que correspon a la distància real

entre una columna d'elèctrodes i la columna emplaçada immediatament a la seva esquerra o a la seva dreta. Les configuracions de contacte dels palatogrames A i C i les dels palatogrames B i D de la figura són idèntiques però, en realitat, el percentatge de contacte sobre la superfície total del paladar és superior en el palatograma A que en el palatograma C i en el palatograma B que en palatograma D atès que l'àrea no contactada de la zona palatal és superior en els palatogrames C i D que en els palatogrames A i B. En el cas que aquests patrons de contacte pertanyessin a parlants de dialectes diferents, llurs diferències no serien dialectals sinó pròpies dels parlants individuals.

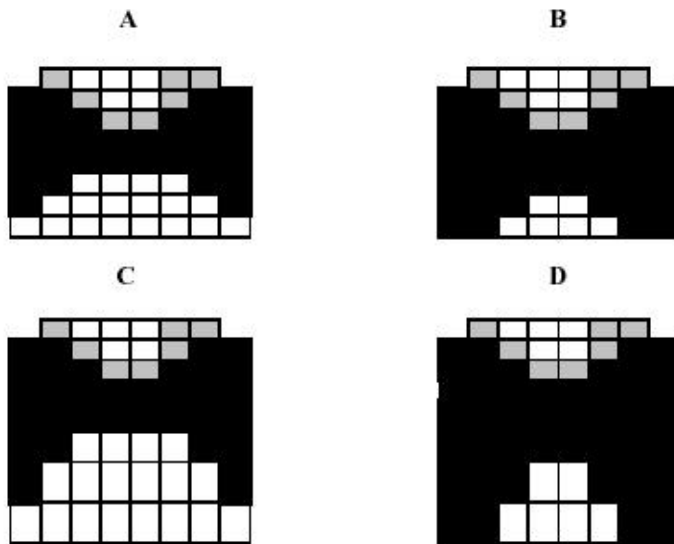


Figura 2. Patrons de contacte linguopalatal que difereixen pel que fa a la configuració de contacte ( $A=C$ ;  $B=D$ ) i a les dimensions dels paladars artificials ( $C, D > A, B$ ). Els electròdes dels paladars s'agrupen en files i columnes, i difereixen pel que fa a la freqüència d'activació.

Una altra representació relativa al contacte linguopalatal amb distàncies normalitzades és inclosa a la figura 3. Les dades representades en el gràfic corresponen a la freqüència del pic espectral del soroll de l'explosió de /k/ del Mallorquí en diferents contextos vocàlics i en les tres posicions inicial de mot

(#CV), intervocàlica (VCV) i final de mot (VC#), segons cinc parlants del dialecte. Cal recordar que la /k/ mallorquina pot realitzar-se oclusiva (alveolo)palatal o velar segons una distribució al·lofònica ben descrita en els manuals de dialectologia catalana. Aquests valors es troben situats en un punt equivalent a la distància en mm entre la fila d'elèctrodes on té lloc l'oclusió i la fila d'elèctrodes més anterior d'un paladar artificial. L'espai articulatori comú a tots els parlants permet de comparar llurs dades entre si. En concret, el gràfic revela que la freqüència del pic augmenta a mesura que l'indret de l'oclusió esdevé més anterior en les files d'elèctrodes 6, 7 i 8 de la zona palatal. Canvis de grau d'anterioritat en files més anteriors no promouen variacions ostensibles de la freqüència del pic.

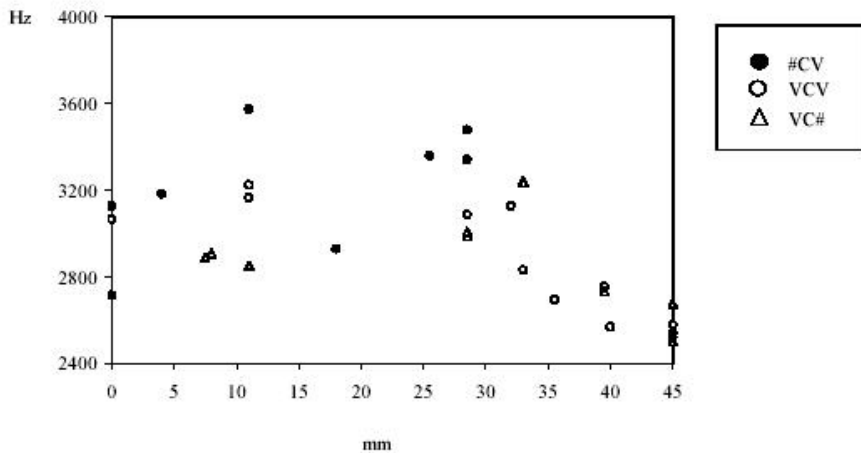


Figura 3. Freqüències del pic espectral del soroll de l'explosió corresponents a realitzacions de /k/ en diferents posicions i contextos segons cinc parlants mallorquins. Els valors freqüencials són representats en funció de la distància entre la fila d'elèctrodes d'un paladar artificial on té lloc l'oclusió i la fila més anterior del mateix paladar.

Un problema amb aquests dos tipus de representació és que la projecció de les dades articulatòries hauria de ser tridimensional per tal com el moviment dels articuladors en el si del tracte vocal no només té lloc al llarg dels eixos longitudinal

i transversal horitzontal sinó també al llarg de l'eix transversal vertical. Representacions tridimensionals com les de la figura 4 reproduïxen millor que les representacions bidimensionals les distàncies reals en l'espai articulari delimitat pel paladar artificial (vegeu Fuchs et al. 2006). La figura mostra dos paladars artificials de formes i dimensions diferents, a saber, un paladar més llarg i amb una bòveda més pronunciada i un altre més curt i més pla. El contacte dorsopalatal resulta suposadament més fàcil d'obtenir en el segon paladar que en el primer mentre que la facilitat per assolir contacte en les files anteriors d'elèctrodes podria ser la mateixa en ambdós casos.

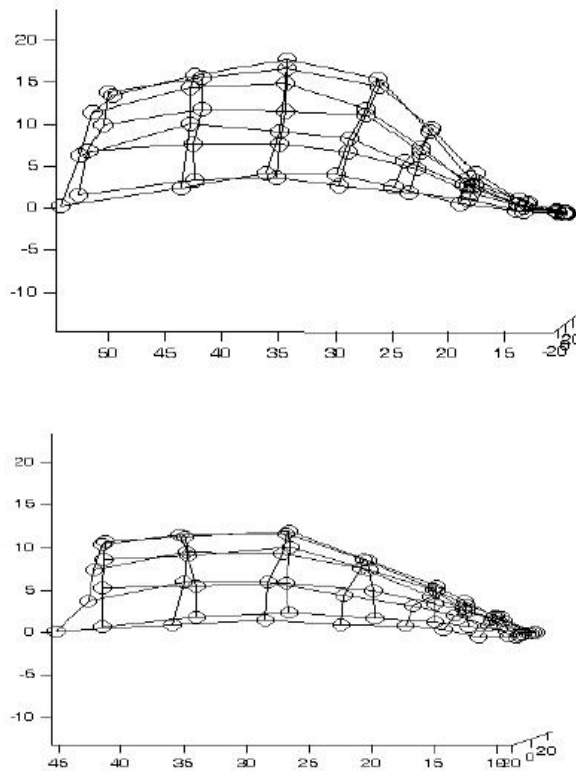


Figura 4. Reproducció tridimensional de dos paladars artificials de diferents formes i dimensions. Els valors dels tres eixos longitudinal, transversal vertical i transversal horitzontal són en mm.



Una altra qüestió és la reconstrucció del perfil superior de la llengua a partir de les dades de contacte linguopalatal i la superposició d'aquest perfil sobre diagrames tridimensionals de les mateixes característiques que els reproduïts en la figura 4. Hom ha provat de reconstruir la configuració del contorn lingual a partir de dades de palatografia directa (Ladefoged, 1957).

Finalment, cal també comentar mètodes de normalització i de representació de dades de moviment articulari obtingudes amb la tècnica EMA (articulometria midsagital electromagnètica). Pel que fa a la representació de dades, cal esmentar la possibilitat de reconstruir la configuració del sostre lingual en un moment temporal determinat unint la posició dels sensors analitzats, e.g., els sensors apical, laminal i dorsal. Cho (2004) presenta indicacions útils pel que fa a la normalització de valors de desplaçament articulari corresponents a parlants diferents i, per tant, obtinguts en espais articularis diferents. En el cas que considerem les tres vocals /i, a, u/, les distàncies articulatòries relatives a parlants diferents poden expressar-se com a percentatges sobre les distàncies entre els valors extrems de X (pla horitzontal) i de Y (pla vertical) de cada parlant, és a dir, sobre les distàncies entre els valors mínim de X per a /i/ i màxim de X per a /u/ i entre els valors mínim de Y per a /a/ i màxim de Y per a /i/. (L'origen dels eixos X i Y es troba a l'alçada de les dents incisives). Així, un valor absolut de X de 15 mm en relació a una distància entre 7.5 mm (valor mínim de /i/) i 22.5 mm (valor màxim de /u/) resultaria equivalent a un valor absolut de X de 20 mm en relació a una distància entre 10 i 30 mm. En ambdós casos l'emplaçament dels dos valors de X seria un 50% de la distància entre els valors extrems considerats, amb la qual cosa podria assegurar-se que en ambdues condicions la llengua està emplaçada a mig camí entre un extrem i l'altre.

### 3. REFERÈNCIES BIBLIOGRÀFIQUES

- ADANK, P. (2003): *Vowel Normalization. A Perceptual-Acoustic Study of Dutch Vowels*, Wageningen, Ponsen and Looijen.
- CHO, T. (2004): «Prosodically conditioned strengthening and vowel-to-vowel coarticulation in English», *Journal of Phonetics*, 32, pp. 141-176.
- FONTDEVILA, J.; M. D. PALLARÈS i D. RECASENS (1994): «The contact index method of EPG data reduction», *Journal of Phonetics*, 22, pp. 141-154.

FUCHS, S.; P. PERRIER; C. GENG i C. MOOSHAMMER (2006): «What role does the palate play in speech motor control?. Insights from tongue kinematics for German alveolar obstruents», in J. Harrington and M. Tabain (eds.): *Speech Production, Models, Phonetic Processes, and Techniques*, Psychology Press, pp. 149-164.

LADEFOGED, P. (1957): «Use of palatography», *Journal of Speech and Hearing Disorders*, 22(5), pp. 764-774.

RECASENS, D. (1991): «El projecte ESPRIT-ACCOR», *A Sol Post*, 3, pp. 237-246.