

Identificació de les vocals tòniques del català.

Mètode d'identificació mitjançant les freqüències relatives dels intervals $F_0 \sim F_1$, $F_1 \sim F_2$ i $F_2 \sim F_3$ i la freqüència d' F_0 de les vocals.

Sílvia Planas Morales
Universitat Rovira i Virgili

Resum

Aquesta recerca desenvolupa un mètode d'identificació de les vocals que controli les variacions entre parlants. Es basa en relacionar els trets distintius amb la distància acústica entre els formants vocàlics, la qual es mesura de manera logarítmica --que és com percep el so el sistema auditiu-- i no pas de manera lineal. La novetat que introdueix és la unitat de mesura: la freqüència relativa. El mètode es va aplicar inicialment al sistema vocàlic tònic del català central, però també es va aplicar al sistema vocàlic de l'anglès americà per fer una anàlisi comparativa. Finalment es va fer un test de percepció de vocals sintetitzades segons els paràmetres de freqüència relativa de cada classe de vocals que vam obtenir a partir del mètode d'identificació.

PARAULES CLAU: vocals, anàlisi acústica, percepció, trets distintius.

Resumen

En esta investigación se desarrolla un método de identificación de las vocales que controle las variaciones entre hablantes. Se basa en la relación de los rasgos distintivos con la distancia acústica de los formantes vocálicos, que se calcula de forma logarítmica --tal y como percibe el sistema auditivo el sonido-- y no de forma lineal. La novedad que introduce es la unidad de cálculo: la frecuencia relativa. El método se aplicó inicialmente al sistema vocálico tónico del catalán central, pero también se aplicó al sistema vocálico del inglés americano para realizar un análisis comparativo. Finalmente se realizó un test de percepción de vocales sintetizadas siguiendo los parámetros de frecuencia relativa para cada clase de vocales que obtuvimos gracias al método de identificación.

PALABRAS CLAVE: vocales, análisis acústico, percepción, rasgos distintivos.

Abstract

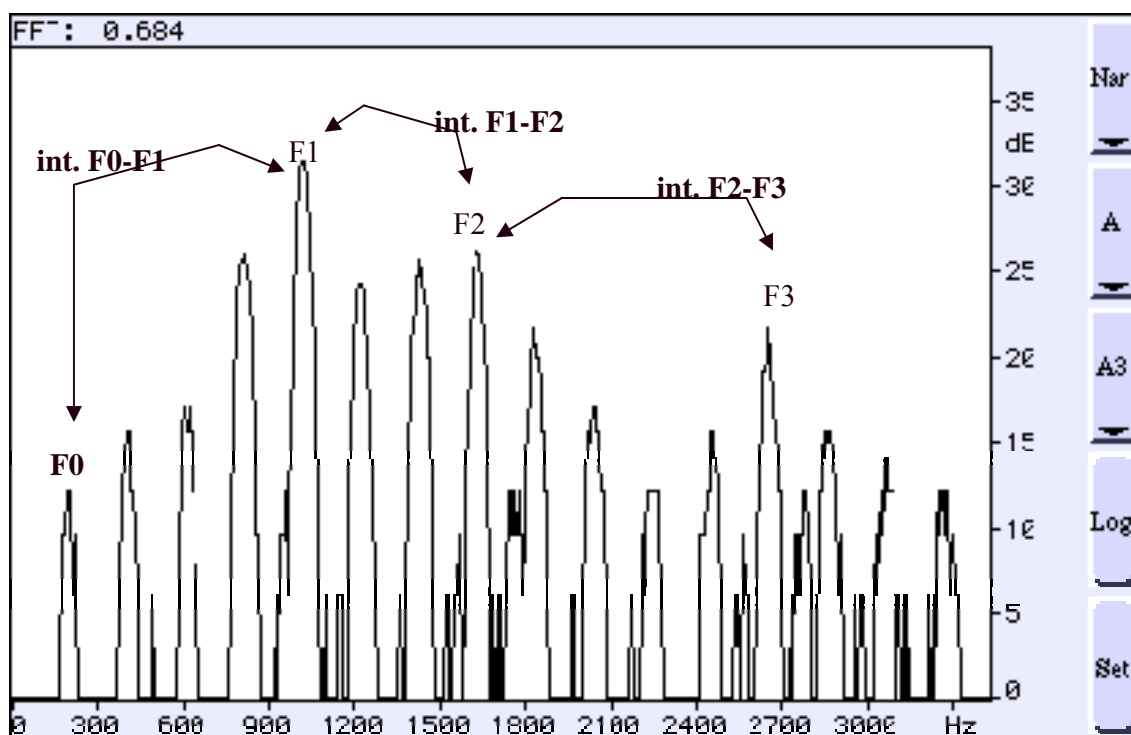
This research develops a method for vowel identification capable of controlling individual variation. It is grounded on the relation between distinctive features and the acoustic distance of vowel formants, which is obtained through a logarithmic scale --which is how the auditory system perceives sound-- and not through a linear scale. The main innovation concerns the unit of measure: relative frequency. The method was initially applied to the tonic vocalic system of Central Catalan, but it was also tested against the vocalic system of American English for comparative purposes. Finally, a perception test was conducted of synthesized vowels following the parameters of relative frequency for each vowel class which were obtained from the identification method.

KEY WORDS: vowels, acoustic analysis, perception, distinctive features

Introducció

L'objectiu de la recerca d'aquesta tesi és demostrar que una vocal es pot identificar a partir del seu F0 i de les freqüències relatives dels intervals F0~F1, F1~F2 i F2~F3. Anomenem intervals a la distància espectral entre F0 i F1, entre F1 i F2 i entre F2 i F3 (gràfic 1). L'elecció de la freqüència relativa (fR) és per dos motius:

1. És la unitat de mesura que dóna la relació de vibració entre dos tons o harmònics; és a dir, entre dos formants i entre F0 i qualsevol formant, la qual cosa ens permet no haver de dependre de divisions arbitràries en funció del sexe i l'edat.
2. El sistema auditiu no percep de manera lineal, sinó logarítmica.



Gràfic 1. Intervals F0~F1, F1~F2 i F2~F3 a l'espectre.

Calculem les distàncies espectrals en freqüències relatives (fR), la qual cosa ens permet accomplir els dos requisits que Disner (1980) exigeix als models de normalització:

- Maximitzar les diferències entre les categories vocàliques.
- Minimitzar les diferències entre parlants.

La investigació s'ha desenvolupat dins el marc del model de percepció de la teoria de la Invariança Acústica (Stevens i Blumstein, 1981), que es resumeix en tres punts:

- (1) La invariança acústica, corresponent a una categoria fonètica en particular o a un tret distintiu, resideix en el senyal acústic.
- (2) La invariança no és pas derivable d'una anàlisi dels components individuals del senyal acústic, sinó que prové de característiques acústiques integrades que poden abraçar diferents components.

En el cas de les vocals, els components individuals característics són les freqüències dels formants i la freqüència fonamental, i la invariança de cada vocal prové de la distribució característica dels formants en l'espai espectral (Fant, 1960). La informació, per tant, de la invariança de les vocals s'ha de buscar en la manera en què es combinen els formants. En aquesta tesi entenem que les característiques acústiques integrades de les vocals són la relació de vibració (fR) entre F0 i F1, entre F1 i F2 i entre F2 i F3, obtinguda a partir de l'espectre promediat sobre una mostra del segment acústic de la vocal que pot oscil·lar entre els 50ms-80ms.

- (3) La naturalesa d'aquestes característiques invariants i la seva relació amb els segments acústics reflecteixen de quina manera:
 - (a) el mecanisme articulatori delimita els sons possibles d'una llengua,
 - (b) el mecanisme perceptiu delimita els sons possibles d'una llengua i
 - (c) els trets defineixen els grups i classes de sons de la llengua.

L'objectiu final d'Stevens (com es pot comprovar a Stevens 1998) és el de postular classes de vocals i consonants naturals amb correlats acústics, articulatoris i perceptius ben definits. Els trets que utilitza Stevens són descrits a Halle (1990). En certa manera, Stevens justifica els trets distintius de base articulatòria (Halle 1990) quan argumenta els correlats acústics (distribució dels formants) i articulatoris (posició i forma dels articuladors) de les configuracions del tracte vocal. Per la nostra banda, la classificació fonològica les vocals tòniques del català emprada en aquest treball és la de Bonet i Lloret (1998:27).

	□	□	□	□	□	□	□
alt	+	-	-	-	-	-	+
baix	-	-	-	+	-	-	-
pos	-	-	-	+	+	+	+
AA	+	+	-	-	-	+	+
lab	-	-	-	-	+	+	+

Metodologia

■ Corpus

El corpus d'aquesta tesi és la base de dades amb l'F0 i les freqüències relatives dels intervals F0-F1, F1-F2 i F2-F3 de mostres de vocals tòniques catalanes analitzades prèviament.

Obtenció de les dades per al corpus

El primer pas per portar a terme la recerca era obtenir les freqüències absolutes d'F0, F1, F2 i F3 de les set vocals tòniques del català. Per aquest motiu es va elaborar un corpus de 36 paraules per obtenir cinc mostres vàlides de □□□□□□□□ tòniques per informant. Les paraules són del tipus 'CV₁CV₂, 'CV₁, 'V₁C i 'CV₁C, on V₁= [i, e, ε, a, ə, o, u].

**bata, be, bé, bota, cec, os, ós, pala, papa, pe, pela, pes, pila, pipa,
pis, poma, por, puma, pupa, ras, rec, res, ric, roc, ros, ruc, sac,
sec, sic, soc, sóc, suc, tos, tus, ven, vent.**

Aquestes paraules es van presentar als informants escrites en diferents formats i agrupades en dos tipus de corpus:

- A. Corpus de parells mínims.
- B. Corpus amb frase marc.

Un cop obtingudes les mostres, es va procedir a l'anàlisi acústica del segment estable de les vocals per extreure'n la freqüència en Hz d'F0, F1, F2 i F3. De cada mostra se'n va fer l'espectre promediat, l'LPC (*Lineal Prediction Code*) i l'estadística de l'F0, amb el programa SoundScope/16.

La base de dades del corpus

El veritable corpus d'aquesta tesi és, per tant, la base de dades amb les freqüències (Hz) d'F0, F1, F2 i F3 de les set vocals tòniques del català central. Les mostres seleccionades són de tres dones i tres homes, d'entre 20 i 35 anys, de nivell universitari.

■ Càlcul de la freqüència relativa dels intervals

Anomenem interval a la distància espectral entre formants adjacents i entre F0 i F1. Per obtenir una escala de freqüències relatives que aquesta distància, partim de l'anomenada *escala natural* o *d'harmònics*, a on 1 és el primer harmònic (F0) i 2 (=2/1) la seva octava i el primer harmònic de l'escala següent, de manera que l'escala es pot concatenar d'octava en octava.

L'octava és el límit de l'escala perquè tot el que excedeix la seva extensió només és la rèplica d'ella mateixa. Les fraccions que completen l'escala indiquen la relació de

vibració que existeix entre el primer harmònic i els altres harmònics que es troben dins de l'escala.

- La fracció $3/2$ indica que mentre el primer harmònic (F_0) realitza 3 cicles per segon, l'harmònic que es troba en aquest interval de l'escala en fa 2.

Per expressar que aquestes fraccions indiquen la freqüència relativa entre dues prominències espectrals s'hi ha adjuntat l'abreviatura fR (freqüència Relativa).

Taula 1. Concatenació d'escapes, essent el primer harmònic (F_0) = 200Hz

1 (F_0)	9/8	5/4	4/3	3/2	5/3	15/8	2	fR
200	225	250	267	300	333	375	400	Hz
2	2·9/8	2·5/4	2·4/3	2·3/2	2·5/3	2·15/8	4	fR
400	450	500	533	600	667	750	800	Hz
4	4·9/8	4·5/4	4·4/3	4·3/2	4·5/3	4·15/8	8	fR
800	900	1000	1067	1200	1333	1500	1600	Hz
8	8·9/8	8·5/4	8·4/3	8·3/2	8·5/3	8·15/8	16	fR
1600	1800	2000	2133	2400	2667	3000	3200	Hz

Si considerem que 1 és F_0 i concatenem les escapes per octaves (taula 1), cada fracció indicarà matemàticament la relació vibratòria del formant vocàlic amb l' F_0 . Traslladat a les vocals, aquest mecanisme ens permet calcular directament la distància entre F_0 i els formants, (si una vocal amb F_0 a de 200Hz té un formant a 1200Hz, la relació de vibració entre el to fonamental i el formant és de $4·3/2$).

Per obtenir la freqüència relativa de l'interval $F_1\sim F_2$ cal dividir la freqüència relativa de l'interval $F_0\sim F_2$ per la freqüència relativa de l'interval $F_0\sim F_1$. I per obtenir la freqüència relativa de l'interval $F_2\sim F_3$ s'ha de dividir la freqüència relativa de l'interval $F_0\sim F_3$ per la freqüència relativa de l'interval $F_0\sim F_2$. El càlcul de les freqüències relatives dels intervals $F_1\sim F_2$ i $F_2\sim F_3$ va generar una escala de freqüències relatives molt més gran¹, perquè no pas sempre els quebrats d'aquests intervals es corresponien amb els de l'escala natural (taula 2).

¹ Fins el punt que apareixen tots els intervals de l'escala musical: 16/15 (segona menor), 9/8 (segona major), 6/5 (tercera menor), 5/4 (tercera major), 4/3 (cuarta), 3/2 (quinta), 2/1 (octava).

Taula 2. Escala de les freqüències relatives, ordenada per octaves (en negreta, les fR corresponents a l'escala inicial)

1 ^a octava	2 ^a octava	3 ^a octava	4 ^a octava
1	2	4	8
16/15	2·16/15	4·16/15	8·16/15
10/9	2·10/9	4·10/9	8·10/9
9/8	2·9/8	4·9/8	8·9/8
32/27	2·32/27	4·32/27	8·32/27
6/5	2·6/5	4·6/5	8·6/5
5/4	2·5/4	4·5/4	8·5/4
4/3	2·4/3	4·4/3	8·4/3
27/20	2·27/20	4·27/20	8·27/20
45/32	2·45/32	4·45/32	8·45/32
64/45	2·64/45	4·64/45	8·64/45
40/27	2·40/27	4·40/27	8·40/27
3/2	2·3/2	4·3/2	8·3/2
8/5	2·8/5	4·8/5	8·8/5
5/3	2·5/3	4·5/3	8·5/3
27/16	2·27/16	4·27/16	8·27/16
16/9	2·16/9	4·16/9	8·16/9
9/5	2·9/5	4·9/5	8·9/5
15/8	2·15/8	4·15/8	8·15/8
2	4	8	16
fR	fR	fR	fR

Quan en el primer interval ($F_0 \sim F_1$) la freqüència relativa és més petita que $2fR$, se li assigna el valor $1fR$ perquè per sota de $2fR$, que és l'octava del primer harmònic de la relació, la freqüència relativa marca relacions de vibració insignificants.

$$(1) F_0 = 210\text{Hz}, F_1 = 300\text{Hz}$$

$$210\text{Hz} \times 2fR = 420\text{Hz};$$

$$300\text{Hz} - 210\text{Hz} = 90\text{Hz}$$

$$420\text{Hz} - 300\text{Hz} = 120\text{Hz}$$

$$\text{Per tant, } F_0 \sim F_1 = 1fR$$

Quan hi ha molt poca distància entre dues freqüències, el que es percep és una freqüència intermèdia, a causa del fenomen perceptiu del centre espectral de gravetat (Chistovich i Lublinskaya, 1979). Aquests casos, que es poden comprovar per observació visual en un gràfic de LPC, es donen a l'interval $F_0 \sim F_1$ de la vocal [i] i a l'interval $F_1 \sim F_2$ de les vocals [u] i [a]. Stevens (1998:273) comenta que les freqüències (Hz) baixes d' F_1 produeixen respostes auditives que sincronitzen amb la freqüència d' F_0 . En conseqüència i per simplificar el mètode, s'opta per donar el valor $1fR$ a les relacions de vibració de l'interval $F_0 \sim F_1$ que surten per sota de $2fR$.

Resultats

Atès que un bon model d'identificació de vocals ha de tenir solucionada la variabilitat d'F0 per poder minimitzar les diferències entre parlants, es va ordenar cada vocal per F0 descendent i es va observar en els següents límits d'F0 es produïa un canvi significatiu de la freqüència relativa de l'interval F0~F1.

- F0 > 200 Hz (to de veu alt)
- F0: 125 Hz – 200 Hz (to de veu mitjà)
- F0 < 125 Hz (to de veu baix)

Al **Quadre de normalització per freqüències relatives** es recopilen les freqüències relatives dels intervals F0~F1 (per extensions d'F0), F1~F2 i F2~F3 que caracteritzen i identifiquen les vocals tòniques del català. Les freqüències relatives que marquen els límits dels trets distintius són octaves de l'escala natural (taula 3)

Quadre de Normalització per freqüències relatives de les vocals tòniques del català					
	F0~F1			F1~F2	F2~F3
	F0 > 200Hz	F0: 125Hz-200Hz	F0 < 125Hz	F0 = x (F0~F1=4fR)	F0 = x
i	1 – 2	2	2·3/2	4·3/2 – 8·5/3	< 2
e	2	2 – 2·3/2	4	4 - 4·4/3	< 2
ɛ	2·3/2 – 4	4 – 4·5/4	4·5/4 – 4·3/2	2·5/4 – 2·3/2	< 2
a	4 – 4·5/4	4·5/4 – 4·5/3	≥ 4·5/3	< 2·5/4	< 2
ɔ	2·3/2 – 4	4 – 4·5/4	4·5/4 – 4·3/2	< 2·5/4 (< 2)	2 – 2·5/3
o	2·3/2	2·3/2	4	≤ 2·5/4 (≥ 2)	2 – 2·5/3
u	2	2	2·3/2	≤ 2·5/4	2 – 4·5/3
	fR	fR	fR	fR	fR

Taula 3. Freqüències relatives de l'escala natural ordenades per octaves (dins els quadrats taronja, les fR que marquen límits en els trets distintius)

1 ^a octava	1	9/8	5/4	4/3	3/2	5/3	15/8	(2)	fR
2 ^a octava	2	2·9/8	2·5/4	2·4/3	2·3/2	2·5/3	2·15/8	(4)	fR
3 ^a octava	4	4·9/8	4·5/4	4·4/3	4·3/2	4·5/3	4·15/8	8	fR

■ **Resultats dels intervals de freqüència relativa per vocal**

[i]
$F_0 \sim F_1 = 1fR - 2fR$, quan $F_0 > 200\text{Hz}$ $F_0 \sim F_1 = 2fR$, quan $F_0: 125\text{Hz} - 200\text{Hz}$ $F_0 \sim F_1 = 2 \cdot 3/2 fR$, quan $F_0 < 125\text{Hz}$
$F_1 \sim F_2 = 4 \cdot 3/2 fR - 8 \cdot 5/3 fR$
$F_2 \sim F_3 < 2fR$

[u]
$F_0 \sim F_1 = 2fR$, quan $F_0 > 200\text{Hz}$ $F_0 \sim F_1 = 2fR$, quan $F_0: 125\text{Hz} - 200\text{Hz}$ $F_0 \sim F_1 = 2 \cdot 3/2 fR$, quan $F_0 < 125\text{H}$
$F_1 \sim F_2 \leq 2 \cdot 5/4 fR$
$F_2 \sim F_3 = 2 fR - 4 \cdot 5/3 fR$ $F_2 \sim F_3 > 2 \cdot 5/3 fR$, quan $F_0 \sim F_1 = 2 \cdot 3/2 fR$

[e]
$F_0 \sim F_1 = 2fR$, quan $F_0 > 200\text{Hz}$ $F_0 \sim F_1 = 2fR - 2 \cdot 3/2 fR$, quan $F_0: 125\text{Hz} - 200\text{Hz}$ $F_0 \sim F_1 = 4fR$, quan $F_0 < 125\text{Hz}$
$F_1 \sim F_2 = 4fR - 4 \cdot 4/3 fR$
$F_2 \sim F_3 < 2fR$

[ε]
$F_0 \sim F_1 = 2 \cdot 3/2 fR - 4fR$, quan $F_0 > 200\text{Hz}$ $F_0 \sim F_1 = 4fR - 4 \cdot 5/4 fR$, quan $F_0: 125\text{Hz} - 200\text{Hz}$ $F_0 \sim F_1 = 4 \cdot 5/4 fR - 4 \cdot 3/2 fR$, quan $F_0 < 125\text{Hz}$
$F_1 \sim F_2 = 2 \cdot 5/4 fR - 2 \cdot 3/2 fR$
$F_2 \sim F_3 < 2fR$

[ɔ]	[o]
$F_0 \sim F_1 = 2 \cdot 3/2fR - 4fR$, quan $F_0 > 200\text{Hz}$ $F_0 \sim F_1 = 4fR - 4 \cdot 5/4fR$, quan $F_0: 125\text{Hz} - 200\text{Hz}$ $F_0 \sim F_1 = 4 \cdot 5/4fR - 4 \cdot 3/2fR$, quan $F_0 < 125\text{Hz}$	$F_0 \sim F_1 = 2 \cdot 3/2 fR$, quan $F_0 > 200\text{Hz}$ $F_0 \sim F_1 = 2 \cdot 3/2fR$, quan $F_0: 125\text{Hz} - 200\text{Hz}$ $F_0 \sim F_1 = 4fR$, quan $F_0 < 125\text{Hz}$
$F_1 \sim F_2 = < 2 \cdot 5/4fR$ $F_1 \sim F_2 = < 2fR$, quan $F_0 \sim F_1 = 4fR$	$F_1 \sim F_2 \leq 2 \cdot 5/4fR$ $F_1 \sim F_2 \geq 2fR$, quan $F_0 \sim F_1 = 4fR$
$F_2 \sim F_3 = 2fR - 2 \cdot 5/3fR$	$F_2 \sim F_3 = 2fR - 2 \cdot 5/3fR$

[a]
$F_0 \sim F_1 = 4fR - 4 \cdot 5/4fR$, quan $F_0 > 200\text{Hz}$ $F_0 \sim F_1 = 4 \cdot 5/4fR - 4 \cdot 5/3fR$, quan $F_0: 125\text{Hz} - 200\text{Hz}$ $F_0 \sim F_1 = \geq 4 \cdot 5/3fR$, quan $F_0 < 125\text{Hz}$
$F_1 \sim F_2 = < 2 \cdot 5/4fR$
$F_2 \sim F_3 < 2fR$

Discussió

■ Els correlats acústics dels trets distintius

L'interval F0~F1, les extensions d'F0 i els trets [alt] i [baix]

Els resultats demostren que:

- (a) No hi ha una fR límit, sinó que hi ha una distribució de les fR relacionada amb l'obertura de les vocals.

Recerques com les de Traunmüller (1981), amb les vocals del Bavarès, i Syrdal (1985), amb vocals de l'anglès americà, avalen el paper cabdal de la distància entre F0 i F1 en la percepció de l'obertura de la vocal. Cal fer constar que si només agafem les tres vocals extremes, sí que trobem unes freqüències relatives límits en la línia dels correlats acústics trobats per als altres trets. Així tenim que si el català només tingués les tres vocals extremes, $4fR$ a l'interval F0~F1 marcaria el límit entre les vocals [+alt, -baix] i la vocal amb els trets [-alt, +baix]. En qualsevol cas, F0~F1 és l'interval predominant de la vocal posterior no labialitzada.

- (b) Els resultats han demostrat que l'interval F0~F1 depèn de l'extensió de to veu a la qual pertany l'F0 de la vocal.
- F0 > 200 Hz (to agut)
 - F0: 125Hz - 200Hz (to mitjà)
 - F0 < 125Hz (to greu)

Els resultats de Di Benedetto (1994), amb vocals sintetitzades de l'anglès americà, corroboren que hi ha una extensió de veu compresa entre 125Hz i 200Hz, en la qual les variacions del to fonamental no varien la identificació de les vocals, mentre que sí que hi ha una identificació diferent si es canvia l'F0 a una freqüència superior a 200Hz. D'altra banda, és més lògic classificar l'F0 per extensions de veu i no pas pel sexe i edat dels parlants, si tenim la capacitat d'analitzar l'evolució d'F0 d'un segment inicial de per extreure'n una estimació del seu valor base com a referència, com defensen algunes teories auditives (Traunmüller, 1998).

L'interval F1~F2 i els trets [posterior] i [arrel avançada] ([AA])

Per a qualsevol F0 ($F0 = x$), les vocals posteriors tenen l'interval F1~F2 per sota de $2.5/4fR$ i les vocals anteriors el tenen igual i per sobre d'aquesta freqüència relativa. Aquests resultats indiquen que $2.5/4fR$ a l'interval F1~F2 és el límit de domini del tret [posterior]. Tant Fant (1983) com Ladefoged (1993) relacionen la distància entre F1 i F2 amb el tret [posterior].

A més, les freqüències relatives límit que s'han trobat als intervals F1~F2 de les vocals /i, e, ε/ permeten classificar-les també pel tret [alt] i pel tret [AA], amb la qual cosa, podem concloure que l'interval més significatiu per a la identificació de les vocals [-pos] del català és l'interval F1~F2. El fet que l'interval F1~F2 distingeix perfectament aquestes tres vocals del català fa pensar que, encara que la identificació d'una vocal

depengui de la combinació de la freqüència relativa dels tres primers intervals, hi hauria un interval predominant que jugaria el paper més important en la identificació d'una classe de vocals.

L'interval F2~F3 i el tret [labial]

Un altre correlat acústic que funciona per a qualsevol valor d' F_0 i per a totes les vocals és l'interval F2~F3 i el tret [labial]. Segons els resultats obtinguts, $2fR$ a l'interval F2~F3 marca el límit del tret [labial] i com més labialitzada és la vocal més alta és la freqüència relativa. Les vocals [o] i [ɔ] no passen de $2\cdot5/3fR$, mentre que [u] arriba a $4\cdot5/3$ (una octava més).

Stevens (1998:292) explica que d'una parella de vocals que ocupen la mateixa posició, tindrà la freqüència d' F_2 més baixa la que és labial. És a dir, si l' F_2 baixa, la distància entre F2 i F3 augmenta en la vocal [+lab] i, per aquest motiu, és lícit establir el correlat acústic del tret [labial] a l'interval F2~F3.

■ Verificació del mètode

Per comprovar-ne la validesa, es va aplicar el mètode a la síntesi de vocals catalanes (experiment 1) i a l'anàlisi de vocals angleses (experiment 2).

Experiment 1

Es van sintetitzar les set vocals tòniques del català amb el sintetitzador de Klatt, que inclou, per defecte, el senyal acústic de la vocal neutra teòrica. Els valors de freqüència en hertz per a F1, F2 i F3 es van extreure del càlcul dels intervals a partir de tres freqüències fonamentals, una de cada extensió d' F_0 .

- $F_0 = 112\text{Hz}$ (extensió $F_0 < 125\text{Hz}$). Arxius: 01, 02, 03, 04, 05, 06 i 07
- $F_0 = 160\text{Hz}$ (extensió $F_0: 125\text{Hz} - 200\text{Hz}$). Arxius: 08, 09, 10, 11, 12, 13, 14 i 15.
- $F_0 = 210\text{Hz}$ (extensió $F_0 > 200\text{Hz}$). Arxius: 16, 17, 18, 19, 20 i 21.

Les vocals sintetitzades² es van ordenar de forma aleatòria en el test de percepció que es va passar a 31 catalans nadius. El percentatge d'identificació positiva supera el 75% excepte a la vocal [ɔ]. Les vocals extremes no s'han pas confós entre elles. Les errades en la identificació han estat entre graus i punts de constricció adjacents, a excepció de la vocal [ɛ] que no s'ha pas confós amb [a] (taula 4).

Els resultats per trets distintius (taula 5) demostren que les freqüències relatives dels intervals assignades als límits de domini dels trets són les adequades, perquè el percentatge de discriminació del tret supera el 90% en tots els casos.

² Si hi ha dificultats per obrir els arxius de *.wav amb un programa normal d'àudio, es poden escoltar amb el PCquiner, Macquiner i el Praat.

Taula 4. Percentatges d'identificació per vocals

		VOCALS IDENTIFICADES						
		i	e	ɛ	a	ɔ	o	u
VOCALS SINTETITZADES	i	100%						
	e	5%	80%	15%				
	ɛ		12%	88%				
	a			9%	84%	7%		
	ɔ				29%	55%	16%	
	o					16%	77%	7%
	u						7%	93%

Taula 5. Resultats del test de percepció per trets distintius

TRET DISTINTIU	DISCRIMINACIÓ
[alt]	97%
[baix]	94%
[pos]	99%
[AA]	92%
[lab]	95%

Experiment 2

Es va aplicar el mètode d'identificació per freqüències relatives dels intervals a les vocals de l'anglès americà que comparteixen trets distintius amb les del català. Atès que el sistema vocàlic de l'anglès americà és més complex, es van escollir les vocals [i, ɛ, a, ɔ, u], les quals comparteixen amb les vocals [i, ɛ, a, ɔ, u] del català els mateixos valors pel que fa a l'altura, posterioritat i labialització i, d'aquesta manera, comprovar els límits quàntics dels trets [posterior], [labial] i t [alt].

Els corpus són les mitjanes de la base de dades de Peterson i Barney (1952) que apareixen a Baken (1991:358). Es va escollir aquest corpus per comparar els resultats amb els de Syrdal i Gopal (1986), els quals també fan servir una escala logarítmica per calcular la distància espectral entre formants: la banda crítica de 3Bark (taula 6).

Taula 6. Classificació de les vocals [i, ε, α, ə, u] de l'anglès americà basada en la diferència crítica de 3 Bark de les distàncies espectrals (dimensions) entre F0 i F1, entre F1 i F2 i entre F2 i F3, segons Syrdal i Gopal (1986).³

Vocals	Dimensions		
<i>Peterson i Barney (1952)</i>	<i>F1—F0</i> □ 3 Bark	<i>F2—F1</i> □ 3 Bark	<i>F3—F2</i> □ 3 Bark
□	+	—	+
□	—	—	+
□	—	+	—
ə		+	—
□	+	—	—

Taula 7. Classificació de les vocals [i, ε, α, ə, u] de l'anglès americà basada en la freqüència relativa de les distàncies espectrals (interval) entre F0 i F1, entre F1 i F2 i entre F2 i F3.

Vocals	Intervals		
<i>Peterson i Barney (1952)</i>	<i>F0~F1</i> ≤ 2 fR	<i>F1~F2</i> ≤ 2·3/2 fR	<i>F2~F3</i> > 2·5/4 fR
i	+	—	+
ε	—	—	+
α	—	+	+
ə	—	+	—
u	+	+	—

Segons els seus resultats, seria la distància espectral entre F2 i F3 (dimensió F3—F2), i no pas la d'entre F1 i F2 (dimensió F2—F1), la que marcaria el punt d'articulació de les vocals, essent [+pos] aquelles vocals que entre F2 i F3 tinguessin la distància espectral superior a 3 Bark (taula 4.5). Segons els nostres resultats (resumits a la taula 7), la distància espectral entre F1 i F2 (interval F1~F2) és la que marca el punt d'articulació de les vocals ([±pos]), mentre que la distància entre F2 i F3 (interval F2~F3) marca la posició del llavis, la qual cosa corrobora els resultats obtinguts amb les vocals tòniques del català i argumenta que el mètode que defensem en aquesta tesi és aplicable a llengües de famílies diferents.

³ [ə] no presenta cap valor +/- en la dimensió F1—F0, perquè, per a Syrdal i Gopal (1986), aquesta vocal no es distingeix en aquesta dimensió.

■ **Recapitulació**

Funció del mètode d'identificació

Establir les freqüències relatives que marquen el límit de domini dels trets distintius

El senyal acústic està fortament estructurat pels correlats acústics dels trets (Blumstein 1986).

Interval F0~F1

- a) *Distribució relacionada amb l'obertura de les vocals.*
- b) *Depèn de l'extensió de to de veu a la qual pertany l'F0 de la vocal:*
 - F0 < 125Hz (to greu)
 - F0: 125Hz-200hz (to mitjà)
 - F0 > 200Hz (to agut)

Interval F0~F1 i les vocals extremes

i, u	a
$F0 \sim F1 \leq 2 \cdot 3/2 fR$	$F0 \sim F1 \geq 4 fR$
[+alt, -baix]	[-alt, +baix]

F0~F1

interval predominant de la vocal [-alt, +baix]

Interval F1~F2

- Per a $F_0 = x$

$F_1 \sim F_2 \rightarrow 2.5/4fR$ límit del tret [posterior]

$F_1 \sim F_2 \leq 2.5/4fR \rightarrow$ vocal [+pos]

$F_1 \sim F_2 > 2.5/4fR \rightarrow$ vocal [-pos]

Interval F1~F2 i les vocals anteriors

vocals [-pos]	[alt]	[AA]
[i] $\rightarrow F_1 \sim F_2 \geq 4 \cdot 3/2fR$	+	+
[e] $\rightarrow F_1 \sim F_2 : 4fR - 4 \cdot 4/3fR$	-	+
[ε] $\rightarrow F_1 \sim F_2 < 4fR$	-	-

F1~F2

interval predominant de les vocals [-pos]

Interval F2~F3

- Per a $F_0 = x$

$F_2 \sim F_3 \rightarrow 2fR$ límit del tret [labial]

$F_2 \sim F_3 > 2fR \rightarrow$ vocal [+lab]

$F_2 \sim F_3 < 2fR \rightarrow$ vocal [-lab]

Conclusions

D'aquesta tesi s'extreuen dues conclusions principals:

- 1. Les vocals es poden identificar a partir de l'F0 i de les freqüències relatives dels intervals F0~F1, F1~F2 i F2~F3.**

$$\text{vocal} = \text{F0 (Hz)} + \text{F0~F1 (fR)} + \text{F1~F2 (fR)} + \text{F2~F3 (fR)}$$

Hi ha uns correlats acústics amb els trets distintius que funcionen per a qualsevol valor d'F0 i per a totes les vocals tòniques del català. El tret [posterior] és el tret dominant de l'interval F1~F2 perquè el seu límit és molt clar i no depèn de cap condició: el canvi quàntic es produeix a $2 \cdot 5/4fR$. Passa el mateix amb el tret [labial] i l'interval F2~F3: el canvi quàntic es produeix a $2fR$.

A l'interval F0~F1, que és on es classifiquen les vocals segons la seva obertura, els límits de domini dels trets [alt] i [baix] no són simples d'establir i depenen de la freqüència del to fonamental (F0). Però si només prenem en consideració les vocals extremes, $2 \cdot 3/2fR$ marca el límit de les vocals [+alt, -baix] i $4fR$ marca el límit de la combinació distintiva [-alt, +baix]

D'altra banda, tot i que la propietat invariant que identifica una vocal és la combinació de la freqüència relativa dels tres intervals, n'hi ha un que predomina per damunt dels altres en la identificació i discriminació de la vocal. Així l'interval F1~F2 és el paràmetre acústic que més important en la identificació de les vocals [i, e, ε] l'interval F2~F3 ho és per a les vocals [ɔ, o u], i l'interval F0~F1, per a la vocal [a].

- 2. Calcular les distàncies espectrals en freqüència relativa i classificar el to fonamental per extensions de veu permet eliminar les variacions entre parlants.**

D'aquesta manera, la identificació perceptiva de la vocal es interpreta segons els referents de la vocal (les freqüències relatives) i no pas els valors absoluts (els hertz) dels paràmetres acústics.

El càlcul logarítmic és més adequat, atès que la relació entre la realitat física i l'auditiva és logarítmica i no pas lineal.

La classificació de les extensions de veu fa possible que les variacions del to fonamental dins l'extensió no afectin la percepció de l'alçada de la vocal perquè el pas d'una extensió de veu a una altra comporta la redistribució de la freqüència relativa de l'interval F0~F1 de totes les vocals. A més, un mateix parlant pot tenir distribuïda la seva tessitura entre dues extensions o gammes tonals, que es distribuïrien de la manera següent:

- F0 > 200 Hz (to de veu alt)
- F0: 125 Hz – 200 Hz (to de veu mitjà)
- F0 < 125 Hz (to de veu baix)

Potser la manera més sintètica de resumir el mètode d'identificació que defenso en aquesta Tesi sigui els dos requisits que imposa Disner (1980) als models de normalització de les vocals, perquè és un mètode que:

- Maximitza les diferències entre categories vocàliques.
- Minimitza les diferències entre parlants.

Referències bibliogràfiques

BAKEN, R.J. (1991). *Clinical Measurement of Speech and Voice*. San Diego, California: College-Hill Press.

BONET, E. i LLORET, M.R. (1998). *Fonologia catalana*. Barcelona: Ariel.

DI BENEDETTO, M. G. (1994). Acoustic and perceptual evidence of a complex relation between F1 and F0 in determining vowel height. *Journal of Phonetics*, 22: 205-224.

DISNER, S. (1980). Evaluation of vowel normalization procedures. *JASA*, 67: 253-261.

FANT, G. (1960). *Acoustic theory of speech production*. La Haya: Mouton [cito per l'edició de 1970].

FANT, G. (1983). Feature analysis of Swedish vowels—a revisit. *Speech Transmission Laboratory: Q.Prog.Stat.Rep. 2-3: 1-19*.

HALLE, M. (1990). Features. Dins de W.Bright (ed.) *International encyclopedia of linguistics*. New York: Oxford University Press. 1-19.

LADEFOGED, P. (1975). *A Course in Phonetics*. Fort Worth: Harcourt Brace College Publishers [cito per l'edició de 1993].

STEVENS, K. N. (1998). *Acoustic Phonetics*. Cambridge, Mass.: MIT Press.

SYRDAL, A. K. (1985). Aspects of a model of the auditory representation of American English vowels. *Speech Communication*, 4: 121-135.

SYRDAL, A. K. i GOPAL, H. S. (1986). A perceptual model of vowel recognition based on the auditory representation of American English vowels. *JASA*, 79: 1086-1100.

TRAUNMÜLLER, H. (1981). Perceptual dimension of openness in vowels. *JASA*, 69: 1465-1475.

TRAUNMÜLLER, H. (1998). The role of F0 in vowel perception. www.ling.su.es/staff/hartmut/i.htm.