

EL CAMP DE DISPERSIO DE LES VOCALS CATALANES
DES DEL PUNT DE VISTA DE LA PERCEPCIO.

JOAQUIM PUIG I RIERA
JUDIT FREIXA I AYMERICH

1. INTRODUCCIO I OBJECTIUS.

El primer objectiu d'aquest treball és l'estudi del camp de dispersió de les vocals catalanes des del punt de vista de la percepció. Amb aquest fi s'ha realitzat una enquesta amb sons vocàlics sintetitzats. Diversos autors han publicat estudis sobre l'articulació de les vocals catalanes: Cerdà (1972), Martí (1983), Recasens (1984), i Llisterrí (1984). Cap d'aquests estudis cerca de delimitar els camps de dispersió sinó que un cop obtingut un cert nombre de mostres efectuen càlculs amitjadors per fer una proposta concreta per a cada vocal. D'altra banda, tots aquests autors mencionats parteixen del punt de vista de l'articulació (excepte Llisterrí que el combina amb el de la percepció).

El segon objectiu deriva del primer: veure en quina mesura s'adopten les nostres propostes per a un nucli del camp de dispersió de les vocals catalanes amb les d'aquests autors que acabem d'esmentar.

Un dels objectius de tot estudi d'aquestes característiques és el suport a la síntesi de veu. Creiem que amb aquest treball, pel fet de partir d'un sintetitzador i de la percepció a través de l'enquesta podem aportar molt a la síntesi de veu ja que de cada so sintetitzat podem afirmar que es percep com a determinada vocal en un percentatge determinat.

2. ASPECTES PREVIS.

2.1. El sintetitzador.

En aquest treball hem utilitzat el sintetitzador de veu MEA 8000 del Laboratori de Fonètica del Departament de Lingüística de la Universitat de Barcelona. El sistema editor de viu que comanda el sintetitzador està constituït per un ordinador personal Apple II i un programa editor dissenyat per Miquel Julián i Lanuza. Per a la creació de sons, el sintetitzador MEA 8000 té els següents paràmetres que cal controlar: F1 (el primer formant), F2 (segon formant), F3 (tercer formant), BW1 (amplada de banda del primer formant), BW2 (del segon formant), BW3 (del tercer formant), BW4 (del quart formant), AM (amplitud), PI (pitch inicial) i IP (increment de pitch).

El sintetitzador ofereix, per a cada un d'aquest paràmetres, un nombre determinat de valors possibles. Els sons que hem utilitzat en l'enquesta del nostre treball, produïts pel sintetitzador que hem descrit, tenen uns paràmetres constants i uns altres paràmetres variables. Cada so té 20 "trames" de 8 ms cadascuna. Els paràmetres que hem mantingut constants són: F3 = 2842 Hz, BW1 = 50, BW2 = 50, BW3 = 125, BW4 = 726, IP = 0, AM = 8, 9, 10, 11, 12, 13, 14 (vuit vegades), 13, 12, 11, 10, 9 i 8 (per a les vint trames del so).

Els paràmetres variables són F1 i F2. Els valors possibles que permet el MEA 8000 per a aquest paràmetres són els que apareixen a continuació:

Formant 1		Formant 2	
1...150 Hz	17...440	1...440 Hz	17...1110
2...162	18...466	2...466	18...1174
3...174	19...494	3...494	19...1254
4...188	20...523	4...523	20...1337
5...202	21...554	5...554	21...1428
6...217	22...587	6...587	22...1528
7...233	23...622	7...622	23...1639
8...250	24...659	8...659	24...1761
9...267	25...698	9...698	25...1897
10...286	26...740	10...740	26...2047
11...305	27...784	11...784	27...2214
12...325	28...830	12...830	28...2400
13...346	29...880	13...880	29...2609
14...368	30...932	14...932	30...2842
15...391	31...988	15...988	31...3105
16...415	32...1047	16...1047	32...3400

2.2. El triangle vocàlic.

El timbre vocàlic ve caracteritzat pel primer i el segon formant. Per tant, variant el primer i/o el segon formant però mantenint els altres paràmetres obtindrem sons vocàlics amb timbre diferent amb una variació molt lleu i gradual.

Si combinem les 32 possibilitats del primer formant amb les 32 possibilitats del segon formant, obtenim 704 possibles sons vocàlics diferents que por crear el MEA 8000 mantenint els altres paràmetres constants.

Vàrem dibuixar una carta vocàlica amb els 704 sons possibles, posant F1 a l'eix de les ordenades i F2 a l'eix de les abcisses. Set-cents quatre estímuls eren massa per a un test de percepció i, a més, era evident que no calia crear tots aquest sons: hi havia zones que es podien despreciar. En primer lloc, vam procedir a crear tots els sons possibles que quedessin englobats dins els triangles vocàlics proposats per Cerdà, Martí, Recasens i Llisterrí i els sons circumdants diverses vegades fins que topàvem amb sons no recognoscibles o molt impresentables.

A continuació, unes observacions sobre les modificacions que ha sofert aquest triangle:

1. La [u] proposada per Cerdà queda fora de la zona explorada perquè és desmesuradament baixa.

2. S'ha ampliat la zona alta del primer formant (zona de [a]) fins molt per dessota de la proposta d'altres autors, perquè s'hi obtenia una qualitat de so encara molt bona.

3. S'ha despreciat l'àrea a partir de la possibilitat 28 del segon formant (2400 Hz) perquè se sobraposa al tercer formant.

4. La resta de sons de zones no incloses en el test foren descartats després que els autors hi efectuassin repetides proves.

3. METODOLOGIA.

3.1. Creació de sons en fitxers.

Una vegada establerts els punts que havíem de cobrir, vam començar a crear els estímuls. El fet d'haver partit des del principi de la consideració només dels dos primers formants ens va facilitar molt la feina a l'hora de crear un sol fitxer sencer (tots tres formants, amplada de banda, pitch inicial, surada, etc.); vam crear la resta de fitxers demanant aquest primer fitxer i canviant el primer formant o el segon o tots dos segons el que convenia. En total vam crear 526 fitxers que anàvem batejant amb una *I* i el número corresponent de les coordenades (així, per exemple, el fitxer I 23-16 fa referència a un so amb 622 Hz en el primer formant i 1047 en el segon). Enregistràvem cada fitxer que creàvem.

3.2. Desordenació i enregistrament de l'enquesta.

El següent pas en la preparació de l'enquesta va ser la desordenació dels fitxers i l'enregistrament en cintes magnetofòniques (amb màxima atenció per evitar la pèrdua d'algun so). Els 526

sons van haver de ser distribuïts en cinc grups (no podem fer enquestes de més de 150 sons i esperar un alt grau de fiabilitat en les respostes).

Vam fer, doncs, cinc llistes de sons desordenats en un bloc i, amb el sintetitzador connectat a l'ordinador, vam anar demanant cada fitxer un per un i el produïem tres vegades; al mateix temps, enregistràvem aquests sons i els ratllàvem en el bloc. El resultat van ser cinc cintes magnetofòniques, cada una de les quals conté un test de 105 sons enregistrats per triplicat i amb una pausa més llarga cada 10 sons on s'indica el número de so a què correspon dins l'enquesta per evitar, d'aquesta manera, que els enquestats tinguin algun problema i desplacin les seves respostes. Aquestes cintes amb l'enquesta estan a disposició del Laboratori de Fonètica de la UB, juntament amb els disquets que contenen els fitxers.

3.3. L'enquesta.

3.3.1. Característiques de l'enquesta.

El model d'enquesta amb què vam treballar eren uns fulls dividits en tantes caselles com sons sentirien els enquestats. Per a més precaucions, les caselles estaven numerades. Tal com vam explicar a l'inici de cada enquesta, a cada casella s'hi havia de posar una de les següents informacions:

1. La transcripció fonètica del so que sentien i un *I* quan consideressin que el so tenia gran claredat.

2. La transcripció fonètica del so i un *II* quan el so no podia considerar-se clar perceptivament, però creien reconèixer-lo.

3. La xifra *III* (sense cap transcripció fonètica) quan el so no es reconeixia com a so de la llengua. En aquesta tercera resposta s'hi trobaven tres tipus de possibilitats: sons que no pertanyen a la parla humana, sons que pertanyen a la parla humana però no a la nostra llengua i, finalment,

sons que pertanyen a la nostra llengua però que són excessivament incomprensibles.

3.3.2. Els enquestats.

Malgrat el fet de partir d'una quantitat d'estímuls tan alta (526 estímuls), vam considerar que calia obtenir trenta respostes de cada un. Els grups dels enquestats amb els qual vam treballar van ser sempre superiors a aquest nombre per tal de poder eliminar posteriorment les enquestes que presentaven problemes (alguna persona de parla únicament castellana, algun cas de no comprensió de que havia de respondre, etc.) i quedar-nos amb les 30 més fiables.

Com ja hem dit anteriorment, l'enquesta estava dividida en cinc tests. Quatre d'aquest tests van ser escoltats i respostos pels joves estudiants de 1er, 3er de BUP i COU de l'Institut de Batxillerat Sants les Corts, i el cinquè test va ser respost per estudiants de la Facultat de Filologia.

Els únics requisits que es demanaven eren coneixements auditius i perceptius de les vocals catalanes; en el full del test els subjectes eren interrogats sobre la seva llengua materna i la procedència, amb la qual cosa vam poder comprendre alguns fenòmens ja que, si bé la majoria dels enquestats eren de llengua materna catalana, el coneixement de la seva procedència ens explicava, per exemple, les confusions de l'obertura de les vocals d'un informant de subdialecte gironí.

4. TRACTAMENT I REPRESENTACIO DE DADES.

4.1. Recompte.

Un cop tinguérem 30 respostes per a cadascun dels 526 estímuls vam procedir al recompte. Donades les característiques de l'enquesta, tenim 17 tipus de resposta per a cada estimul:

i I, i II, e tancada I, e tancada II, e oberta I, e oberta II, a I, a II, vocal neutra I, vocal neutra II, o oberta I, o oberta II, o tancada I, o tancada II, u I, u II, i III.

Després de sumar els tipus de resposta obtenim uns resultats d'aquest tipus, per exemple:

Estímul 92: 12 aI, 6 aII, 7 o oberta I, 7 o oberta II, i vocal neutra II, i III.

4.2. Representació gràfica.

Els resultats de l'enquesta ens proporcionen un corpus de dades certament voluminós. Existeixen moltes formes possibles d'aprofitar aquestes dades mitjançant representacions gràfiques i tractaments estadístics, però en el present treball ens restringim a només una d'aquestes possibilitats i deixem esperant altres possibilitats de treballar.

El primer pas fou la confecció d'unes coordenades immenses amb F1 a les abscisses i F2 a les ordenades, en què a cada punt corresponent a un so o estímul sintetitzat de l'enquesta hi escrivíem el resultat de la suma dels tipus de resposta obtingut, tal com acabem de descriure. Formant part d'un treball de curs, vàrem confeccionar també unes altres coordenades de semblants característiques on utilitzàvem 8 colors diferents per a representar els diferents timbres vocàlics i la quantitat de zona pintada representava el nombre de respostes obtingudes. Aquest dos gràfics de coordenades descrits no són inclosos en el treball present per problemes obvis. En tot cas, el dipositari d'aquest gràfics és el Laboratori de Fonètica de la UB.

El fet que el subjecte qualifiqui un estímul amb *I* o amb *II* ha d'esser considerat com una valoració qualitativa. Per la qual cosa hem aplicat el següent procés de ponderació:

Les respostes *I* sumen 2 punts
Les respostes *II* sumen 1 punt.

A partir d'aquí representarem un gràfic en coordenades per a cada una de les 8 vocals del català central on, de tots els estímuls sintetitzats de l'enquesta, es representen els que han estat sentits com a determinada vocal, concretant, d'aquesta manera, el camp de dispersió auditiu de cada vocal.

Els valors dels símbols emprats es detallen en el quadre 1 (veure pàg. 135).

4.3. Problemes.

Volem fer observar tres problemes els quals, malgrat que en siguem conscients, no s'han pogut eludir.

1. El mètode aplicat, per la seva complexitat i precisió, permetria d'observar les petites diferències en els triangles vocàlics de cada dialecte, subdialecte, o parlar d'una localitat determinada. Però resulta summament difícil de trobar un nombre adequat d'informants d'una mateixa població i amb parlar homogeni. En el nostre cas tenim una barreja d'informants, tots pertanyents al català central (això sí) i amb predomini clar de Barcelona ciutat.

2. S'observà certa dificultat entre els enquestats a distingir les e i o obertes de les tancades. Si bé tots eren estudiants i, per tant, haurien de conèixer-ne la diferència, creiem que la dificultat ve principalment del fet que en la seva parla pròpia no es distingeixen prou bé (per influència del castellà).

3. Ja hem explicat que l'enquesta es va dividir en cinc grups d'informants. En cada grup podria haver-hi tendències diferents a l'hora d'interpretar un so, de manera que un mateix estímül hauria pogut rebre respostes diferents. D'altra banda, dins d'un mateix grup, un mateix so pot arribar a rebre respostes diferents depenent del context, perquè un so fa una impressió diferent segons els últims sons vocàlics que s'han sentit. Per exemple, si l'estímül va precedit d'una sèrie de sons molt foscos o dolents, pot fer una sensa-

ció millor que si va precedit d'una sèrie de sons clars i oberts.

5. CONCLUSIONS.

5.1. Observacions

En primer lloc observem que la [a] té una àrea de dispersió clara i nítida. Per contra, per al camp de dispersió de la [u], és gairebé impossible traçar uns límits, i a més no només s'hi observa inseguretats (respostes *II*), sinó que conté un alt grau de resposta *III*. Aquest fet es pot interpretar com a fruit de dues possibles causes:

1. El sintetitzador produeix sons més bons, clars o identificables a la zona de la [a] que a la zona de la [u], agreujat, això, pel fet que els amples de banda (BW) dels formants de la [u] sintetitzada haurien de ser, sembla, tots quatre iguals per a obtenir un bon so; nosaltres, en canvi, els vàrem mantenir constants per no caure en l'apriorisme de decidir arbitràriament a partir d'on els canviàvem.

2. Essent la [a] una vocal oberta i clara i la [u] una vocal tancada i fosca per excel·lència, ens va fer la impressió que els enquestats tendien a rebutjar més els estímuls tipus [u], qualificant-los amb *II* i *III*, (però no en nombre suficient per arribar a ser representants).

Hem observat també que, en sons que potser haurien de rebre una resposta *III* (és a dir, que són fora del camp de dispersió de qualsevol vocal catalana), s'hi troben respostes del tipus *u II*, --sovint en quantitat no representable--, cosa que confirmaria la identificació de la [u] amb qualsevol so fosc, poc clar, per allunyat que sigui del camp de dispersió de la [u].

A mida que F2 va pujant, es va acostant a F3, situat a 2842 Hz, a banda palatal. De manera que a l'abscissa 30 (2842 Hz) els dos formants coincidirien. Això ens va fer aturar l'exploració a 28 de F2, per mantenir una distància entre els for-

nants. Però els bons resultats de [i] amb F2 més alt ens fa pensar que si pujàvem F3 podríem fer pujar F2 i obtindriem igualment bons resultats de [i]. D'altra banda creiem que no passaria el mateix si continuàvem l'exploració elevant F1, ja que a la part baixa i esquerra del camp de dispersió de e oberta (sobretot amb F2 a 2400 Hz) hem trobat una resposta important de tipus e tancada, la qual cosa ens fa pensar que aquí tenim ja un so distorsionat i espuri per a la llengua catalana.

Pel que fa al camp de dispersió de la vocal neutra, observem que té els límits poc definits i que conté puntuacions baixes, que expressen inseguretats. Això pot tenir dues interpretacions:

1. El camp de dispersió de la vocal neutra es toca amb el de totes les vocals obertes i mitjanes, i les superposicions de camps de dispersió es tradueix en inseguretats en les respostes.

2. S'hi podria aplicar quelcom de semblant al que hem dit de la [u] la vocal neutra és una vocal no clara, "neutra", i a l'informant li costa més d'identificar-la o de decidir-se que davant d'una vocal clara i oberta com serien la a i la e i o obertes.

5.2. Comparació amb altres autors.

En el gràfic 9 exposem, el resultat de posar en comparació els nostres resultats amb el d'altres autors (sobreposant les seves propostes amb els nostres resultats).

Una observació a destacar: la [a] ens surt notablement més baixa. Tota la zona a partir de F1 a 28 (830 Hz) fins a 32 (1047 Hz) pot semblar exageradament alta per al primer formant, però les respostes són claríssimes. Com s'ha d'interpretar això?

1. Potser el so correspondria a una realització forçada, no natural dins la llengua, però el color de la vocal s'hi manté.

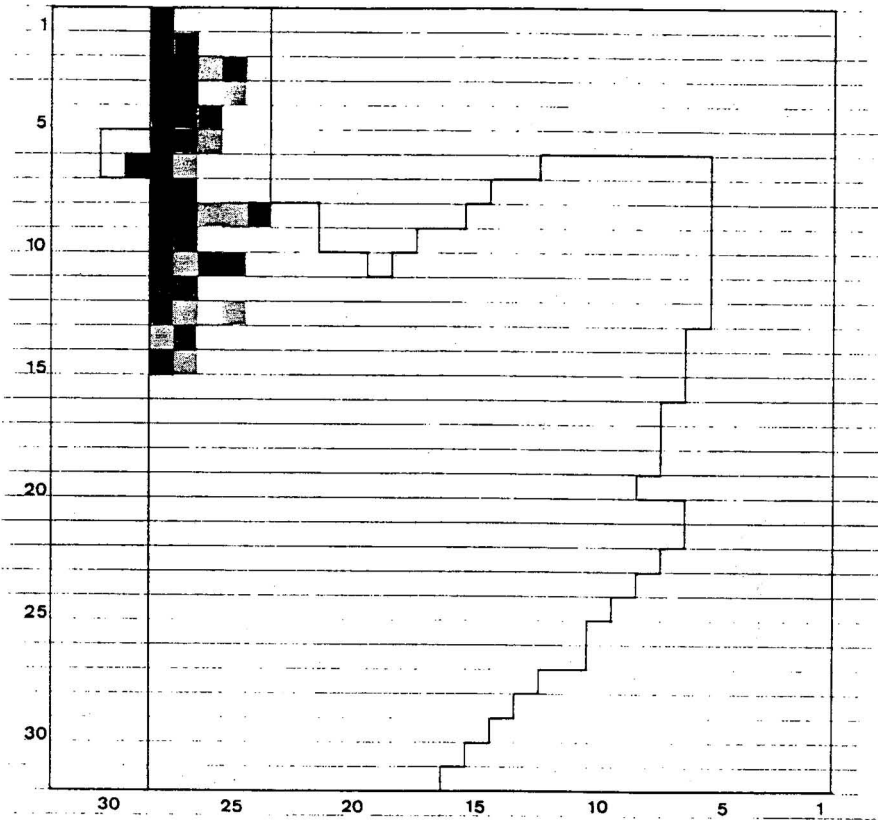
2. Podria tractar-se de sons que no es produeixen en la parla normal però sí quan s'emfatitza la pronunciació (per exemple, quan es vol dir el nom de la vocal). Aquesta segona interpretació ens sembla més convincent. L'exageració de les característiques de la vocal (per a la [a], un F1 alt) augmenta encara la seguretat de la identificació de la vocal, i això explicaria la quantitat de respostes I que hi trobem.

A part d'això, pensem que es podria donar el cas que un so sintetitzat contingui uns formants fisiològicament impossibles de produir en la parla humana i malgrat això l'oient els percebi com a sons de la llengua amb una nitidesa i qualitat acceptables.

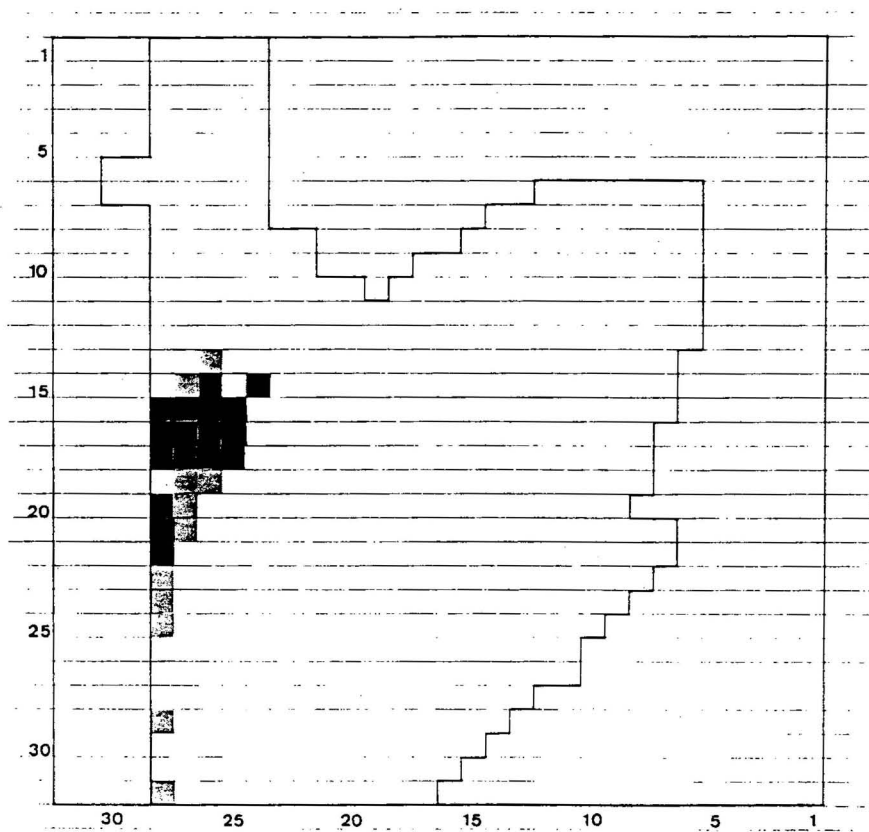
6. GRAFICS.

■	de 11 a 21 punts --	16.3 %	a	35 %
■	de 22 a 33 punts --	36.6 %	a	55 %
■	de 34 a 44 punts --	56.6 %	a	73.3 %
■	de 45 a 60 punts --	75 %	a	100 %

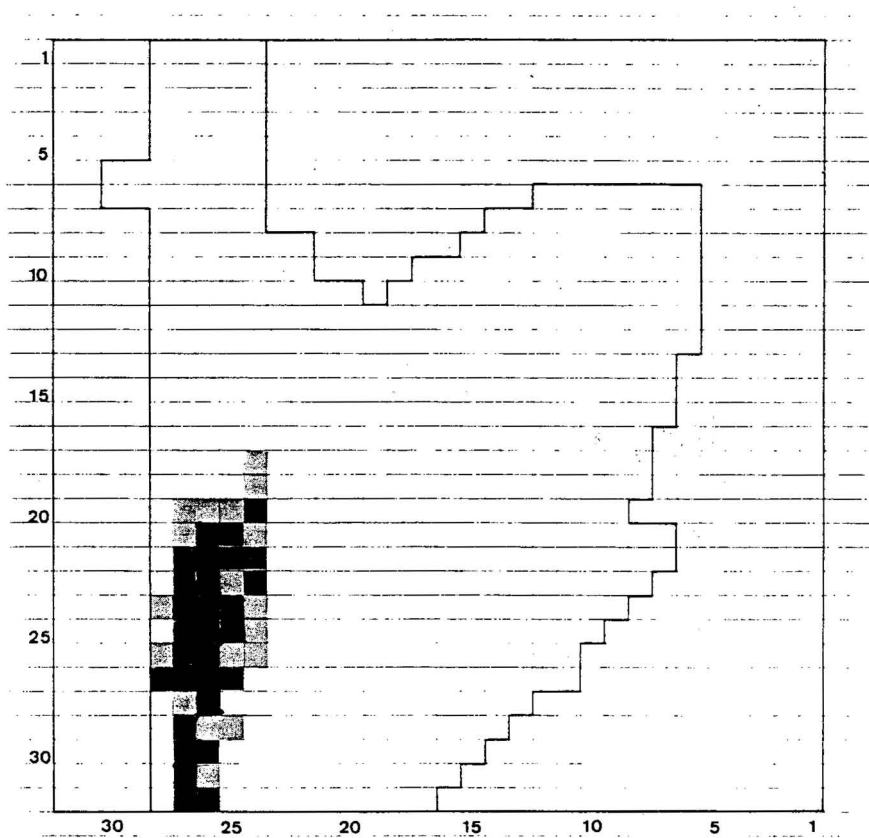
Quadre 1: valors del símbols emprats.



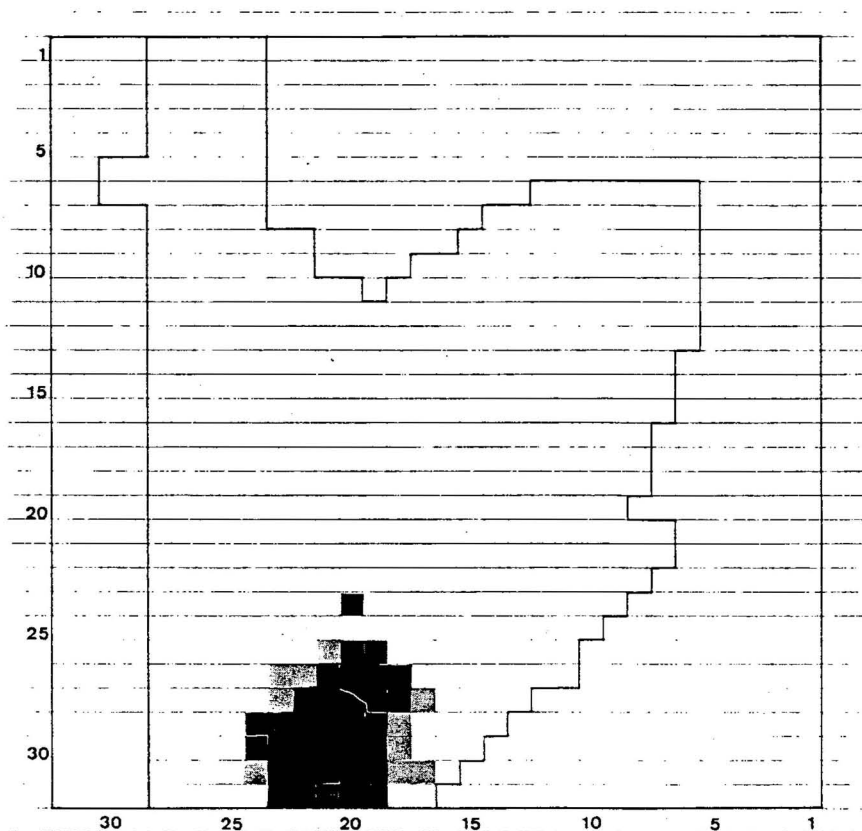
Gràfic 1. (i)



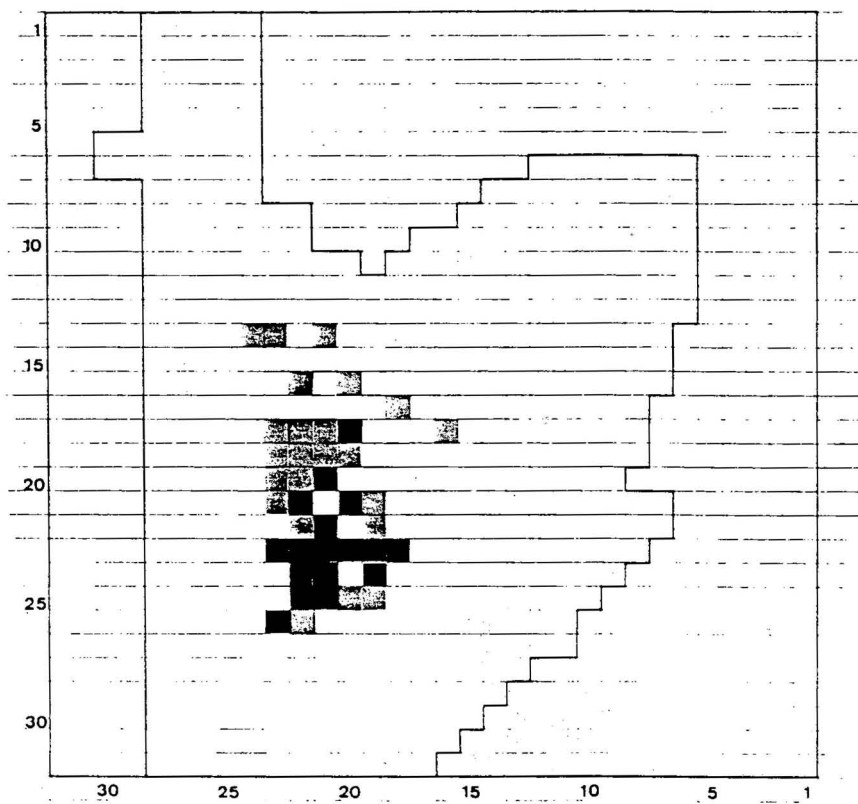
Gràfic 2. (e)



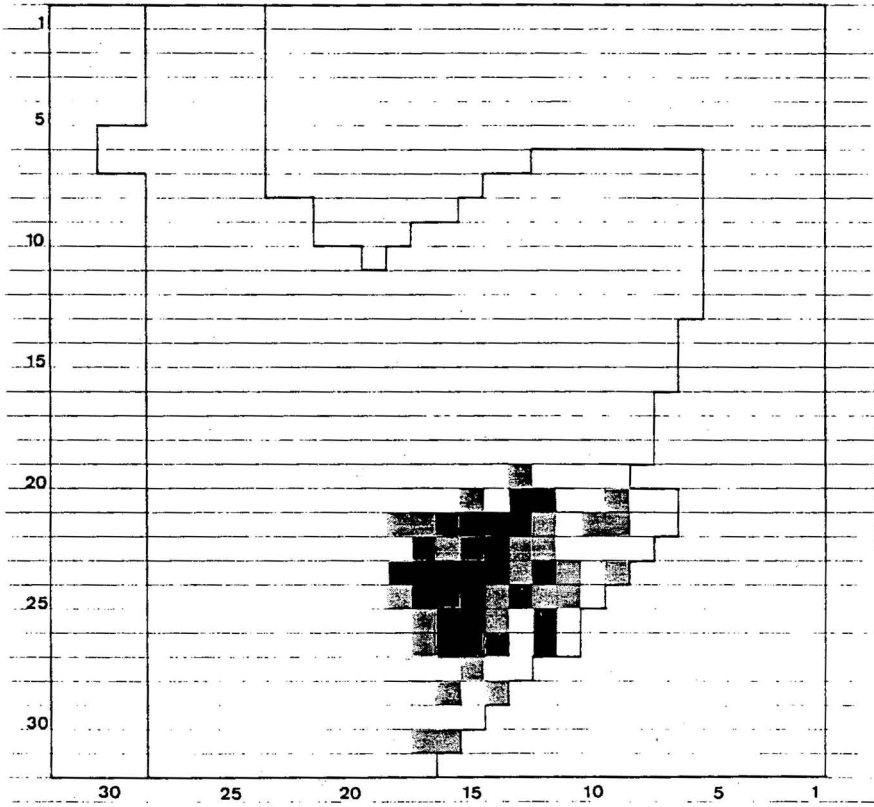
Gràfic 3. (e oberta)



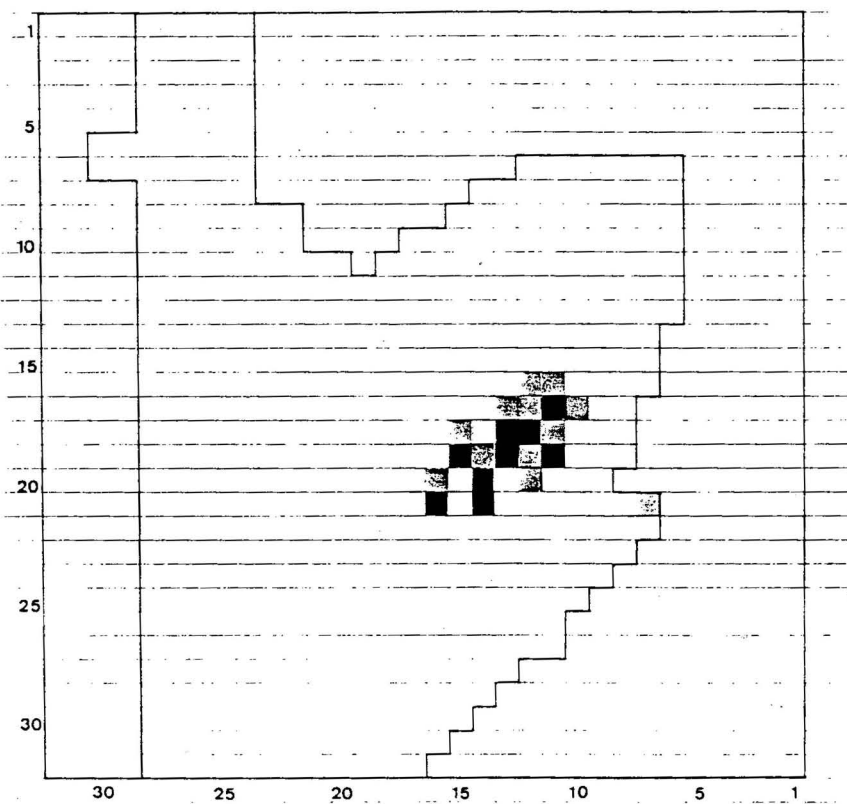
Gràfic 4. (a)



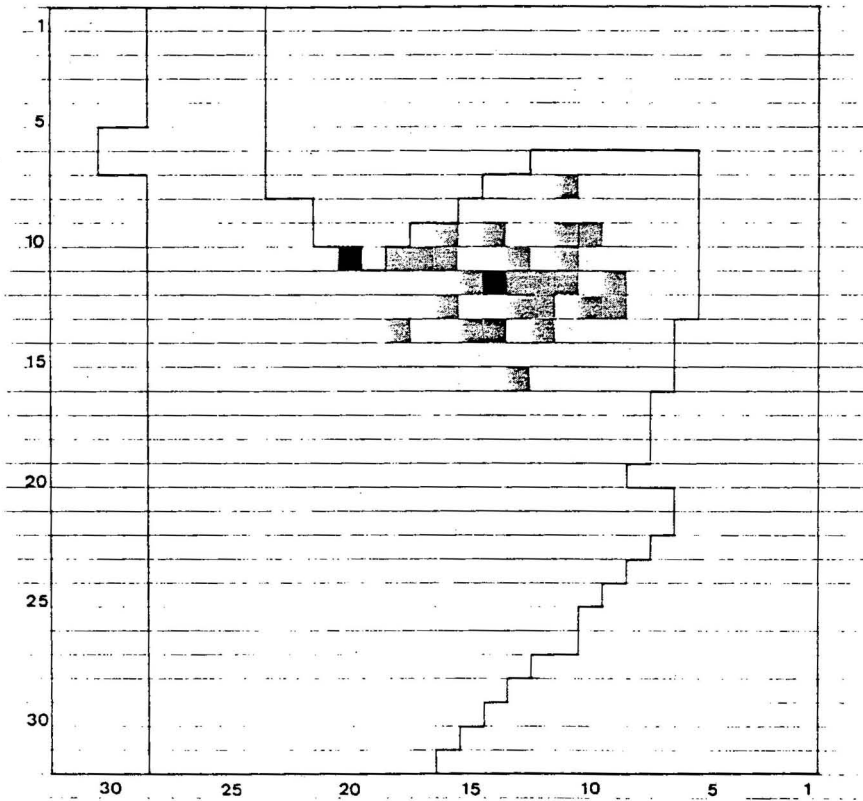
Gràfic 5. vocal neutra.



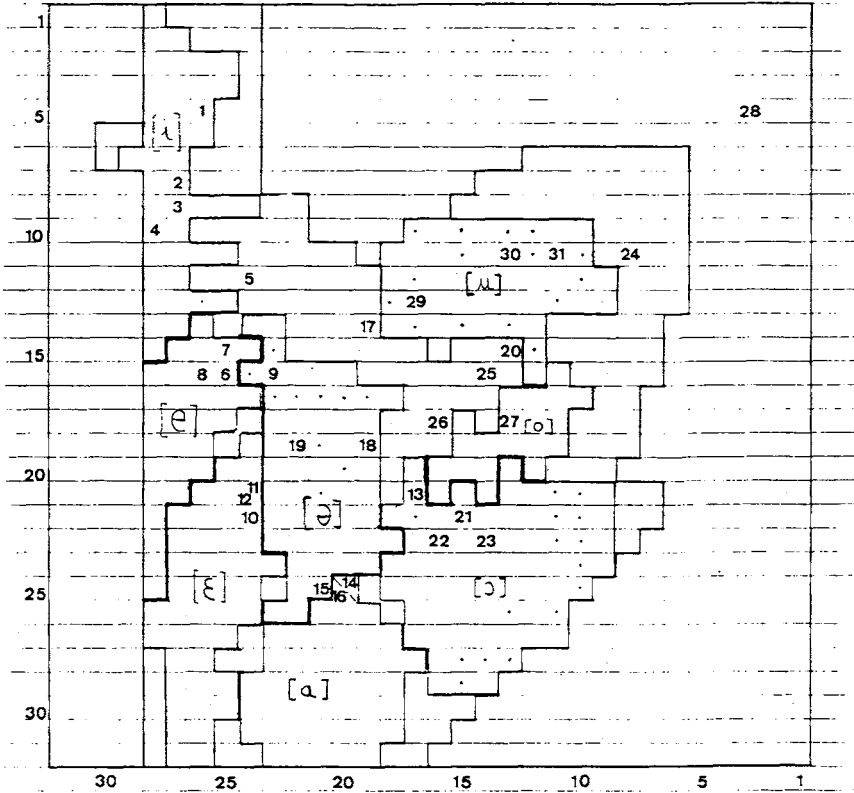
Gràfic 6. (o oberta)



Gràfic 7. (o)



Gràfic 8. (u)



Gràfic 9. Comparació amb altres autors.

En aquest gràfic hem delimitat els camps de dispersió de les 8 vocals del català central segons els resultats de la nostra enquesta. Hem encerclat amb una línia els sons que eren considerats com a mínim pel 50 % dels enquestats com a determinada vocal, és a dir, que la resposta de 15 o més persones era coincident. (En els camps que presenten buits o irregularitats hem assenyalat amb un punt aquells sons que no arriben a aquest 50 %, concretament la vocal neutra, la [u] i la o oberta). A continuació hi hem sobreposat, en els llocs que les correspondrien, les propostes de Cerdà (1972), Martí (1983), Recasens (1984) i Llisterrí (1984), representades cadascuna per una xifra.

6.1. Codis.

No.	Vocal	F1	F2	Autor
1	(i)	195 Hz	2050 Hz	C (Cerdà)
2		258.3	2179.8	M (Martí)
3		276	2156	R(Recasens)
4		291	2397	L(Llisterri)
5	(e)	330 Hz	1775 Hz	C
6		404.8	1929.8	M
7		397	1892	R
8		412	2075	L
9	(e oberta)	415 Hz	1625 Hz	C
10		581.0	1365.5	M
11		544	1811	R
12		550	1745	L
13		540 Hz	1125 Hz	C
14		694.5	1287.5	M
15		697	1420	R
16		692	1307	L
17	vocal neutre	370 Hz	1225 Hz	C
18		504.8	1287.5	M
19		490	1510	L
20	(o oberta)	400 Hz	890 Hz	C
21		586.9	1014.3	M
22		611	1074	R
23		617	924	L
24	(o)	300 Hz	670 Hz	C
25		417.4	947.6	M
26		462	1024	R
27		462	848	L
28	(u)	195Hz	480 Hz	C
29		329.8	1097.6	M
30		306	895	R
31		300	800	L

7. BIBLIOGRAFIA.

- Cerdà, R. 1972. *Parámetros vocálicos del catalán*. Madrid. C.S.I.C
- Llisterri, J. 1984. Aproximación a la síntesi de les vocals del català, *Folia Phonetica*. Estudi General de Lleida, UB. Lleida.
- Martí, J. 1983. *Estudi acústic de les vocals catalanes*. Memòria de Llicenciatura no publicada. Universitat de València. Facultat de Ciències Físiques.
- Martí, J. 1984. Paràmetres vocàlics del català, *Folia Phonetica*. Estudi General de Lleida, UB. LLeida.
- Recasens, D. 1981. Futur dels estudis de fonètica experimental del català, *Els Marges* 21.
- Recasens, D. 1984. *Producció i coarticulació de vocals i consonants en del decurs*. Tesi doctoral inèdita. UB.
- Soler i Sabater, M.J. 1984. La experimentación en fonética i fonología, *Estudios de Fonética Experimental* I. PPU. Barcelona.

AGRAIMENTS.

Expressem el nostre agraïment als alumnes de l'IB Sants-Les Corts que varen col·laborar en l'enquesta, i especialment al seu professor de Català, en Josep Maria, per la seva ajuda que ens ha resultat imprescindible. També volem donar les gràcies per l'ajuda i la bona disposició en tot moment als professors de fonètica de la Facultat de Filologia de la UB, i d'una manera especial a Eugenio Martínez Celdran.