

## CRITERIS D'ELABORACIÓ D'UN CORPUS ORAL PER A L'ESTUDI DE L'ENTONACIÓ

En aquest article exposem els criteris i el procés d'elaboració que hem seguit per crear un corpus oral de parla espontània ampli, significatiu, aleatori i compensat, amb la finalitat de fer una descripció de l'entonació de la llengua catalana.<sup>1</sup>

Entenem que el corpus ha de basar-se en parla espontània per tal d'obtenir models entonatius reals i genuïns, idonis per descriure la realitat lingüística; ha de comptar amb un nombre d'informants ampli i alhora significatiu, és a dir, que siguin suficients i que siguin variats pel que fa al sexe, l'edat, i la procedència socio-cultural i dialectal; ha de ser aleatori o, dit d'una altra manera, regit per les lleis de l'atzar; i, finalment, ha de ser compensat, és a dir, ha de contenir una representació equilibrada dels diversos perfils d'informants i dels tipus d'enunciats, a la qual es pot tendir neutralitzant les tendències més marcades que presenta el corpus: ignorant el que és més abundant i afavorint el menys representat.

Un corpus d'aquestes característiques ens permetrà, després d'una llarga i acurada selecció dels enunciats i de sotmetre'ls a una anàlisi formal, fer comparacions entre les diverses melodies dels enunciats, constatar-ne les coincidències i establir generalitzacions lingüístiques.

### EL CORPUS DE GRAVACIONS

#### *Model lingüístic*

Per a la creació del corpus, el nostre interès se centra a obtenir mostres de llengua oral espontània, informal. Entenem la llengua espontània com aquella expressió lingüística oral no planificada; és a dir, la llengua parlada i d'ús habitual que podem trobar en els diversos contextos no formals.

1. Els resultats obtinguts a partir d'aquest corpus es poden trobar a Font (2004 i 2005).

Vam optar per aquest model perquè la llengua, per naturalesa, és oral i espontània. Els altres registres orals, més formals, presenten una preparació prèvia que ens desviaria del nostre objectiu: descriure l'entonació del català. Un discurs oral formal ha estat prèviament escrit o parteix d'un guió molt detallat, la qual cosa implica que el locutor s'ha plantejat el que dirà i com ho dirà, usant totes les estratègies i habilitats lingüístiques al seu abast. L'orador pot decidir el lloc on farà èmfasis, la velocitat i el volum d'elocució, la llargària de les diverses pauses, el ritme i les característiques de l'entonació. És evident, doncs, que aquests models formals no poden ser la base per descriure l'entonació, pel fet que el model lingüístic que ens aporten pot haver sofert moltes modificacions, condicionades pels coneixements lingüístics del locutor, molt sovint adquirits en un context escolar on el model de llengua que s'aprèn és mediatitzat per la llengua escrita.

La llengua oral espontània es manifesta bàsicament en el diàleg, que té lloc entre dues o més persones en qualsevol moment de la vida quotidiana, ja sigui per mantenir una conversa, parlar per telèfon, comprar o demanar informació, etc. Per tant, doncs, per a la creació del corpus, calia buscar contextos en què es produïssin diàlegs.

### *Els informants*

Vam partir de la base que, si volíem aconseguir dades suficients i fiables per fer el nostre estudi, el nombre d'informants havia de ser ampli i que, en la mesura que fos possible, la mostra havia de presentar individus d'ambdós sexes, d'edat, de professió, i de procedència sociocultural i dialectal diversa, a diferència dels estudis sobre entonació elaborats a partir d'enregistraments fets en presència de l'autor (i gairebé sempre al laboratori) que tenen un nombre d'informants limitat i poc variat. El perfil d'informant que crèiem adequat per a aquest treball d'investigació és el que exposem a continuació.

Els informants havien d'utilitzar el català com a llengua familiar i d'ús social i presentar una pronúncia genuïna. També havien de vocalitzar bé, no presentar cap defecte de parla, ni una veu amb característiques peculiars del tipus nasalitzada o engolada, entre d'altres, que poguessin desviar els resultats de l'anàlisi de les dades.

També es demanava que el lloc de treball que ocupessin no impliqués tenir una bona preparació en llengua oral, com seria el cas de mestres, locutors, polítics, actors o advocats. Els discursos d'aquests professionals, encara que no tinguin el text escrit davant, presenten un grau menor d'espontaneïtat, a causa de les habilitats que han anat desenvolupant amb la pràctica. Ara bé, quan aquests informants es troben en determinats contextos, fora del seu àmbit professional, com seria el cas d'una

conversa amb amics o un debat, poden esdevenir informants del nostre corpus perquè fan ús d'un model de llengua col·loquial i espontani en un context real.

Finalment, els informants s'havien d'expressar amb naturalitat i sense nerviosisme; en aquest sentit, ens calia trobar un context en què no sabessin que se'ls estava escoltant i enregistrant per fer un estudi lingüístic.

En resum, els informants havien de ser catalans, anònims, tenir com a llengua materna i habitual el català, que, preferiblement, no es dediquessin a professions que impliquin tenir una bona preparació en llengua oral i que s'expressessin amb naturalitat en una conversa genuïna.

### *Obtenció de les dades*

Seguint Cantero (2002: 167), vam creure que la televisió era un mitjà idoni per obtenir les dades necessàries per a la creació del corpus. La televisió d'avui dia presenta tota una sèrie d'avantatges, com poden ser varietat de programes representatius de gèneres molt diversos, multiplicitat d'informants i bona qualitat acústica del senyal de veu, indispensable per poder ser analitzada, perquè s'emet des d'un estudi tancat i ben sonoritzat i els participants parlen amb micròfon.

Pel que fa a la varietat de programes, si tenim en consideració els canals de més audiència, TV3, TV2 i Canal 33, ens n'ofereixen una àmplia gamma per poder escollir. Pel tipus de model lingüístic i d'informants que necessitàvem, vam creure convenient de buscar programes que continguessin gèneres com el debat, l'entrevista, la tertúlia, el reportatge o el concurs. Aquest tipus de programes són molt útils per a la creació del nostre corpus perquè presenten models de llengua espontània usats en un context pragmàtic real i hi participen moltes persones, la majoria d'elles anònimes, el model de llengua de les quals ens permetrà elaborar una mostra rica, variada i representativa del català.

Hem cregut que servir-nos de les dades que ens ofereix la televisió per crear el corpus oral era més apropiat que basar-nos en les d'altres mitjans. Així, doncs, la ràdio, ens podria brindar un material lingüístic idoni, que podria ser enregistrat en cinta de casset, però que presenta l'inconvenient que la gran majoria de persones que hi participen, per no dir totes, tenen un cert grau de professionalització i, en conseqüència, el seu model no és espontani.

També podem obtenir les dades enregistrant l'informant en un espai tancat a partir de frases que l'indueixin a produir aquells enunciats que ens interessen per al nostre estudi o mantenir una conversa deseixida i informal. Aquest tipus de plantejament és adequat, però té la seva dificultat a l'hora de crear un clima relaxant a l'interlocutor per tal que no perdi naturalitat i és costós obtenir un gran nombre d'enunciats a partir de múltiples informants.

Menys indicat és obtenir dades a partir de converses telefòniques que es produeixen en contextos reals, les quals són una bona mostra de llengua espontània, perquè el que es diu a través del fil telefònic no sempre es pot analitzar acústicament. I pot resultar gairebé impossible analitzar material lingüístic obtingut en espais oberts, per l'encavallament dels sons dels informants amb els sorolls de l'entorn, com poden ser els provocats pels automòbils, per les botzines o per d'altres persones.

### *Material audiovisual*

Per elaborar el corpus de gravacions, ens vam basar en programes de televisió emesos en els tres canals públics de més audiència i que tenen programació diària en llengua catalana: TV3, Canal 33 i TV2. Mentre que TV3 i Canal 33 tenen tota la programació en llengua catalana, TV2 només en té en una petita franja del vespre, aproximadament de 19:30 a 21h. i algun dia a la nit.

Dividim el procés de recollida de les gravacions dels programes en tres etapes:

a) De setembre del 2000 fins a mitjan desembre del 2000. Vam seleccionar de cada canal els programes que més s'adeïen amb les nostres necessitats i vam enregistrar 24 emissions —es troben entre parèntesi al costat del nom del programa— de 13 programes diferents. Es tracta dels programes de TV3 *Bon dia, Catalunya* (1), *Connexions* (1), *a+a+* (3), *La columna* (3), *Jutjats* (2), *Entre línies* (3) i *Temps era temps* (1); dels programes de TV2 *Catalunya Avui* (5), *La vida és bella* (1), *Tic, tac, toe* (1) i *El rondo* (1); i dels programes del Canal 33 *Els llibres i la vida* (1) i *Gol a gol* (1). Les gravacions d'aquests programes sumen un total de 23h 35min.

b) De mitjan desembre del 2000 fins a mitjan gener del 2001. Pel fet que en aquell moment no hi havia cap programa de debat en la programació, vam obtenir cinc programes de debat de *L'ou o la gallina*, emesos els anys 1996-97, de la secció de documentació de TV2,<sup>2</sup> i també una edició del *Catalunya Avui* de l'agost del 2000 i una de *El Rondo* del setembre del 2000. Aquests programes sumen un total de 11h 20 min. d'enregistrament.

c) De mitjan gener del 2001 fins a principis de febrer del 2001. Vam obtenir una sèrie de gravacions de programes de TV3 i de Canal 33 dels mesos de maig, juny i setembre del 1999.<sup>3</sup> Es tracta dels programes de TV3 *a+a+* (1), *Punt de*

2. Agraïm als responsables de TV2 i, especialment, a A. Zamora, les facilitats que ens van donar per accedir a les fonts documentals del canal; també a José Luís López, de la secció de Documentació, per haver-nos ajudat a seleccionar el material i a gravar-lo.

3. Aquests programes van ser enregistrats per A. Canals, G. Ester, i A. Hermoso i ens van servir per fer un primer estudi exploratori sobre les interrogatives del català, presentat com a comunicació al II Congreso de Fonética Experimental que es va celebrar a Sevilla l'any 2001 (Font *et alii* 2002).

*trobada* (3), *Si l'encerto, l'endevino* (1), *Això no és cap anunci* (3), *La cosa nostra* (1), *Malalts de tele* (1); i del programa del Canal 33 *Paral·lel* (1). Aquests programes sumen un total de 12h de gravació.

### *Descripció del corpus*

El corpus de gravacions és format per 46h 55m de material audiovisual, procedent de 20 programes diferents obtinguts en els tres canals de televisió: 12 programes de TV3, 5 programes de TV2 i 3 del Canal 33. En total són 42 lliuraments que es van emetre en tres períodes diferents: de novembre de 1996 a febrer de 1997; maig, juny i setembre del 1999; i d'agost a desembre del 2000. Els tipus de gènere que representen aquests programes (reportatge, entrevista, enquestes al carrer, debat, concurs, tertúlia i docudrama) ens ofereixen models de llengua oral espontània i hi abunda el diàleg.

Davant la dificultat de treballar amb un corpus oral d'aquestes característiques, vam procedir a una primera selecció per tal de poder tenir un corpus de gravacions més reduït que ens permetés obtenir els enunciatos de forma més àgil. En aquesta selecció, en una primera etapa, vam excloure del corpus fragments en què els informants no complien els requisits mínims que havíem establert en els criteris de model lingüístic i d'informants, ja sigui perquè no parlaven en català, o perquè no eren parlants genuïns o perquè el registre que usaven no era del tot espontani, o perquè simplement llegien el guió. En una segona etapa, essent més estrictes, vam eliminar intervencions d'informants diversos, que presentaven melodies repetitives, i també vam escurçar i, a vegades, suprimir intervencions molt llargues i reiterades de determinats participants.

Com a conseqüència d'aquest procés, vam obtenir el corpus de gravacions definitiu, el qual compta amb 21h de material audiovisual i conté la mateixa riquesa i varietat de programes i d'informants que l'anterior.

### EL CORPUS D'ENUNCIATS

En aquest apartat, expliquem el procés que es va seguir per crear el corpus d'enunciats dividit en fases, els problemes que ens van sorgir i com els vam resoldre; presentem la fitxa model amb la informació bàsica que havia de contenir i els criteris que vam utilitzar per omplir-la; i descrivim el model lingüístic i les característiques dels informants, com també la procedència i la classificació dels enunciats.

*Procés d'elaboració*

El corpus de gravacions que hem elaborat ens garanteix l'acompliment dels criteris que ens havíem proposat: que el model de llengua fos de llengua oral i espontània, basat en el diàleg, que presentés multiplicitat i varietat d'informants de parla catalana i que els enregistraments tinguessin una bona qualitat acústica. Ara bé, el fet d'obtenir dades per mitjà de programes de televisió també planteja alguns problemes per a l'investigador.

Així, doncs, ens trobem que la majoria de persones que participen en aquest tipus de programes són anònimes. Això implica no disposar d'informació detallada i objectiva sobre on viuen, quina edat tenen, a què es dediquen, si són catalanoparlants o si usen el català com a llengua d'ús habitual. D'alguns d'aquests informants potencials, no en sabem res, només en tenim la imatge i els fragments en què parlen; d'altres, n'arribem a saber el nom de pila o, a vegades, fins i tot el cognom; i encara, d'uns altres, sabem a què es dediquen o el lloc on viuen. En conclusió, ens manquen dades objectives que ens els defineixin.

Davant d'aquesta constatació, vam haver d'observar i d'analitzar cada informant individualment a partir de tot el material audiovisual que teníem. Seguidament, decidir si la persona tenia el perfil adequat per formar part del nostre corpus; és a dir, si parlava un català genuí, vocalitzava bé i no tenia defectes de parla. En el cas que reunís aquestes característiques, aleshores, es va haver de recollir tota la informació que teníem sobre el nom, el sexe, l'edat, la professió, el lloc de residència i la procedència dialectal, la qual es pot obtenir en subtítols que s'impressionen en la pantalla, en els continguts que exposa la persona directament o el locutor quan el presenta i en informació implícita que es dona en les intervencions.

Com que en la majoria de vegades no disposàvem d'aquesta informació, en el cas de l'edat, en fèiem una estimació aproximada, i en el del lloc de residència i de la procedència dialectal, els classificàvem nosaltres, tot partint de la varietat que parlava. La resta d'informació no podíem presuposar-la i, com a conseqüència, no consta en la fitxa de l'informant que vam elaborar.

Tota aquesta problemàtica va provocar que només acceptéssim, del gran nombre d'informants que teníem i de les múltiples intervencions que podien formar part del corpus, aquells que estàvem segurs que complien els criteris que ens vam proposar. Si en l'observació i l'anàlisi de l'informant ens quedava algun dubte, el desestimàvem.

A part dels informants anònims, en les gravacions també hi havia altres persones conegudes, ja sigui per la seva professió o perquè són habituals en els mitjans de comunicació. Érem conscients que s'havia d'anar molt amb compte a seleccionar aquestes persones com a informants, perquè coneixien molt bé el medi i tenien

una gran capacitat de planificació del seu discurs oral. Ara bé, escoltant i analitzant detalladament les seves intervencions, vam adonar-nos que algunes eren, en major o menor grau, controlades, però, d'altres, en què la persona havia d'intervenir molt ràpidament, sense tenir ni un segon per pensar el que i el com ho volia dir, eren totalment espontànies. Són aquestes intervencions les que vam considerar adequades per formar part de la nostra investigació. Així, per exemple, vam seleccionar enunciats de Mercè Comes, actriu, quan s'indignà davant unes afirmacions d'un participant en un debat; de Ramon Miravittles, presentador i moderador d'un programa de debat, quan en un moment va intervenir amb caràcter d'urgència perquè el públic de les graderies estava molt esverat; o del publicista i empresari Lluís Basat, quan va haver d'improvisar enmig d'una entrevista.

Vam procurar no escollir la primera intervenció dels informants, pel fet que podia haver estat preparada prèviament. A més a més, el nerviosisme que presenta el participant en aquest primer moment es va dissipant a mesura que va parlant i fent successives intervencions. L'informant, doncs, cada cop se sent més còmode amb el mitjà i s'expressa amb més naturalitat. Per seleccionar, cal saber esperar el moment en què per a l'informant és més important el tema que tracta que la resta de condicionants del mitjà.

Vam establir el procés de selecció dels enunciats en les següents fases:

a) Observar i analitzar cada informant en totes les seves intervencions. Assegurar-nos que parla un català genuí, que vocalitza bé i que no té cap defecte de parla.

b) Elaborar la llista d'informants.

c) Escollir els fragments en què cada informant triat s'expressa amb més naturalitat, sense nervis ni tensió i, alhora, parla deseixit, espontani.

d) Seleccionar enunciats variats: els d'ús més habitual i els més destacables per la seva singularitat. Els classificarem segons aquests criteris:  $\pm$ interrogatiu;  $\pm$ emfàtic i  $\pm$ suspès.

e) Comprovar que els enunciats triats no presenten solapaments de dos o més parlants, ni músiques de fons, ni sorolls ambientals.

f) Emplenar una fitxa per a cada enunciat amb un codi i una localització en la cinta.

### *La fitxa del corpus*

Vam haver d'elaborar una fitxa que ens servís de base per consignar tots els enunciats. Aquesta fitxa havia de contenir el codi de l'enunciat, l'enunciat, la localització en la cinta, el programa, el canal de televisió, el dia i l'hora d'emissió, la descripció dels informants i la descripció del context.

El codi de l'enunciat de la fitxa és format pel codi de la cinta (de D1 a D6), la posició que ocupa el programa dins la cinta, separat per un guió (de l'1 en endavant), i, finalment, el número d'ordre en què apareix l'enunciat dins el programa (de l'1 en endavant), separat per un guió. A vegades, seleccionàvem fragments que estaven formats per més d'un enunciat. Per distingir-los, vam decidir afegir una lletra al nombre de l'enunciat (de la *a* en endavant).

Era molt important que anotéssim la localització de l'enunciat en la cinta per tal que fos fàcil trobar-lo. Aquesta informació ens la proporcionava l'aparell de vídeo ja que ens facilitava el lloc on es trobava l'enunciat per mitjà del temps recorregut des del principi de la cinta; és a dir, l'hora, els minuts i els segons.

La resta d'apartats s'havien d'anar omplint amb la informació al màxim de fidedigna i objectiva. Un apartat on calia posar especial atenció era el de la descripció del context. En aquest apartat, s'havia de procurar explicar en poques paraules en quina situació discursiva i pragmàtica es donava l'enunciat: si era una pregunta, una resposta, una afirmació rotunda i categòrica, si hi havia ironia, agressivitat o simpatia, per dir algunes de les múltiples possibilitats que es poden donar. Pel que fa a la descripció dels informants, consignàvem tota la informació que teníem, sobre l'edat, el sexe, la professió i la procedència. Finalment, després de tenir tots els elements bàsics, ens calia classificar l'enunciat, segons si en el context era Interrogatiu o no, si contenia un Èmfasi o no i si era Suspès o no.

Seguidament, presentem un exemple de fitxa plena. En aquesta fitxa del corpus trobem, en primer lloc, el codi D4-1-25b, el qual significa que aquest enunciat es troba a la cinta D4, al primer programa de la cinta -1-, i l'enunciat és el número 25. Hi hem afegit la lletra *b* perquè en el fragment que hem triat hi ha dos enunciat. En aquests casos, s'escriuen els grups fònics seleccionats i es marca l'enunciat que analitzarem entre cometes; la resta d'enunciat d'un mateix fragment es troben en fitxes successives, també entre cometes i codificats amb les lletres successives de l'abecedari (*a*, *b*, *c*, *d*, etc.). Pel que fa a la localització dins la cinta, sabem que l'enunciat és a 20 minuts 21 segons del començament de la cinta.

#### FITXA DEL CORPUS DE PARLA ESPONTÀNIA

*Criteris formals de classificació:* -Interrogatiu +Emfàtic -Suspès

*Codi:* D4-1-25b *Localització:* 00:20:21

*Programa:* L'ou o la gallina *Gènere:* debat

*Canal:* TV2 *Dia d'emissió:* 18/01/97 *Hora:* 22:15-00:15

*Enunciat:* «Sisplau, tinguem una obertura més de mira!»

*Descripció del context:* es tracta d'un participant del públic en un debat que fa una intervenció llarga. Al final emet aquest prec adreçant-se a tota la graderia de davant, que té una opinió contraposada a la que ell defensa.

*Descripció de l'informant:* Home d'edat compresa entre els 51 i els 60 anys. Parla català central, però desconeixem on viu i de què treballa.



Hi ha també el títol del programa, el gènere, el canal i l'hora d'emissió. A continuació, trobem la descripció del context, que cal que sigui molt explícita; en aquest cas, es tracta d'un prec a altres interlocutors per tal que reconsiderin la seva posició contrària respecte al tema de debat: *Els homosexuals poden formar família?* Al darrer apartat, hi ha totes les característiques que tenim d'aquest informant: es tracta d'un home d'entre 51 i 60 anys que parla català central.

Finalment, veiem la classificació de l'enunciat en: -Interrogatiu, perquè no formula cap pregunta; +Emfàtic perquè apuja el volum en emetre l'enunciat; i -Suspès perquè és acabat.

Per reunir tots els enunciats que hem obtingut, vam fer tres visualitzacions completes de les cintes. En la primera visualització, vam crear el cos del corpus d'enunciats bàsic, i en la segona i la tercera, vam completar-lo, afegint-hi més enunciats als quals no havíem parat atenció anteriorment. Com a resultat de tot aquest procés de codificació i consignació dels enunciats, vam obtenir 586 fitxes, que es corresponen a 586 enunciats diferents, els quals ja estaven a punt per ser digitalitzats.

Després de tot el procés de digitalització i anàlisi acústica, el corpus va sofrir unes modificacions (com veurem a l'apartat següent) i el nombre d'enunciats es va reduir a 580. Aquests enunciats constitueixen el corpus definitiu.

### *Model lingüístic i informants dels enunciats*

Els 580 enunciats que hem obtingut són mostres de català oral espontani, recollits en un context de diàleg. Aquests enunciats provenen de 160 informants, entre homes i dones i en edats compreses entre 18 i 70 anys. Hem classificat els informants en 5 grups d'edat, cadascun dels quals comprèn un tram de 10 anys, excepte el primer grup que n'abasta 13.

QUADRE 1 · ELS INFORMANTS: SEXE I EDAT

Franges d'edat	Home		Dona		Nombre d'informants	
	n.	%	n.	%	n.	%
18-30	17	46%	20	54%	37	23%
31-40	12	40%	18	60%	30	19%
41-50	22	67%	11	33%	33	20,5%
51-60	26	79%	7	21%	33	20,5%
61-70	16	59%	11	41%	27	17%
Total	93	58%	67	42%	160	100%

Com podem veure al quadre 1, tenim un nombre important d'informants, 160, que reuneixen les característiques bàsiques que es demanaven per crear el corpus.

Podem veure que hi ha un 58% d'homes i un 42% de dones, que estan repartits en els diversos trams d'edat: el que té més informants és el tram de 18 a 30 anys amb un 23%; fixem-nos que en aquest grup els informants tenen un diferència de 13 anys, mentre que a la resta, la diferència és de 10 anys; la resta de trams se situen entre el 20,5% i el 17%. Amb tot, aquestes dades podrien sofrir lleugeres modificacions perquè ja vam comentar que en la majoria de casos s'ha establert l'edat dels informants a partir d'una estimació subjectiva de l'investigador.

Un fenomen semblant s'ha produït a l'hora de conèixer la professió dels informants. Sabem a què es dediquen un 30% dels informants. De la resta, un 70%, no tenim elements que ens ajudin a saber-ho. El tipus de professions que apareixen són molt variades i de diverses categories: metge, guàrdia urbà, administrativa, gimnasta, escriptor, fiscal, estudiant, jutge, carnisser, monitora de cavalls, polític, psiquiatre, actor/actriu, porter, àrbitre, jubilat, mestressa de casa, ballarina, professor, ginecòloga, agent immobiliari, empresari, comentarista esportiu, cuiner, pagès, cantant, obrer, periodista, advocat, pastor.

Els informants de la mostra són bàsicament de Catalunya i procedeixen de diverses contrades. En alguns casos, sabem exactament el poble o ciutat on s'estan, en d'altres només la zona. Tenim informants de parla oriental de Barcelona, el Masnou, Sabadell, Reus, l'Escala, Tossa, Vilafranca del Penedès, Valls; i també de parla occidental de llocs com Lleida, Fraga, Alta Ribagorça i Delta de l'Ebre; finalment, tenim alguns informants procedents de València i de Palma.

L'opció que la majoria dels informants fossin anònims, ja sabíem que comportaria la sensació de no tenir informació detallada sobre la seva edat, procedència i professió. Això, però, des del nostre punt de vista no és un problema, perquè entenem que el sexe, l'edat o la professió són factors que poden fer variar el model de llengua que parlen els informants des d'un punt de vista de la pronúncia, de la morfologia, de la sintaxi o del lèxic, però en cap cas es modifica la melodia dels enunciat. La melodia és clau per interpretar els enunciat en un context sociopragmàtic: hem de poder percebre si ens estan fent una pregunta o un suggeriment, explicant o informant i saber actuar adequadament.

### *Classificació dels enunciat*

Per fer una primera classificació dels enunciat vam partir d'uns criteris formals del tipus: ±Interrogatiu, ±Emfàtic i ±Suspès, terminologia que hem extret de Cantero (2002: 137-143). Entenem, doncs, que un enunciat és +Interrogatiu quan en el context equival a una pregunta, ja sigui retòrica, ja sigui directa. La resta d'enunciat seran -Interrogatiu. Considerem que un enunciat és +Emfàtic quan conté alguna marca d'emotivitat, sense tenir en compte el significat pragmàtic. La resta

d'enunciats seran -Emfàtic. Finalment, classifiquem un enunciat com a +Suspès quan no és acabat i -Suspès quan és acabat.

Si fem servir aquests tres criteris per definir un enunciat i els combinem, obtindrem 8 grups amb característiques formals diferents, els quals ens permetran organitzar els enunciats. Aquesta classificació aproximativa és simplement una manera d'organitzar els enunciats i fer-los més manejables a l'hora d'estudiar-los (Quadre 2). Es tracta d'una primera classificació només a partir de les dades lingüístiques (gramaticals i semàntiques) i de les contextuals que tenim.

QUADRE 2. CLASSIFICACIÓ FORMAL DELS ENUNCIATS

Grups formals	N. d'enunciats
+Interrogatiu, +Emfàtic, +Suspès	44
+Interrogatiu, +Emfàtic, -Suspès	29
+Interrogatiu, -Emfàtic, +Suspès	35
+Interrogatiu, -Emfàtic, -Suspès	66
-Interrogatiu, +Emfàtic, +Suspès	135
-Interrogatiu, +Emfàtic, -Suspès	117
-Interrogatiu, -Emfàtic, +Suspès	94
-Interrogatiu, -Emfàtic, -Suspès	60

Aquesta classificació ens servirà per organitzar a partir d'ara els enunciats. Hem de dir, però, que només tenim 172 enunciats +Interrogatiu enfront de 408 -Interrogatiu. Aquesta descompensació s'explica per la dificultat de trobar enunciats +Interrogatiu, ja que aquests no són pas els més abundants en la llengua.

## RECAPITULACIÓ

En crear un corpus per fer un estudi de l'entonació, les dades han de procedir de contextos reals on les persones s'expressin de forma oral i espontània amb la finalitat d'obtenir models lingüístics genuïns que no hagin estat mediatitzats per la llengua escrita.

Un d'aquests contextos reals on es poden trobar múltiples informants és el mitjà televisiu, del qual hem obtingut 46h 55m de material audiovisual, procedent de 20 programes diferents, amb una bona qualitat acústica i amb informació visual sobre qui parla i sobre la seva expressió facial i els gestos. D'aquest material audiovisual, hem seleccionat 580 enunciats que formen el corpus. Són mostres de parla espontània, obtingudes en un context de diàleg i emeses per 160 informants catalanoparlants, de Catalunya, Palma i València, d'ambdós sexes i d'edat compresa entre els 18 i 70 anys, aproximadament.

Els enunciat del corpus, un cop digitalitzats i analitzats amb mitjans instrumentals, ens permetran establir comparacions entre les diverses melodies resultants, constatar-ne les coincidències i establir generalitzacions lingüístiques, amb la finalitat de fer una descripció de l'entonació de la llengua catalana.

DOLORS FONT I ROTCHÉS  
Universitat de Barcelona

#### REFERÈNCIES BIBLIOGRÀFIQUES

- CANTERO 2002: F. J. Cantero Serena, *Teoría y análisis de la entonación*, Barcelona, Edicions de la Universitat de Barcelona.
- FONT 2004: Dolors Font Rotchés, «Prosòdia», a *Veu i locució. Articles de dicàctica de la llengua i la literatura*, 32, 25-38.
- FONT 2005: Dolors Font Rotchés, *L'entonació del català. Patrons melòdics, tonemes i marges de dispersió*, tesi doctoral inèdita, Universitat de Barcelona, Departament de Didàctica de la Llengua i la Literatura.
- FONT *et alii* 2002: Dolors Font *et alii*, «Patrones melódicos de la entonación interrogativa del catalán en habla espontánea», a *Actas del II Congreso de Fonética Experimental*, Sevilla, Universidad de Sevilla, 192-197.