

LES VARIETATS SOCIOCULTURALS DE LA REGIÓ METROPOLITANA DE BARCELONA: EL COS (CORPUS DE VARIETATS SOCIALS) DE LA UNIVERSITAT DE BARCELONA

Els barcelonins viuen en una gran ciutat on hi ha tants fenòmens de variació interna, que ningú pot parlar en nom de tota la conurbació. Els canvis demogràfics, socials i econòmics, polítics i comunicatius han estat tan descomunals a Catalunya i a Barcelona, que han provocat canvis sociolingüístics també radicals.

Aquests canvis fan necessaris corpus com el que descrivim, si el que es vol de debò és conèixer l'ús majoritari dels parlants de la llengua, que són els que al capdavall «manen» (vegeu Solà 2003, que relativitza aquest comandament en llengües poc segures i poc autònomes com la catalana). A partir de corpus com aquest es poden establir comparacions amb altres varietats lingüístiques a la mateixa zona i a altres zones

I. OBJECTIUS DEL COS

El 1991, el departament de Filologia Catalana de la Universitat de Barcelona inicià, basant-se en les recerques anteriors, el Corpus de la Universitat de Barcelona (CUB),¹ del qual formen part el COS, que aquí presentem, el corpus oral de català col·loquial (COC) i el corpus de registres (COR). El corpus de varietats socioculturals (COS) del CUB, tingué com a objectiu constituir un corpus de català oral de la regió metropolitana de Barcelona, estratificat a partir de criteris socioculturals. El projecte va preveure recollir i transcriure mostres de parla comparables de parlants de primera llengua catalana, socialitzats i residents a Barcelona i la seva zona, per tal de poder-ne estudiar l'estructura, els usos, i les ideologies lingüístics en la dècada dels anys noranta. Els procediments de transcripció i digitalització han de permetre

1. El 1991 Emili Boix-Fuster presenta la primera proposta del CUB, inicialment centrada en els tres milions i mig d'habitants de la regió metropolitana de Barcelona, que s'inspirava en les recerques del British National Corpus i de Thibault & Vincent (1990). Dues visites d'estudi a Gran Bretanya i al Quebec, el 1992, van permetre conèixer de primera mà aquests projectes. Per a més detalls del conjunt del CUB remetem els lectors a Boix (1996) i a Oller *et alii* (2000).

la comparació d'aquestes mostres de parla amb les del corpus de varietats funcionals i de conversa de la Universitat de Barcelona (COR i COC respectivament).

El COS s'inicia esperonat per la publicació de les primeres enquestes metropolitanas de la regió metropolitana de Barcelona, que inclouen els usos lingüístics com un factor més de l'estructura social (Subirats 1992). Es trencava així una mica la impermeabilitat dels grans *surveys* a Catalunya a considerar rellevants els usos i coneixements lingüístics.² Seguint una tradició ja clàssica sociolingüística (Labov 1966, en el seu estudi pioner sobre l'East Side de Nova York, es basa en un estudi sociològic previ), també a Barcelona, l'aparició de l'Enquesta Metropolitana, va estimular que el 1991 comencés el projecte d'un corpus de varietats socioculturals i funcionals de la zona central del domini lingüístic català (Boix 1993, 1996, Sánchez 1996), del qual el COS és una part.³

Els punts de partida teòrics del COS són els característics de bona part de la sociolingüística. Les varietats *vernaculars*, tal com assenyala la sociolingüística variacionista, fonamentalment laboviana (Labov 1994; vegeu també Turell 1995, i Pla 1995) són les més apropiades per estudiar l'heterogeneïtat estructurada, que és condició necessària perquè hi hagi canvi lingüístic, un canvi sempre relacionat amb aspectes socials. Les altres varietats no apreses en la conversa familiar, són molt menys habituals, i apreses posteriorment i de manera desigual pels membres d'un grup lingüístic.⁴

Els trets i posicions socioculturals dels parlants, el que anomenem capitals econòmics, socials, culturals i simbòlics seguint la terminologia de Pierre Bourdieu, estan interrelacionats amb bona part de les característiques de la seva parla, i amb les seves ideologies envers aquesta parla. La parla comunica, però també indicia algun aspecte de les identitats socials dels parlants. Així mateix, la parla no només

2. La recerca majoritària en lingüística no mostra gaire interès per la variació sociocultural, en la mesura que deu ser considerada perifèrica per al coneixement del sistema lingüístic (Boix 1998 i 2001). De manera complementària, la majoria de la recerca sociològica, antropològica i de geografia urbana sol percebre la diversitat lingüística com un afer massa secundari, per raons que no correspon de tractar aquí. Des de les ciències del llenguatge als països de llengua catalana, falta encara un treball interdisciplinari, teòric, no sols instrumental, amb les ciències socials. Així, per posar dos exemples, la dialectologia ha sabut començar a incorporar elements de la dialectologia urbana i de les seves metodologies sociolingüístiques (Viaplana 1997), i l'estudi de la normativa i la prescripció se serveix dels conceptes de norma social (Vilà & Fargas 1999, Solà 1999), sense que es pui de debò en els conceptes sociològics de base. Des de les ciències socials, passa el mateix en sentit invers: les diferències sociolingüístiques de la població barcelonina són absents de molts treballs sociològics.

3. Al mateix temps, esperonat per la mateixa recerca s'inicia el projecte d'un corpus de les varietats del castellà també de l'àrea metropolitana de Barcelona, també des de la Universitat de Barcelona, i amb una base sociològica molt similar a la del COS (Hernández 1998).

4. Badia (1995: 37) també coincideix amb aquests principis. La llengua corrent o habitual, en la seva terminologia, és la que li mereix prioritat: «el patró de la llengua han d'ésser les modalitats més emprades d'un cap de dia a l'altre».

designa, sinó que assigna, crea ella mateixa identitats. Les varietats socioculturals reflecteixen i creen, doncs, identitats i desigualtats socials.

Ara bé: les tradicions variacionistes i de lingüística de corpus han treballat sobretot en poblacions força homogènies lingüísticament, ben diferents de Barcelona i la seva àrea d'influència, una conurbació urbana d'aiguabarreig de varietats del català, i alhora lloc de confluència i interferència mútua amb la llengua castellana. D'una banda, bona part del *català que ara es parla* és un català de persones que no l'han après a casa (i a la inversa per al castellà que s'hi empra). De l'altra, hi ha un nombre creixent de situacions socials d'intersecció entre els dos principals grups lingüístics, començant pels entorns escolars, les associacions i grups polítics, els llocs de treball de cara al públic, les places i els mercats, els mitjans de comunicació locals... fins arribar al clos familiar, on creix el nombre de parelles d'origen lingüístic mixt. Aquesta considerable fluïdesa del repertori lingüístic a la conurbació barcelonina (facilitada per la política d'esmoreïment, en situacions públiques, de les fronteres etnolingüístiques internes tradicionals entre *castellans* i *catalans*), en especial entre els grups joves, planteja dubtes seriosos sobre quin tipus de comunitat i comunitats lingüístiques hi podem trobar (Boix 2002). Entre alguns dels joves, que alternen amb gran facilitat una i altra llengua, potser més que dels conceptes de comunitat, caldrà servir-se del d'actes d'identitat, amb què Le Page i Tabouret-Keller (1985) han descrit el plurilingüisme de societats postcoloniales i mestisses.

Com indica Badia (1995: 73) en la seva darrera gramàtica, «distingir matisos diastràtics en l'ús de la llengua és una comesa altament delicada, [...] un camp per artigar», i difícil per a) la subjectivitat de la tria, b) la gran diferència entre llengua escrita i parlada, i c) la diferenciació regional en la consideració social de la gramàtica de la llengua. En el corpus que presentem, no són rellevants les dificultats b) i c) ja que ens basem en la llengua oral i sols en una varietat regional determinada, la barcelonina. La dificultat a) vol ser en part superada tot recorrent a la informació intersubjectiva de mostres de parla amb una certa representativitat (vegeu l'apartat sobre el disseny del COS). A l'hora de voler construir el corpus de varietats socioculturals barcelonines com aquest, la dificultat principal prové precisament de l'excessiu i exclusiu caràcter mesocràtic de les descripcions del barceloní, que no tenen en consideració el català popular, com succeeix en la coneguda gramàtica catalana de Badia (1962).⁵ Aquesta intensitat de canvis socials i lingüístics no hauria

5. En aquesta *Gramàtica* (1962) no són clars els límits entre el català normatiu, comú i el del català barceloní. Com demanava el filòleg de Castelló de la Plana, Germà Colon, (1973: 236) a Badia i Margarit, arran d'una descripció de la morfosintaxi catalana, «est-ce donc le barcelonais ou la langue littéraire qu'il a voulu nous présenter?» No debades, Badia (1962) considera els dos termes (*barceloní* i *llengua literària*) gairebé equivalents: «La presente Gramàtica catalana aspira a ser expresión de la modalidad barcelonesa en el conjunto del dominio catalán, modalidad que el uso y la tradición identifican con la lengua común y literaria» (7);

de ser un obstacle sinó un estímul per estudiar-lo en el conjunt del seu repertori, tant català com castellà.

2. DESCRIPCIÓ DEL COS

2.1 *Disseny i obtenció de les dades del COS*

2.1.1 Delimitació de la zona i selecció dels informants

L'univers dels informants se circumscriu als parlants que tenen el català com a llengua primera, provinguin tant de llars bilingües homogènies lingüísticament com de bilingües familiars, socialitzats a les comarques del Barcelonès, Baix Llobregat, Maresme i Vallès Occidental i Oriental.⁶ La tipologia social dels parlants entrevistats es basa en la mostra i les classificacions de l'enquesta metropolitana de l'Institut d'Estudis Metropolitans de la Universitat Autònoma de Barcelona (Subirats 1992; Sánchez 1996). Segons l'enquesta metropolitana del 1991, el percentatge de població de primera llengua catalana de la regió de Barcelona se situa entorn del 35%, és a dir més d'un milió d'habitants, mentre que els bilingües familiars, se situaven

«Consecuencia obligada de la erección de la modalidad barcelonesa como lengua común ha sido decidir en favor del uso de la capital soluciones lingüísticas oscilantes» (16). Badia assenyala que el barceloní, més enllà de la norma explícita acadèmica (de poca o nul·la influència social l'any 1962, per raons òbvies) és la varietat de referència que guia les normes d'ús dels parlants, atès «el hecho evidente de la capitalidad efectiva que en todos los órdenes ejerce Barcelona» (16). I conclou: «todo ello ha ido formando una *conciencia de superioridad que sienten los catalanes del dialecto central y más concretamente de la modalidad barcelonesa*, y desde ella se observa que los hablantes de otros dialectos tienden a acercarse a las maneras de la capital o de una u otra forma se sienten cohibidos por ellas» (16, la cursiva és nostra). Recordem, per exemple, la difusió, suposadament des de Barcelona, de l'antiieisme i de l'argot d'origen caló. Ni que sigui oposant-s'hi o rebutjant-los, els parlants d'altres varietats fora de Barcelona, sobretot del Principat, s'hi han vist influïts. I en quins sectors socials barcelonins es basava la descripció de Badia (1962)? Com afirma ell mateix anys després (1995: 12), «l'atenció hi era concentrada en la modalitat barcelonina de la llengua en medis més aviat cultes», és a dir, en la llengua de les classes mitjanes barcelonines que havien tingut contacte amb els models normatius fabrians, o, el que és el mateix: la varietat social del mateix lingüista. Aquesta selecció social dels informants és la mateixa que la recerca de la norma lingüística culta en castellà (Alvar 1972). En la gramàtica de Badia (1995), en canvi, que duu el subtítol de diatràtica, es fa referència constant a tres nivells de la llengua: elevat, corrent i col·loquial.

6. L'expansió de la població barcelonina al llarg dels noranta, resseguint els eixos de comunicació per carretera i ferroviaris desborda probablement aquests límits. Ens restringim a aquestes comarques per congruència amb els límits dels primers estudis de l'Enquesta Metropolitana de la Regió de Barcelona (Subirats 1992). Aquest nucli pot permetre posteriorment estudiar les altres varietats del repertori barceloní, sobretot el català emprat com a segona llengua per castellanoparlants. Briz (1996) en el seu corpus de castellà de l'àrea metropolitana de València no dóna importància a aquest tret, i hi inclou tant parlants monolingües del castellà com bilingües català-castellà.

entorn del 7% del total de la població. Aquests parlants del català com a primera llengua es concentraven en les capes mitjanes i altes, mentre que en els sectors populars, un 18% dels informants tenien el català com a primera llengua. No hem triat el sexe com a criteri rellevant de selecció dels informants.

Tenim dos tipus de materials: els materials 1, nuclears del COS en l'actualitat, i els materials 2, centrats en les ideologies lingüístiques.⁷

Materials 1: Consisteixen en un corpus de setanta entrevistes a parlants de primera llengua catalana de la Regió Metropolitana de Barcelona, recollits en la dècada dels noranta, de tots els sectors socials, a partir dels divuit anys d'edat (vegeu quadre 1).⁸ La finalitat declarada de l'entrevista, a l'hora de concertar-les amb els entrevistats, era la recollida d'enregistraments per a un estudi sobre experiències de vida de les diferents generacions, a la manera de recerques antropològiques o sociològiques. Es volia evitar, així, córrer el risc que l'entrevista es veiés relacionada amb la llengua —suposem que, si fos així, podrien aparèixer reaccions de control

7. Els materials provinents de *Ciudadans*, de què disposem, constitueixen unes dades laterals, que convé esmentar perquè ofereixen possibilitats de recerca clares. Consisteixen en 110 cintes de 45 minuts cadascuna d'enregistraments audiovisuals d'entrevistes a informants de tot Catalunya, provinents del programa *Ciudadans*, dirigit per Francesc Escribano i Joan Ubeda (TVC). Les entrevistes foren fetes pel periodista Francesc Escribano i es feren a sectors que havien viscut experiències vitals importants: separació matrimonial, malaltia greu amb risc de mort, primer fill, decisió d'aïllar-se de la vida pública, etc., de tal manera que el grau de desinhibició és força alt. Sols una petita part d'aquests enregistraments (10 hores) foren emeses per al públic. Disposem de transcripcions no corregides de totes aquestes hores d'enregistraments, ni homogeneïtzades d'acord amb la resta del corpus, i d'una base de dades dels entrevistats en què consten les informacions següents: lloc de naixement, professió, sexe, edat. La majoria de les entrevistes són en llengua catalana.

8. Els membres de l'equip de recerca del COS han estat els següents. Professors universitaris: Núria Alturo, Emili Boix i F. Xavier Vila. Becaris de segon i tercer cicles: R. Aponte, M. Carbó, E. Casasús, E. Curt, M. Galindo, A. Oller, B. Puig. Col·laboradors voluntaris i estudiants de segon i tercer cicle que, a partir dels treballs de curs, han aportat material al COS: I. Acosta, A. M. Albà, M. Alonso, R. Aponte, J. Badiella, G. Barceló, T. Barenys, E. Benito, C. Bertomeu, Ò. Bladas, A. Calzada, M. Cañellas, A. Caparrós, J. Capdevila, M. Carbó, S. Carreras, D. Casals, M. Casals, E. Casasús, M. Cervera, S. Cirbián, O. Clapés, R. Claverol, L. Clotet, E. Collellmir, S. Colomina, C. Corcoll, B. Cormand, E. Coromina, E. Cortés, E. Curt, P. Cuscó, N. Cuxart, A. Deiser, V. Dell'Aquila, M. Dobón, A. Domènech, J. Dorca, M. A. Durà, J. Espuny, G. Espuña, R. Estopà, A. Fargas, S. Farré, M. Fernández, S. Font, J. Fort, J. Freixa, M. Freixinet, E. Fresneda, M. Galindo, J. García, C. Gassó, C. Gelpí, J. Gort, M. Grau, M. de la Guàrdia, R. Guasch, M. Hernández, R. Iglesias, J. Rifà, J. Janariz, G. Jordana, M. Lladó, A. Lladó, Llinàs; M. Lloret, M. M. Llurba, J. Lon, J. Manrubia, L. Martínez, R. Martínez, S. Martínez, G. Martínez, M. M. Massanell, M. M. Mayoral, E. Melero, A. Moll, A. Moneo, X. Monjo, J. C. Mora, J. Moreno, H. Moret, E. Munné, M. Nadal, N. Nogué, E. Padrós, S. Pardas, S. Pascual, R. Pavia, M. Payà, M. P. Perea, J. Pérez, G. Perich, B. Pi, M. Piñol, À. Piró, M. Planell, L. Poblet, M. Prieto, Pujans, B. Pujol, P. Pujol, M. Pujolàs, O. Ramon, E. Ribera, O. Rodríguez, L. Rodríguez, L. Roig, E. Roma, N. Roma, S. Romero, A. Ros, C. de Rosselló, E. Rovira, F. Ruaix, H. Ruiz, T. Ruiz, N. Ruíz, E. Rusinés, E. Sàbat, E. Saiz, J. Salvanyà, N. Sancho, A. Segura, C. Serra, Ll. Serra, M. Solé, F. Soler, R. Soler, M. Sopena, M. Tarradellas, K. Tazawa, C. Teixidó, M. Toro, R. M. Torrens, Torrent, N. Ubach, S. Valero, F. Vall-Llovera, N. Valldeperas, R. Vallès, E. Vendrell, O. Vidal, E. Vide, F. X. Vila, J. M. Virgili, M. Xufré, M. Zapata.

i vigilància metalingüístics, probables en societats amb contacte de llengües i amb polítiques lingüístiques explícites com la catalana.

Materials 2: Consisteixen en un corpus d'entrevistes, encara no completat, al membre de llengua catalana de parelles lingüísticament mixtes (català-castellà), també de la Regió Metropolitana de Barcelona, amb un infant com a mínim entre tres i onze anys d'edat. La finalitat declarada de l'entrevista era un estudi sobre les experiències humanes, socials i educatives de famílies lingüísticament mixtes, tant pel que fa a la formació prèvia de la parella com pel que fa als seus usos i ideologies lingüístics declarats en la parella i en la relació amb els fills, i entre aquests. L'ús i les ideologies lingüístics eren, doncs, un objectiu explícit i declarat de la investigació. Una part de les entrevistes són en castellà, i són excloses del COS.

QUADRE 1 · GRAELLA D'ENTREVISTATS DEL COS

Edat	18-24	25-34	35-44	45-54	55-64	65-75	
Any de naixement Fracció de classe	1980-74	1973-64	1963-54	1953-44	1943-34	1933-23	
Elits, alta burgesia			1		1		2
Tècnics. alts i professionals liberals	1	2	2	1	1	1	8
Empleats i tècnics mitjans	3	1	4	3	2	1	14
Petits empresaris	2	1	2	1	1		7
Comerciants i autònoms		1	3	1	2	2	9
Contramestres i capatassos	1	1	1	1		1	5
Treballadors manuals en actiu	2	4	3	4	2	1	16
Treball. manuals vells i atur					3	5	18
Lumpen					1		1
	9	10	15	12	11	13	70

2.1.2 L'eina de recollida de dades

Usem l'entrevista semidirigida de final obert. En el plantejament del COS s'ha optat per recollir un mateix tipus d'esdeveniment comunicatiu: l'entrevista. Com succeeix en la sociolingüística variacionista i en els treballs de lingüística de corpus (Llisterri 1996), la tria del gènere *entrevista* per a la recollida de dades és el fruit d'una transacció entre una certa homogeneïtzació i una certa qualitat del so, i una

certa representativitat social i el sentit pràctic. S'ha minimitzat al màxim, mitjançant un seguiment molt lax d'un guió entorn de la història de vida dels entrevistats, el caràcter distanciador que és motivat pel fet comunicatiu *entrevista* (Hammer & Wildavsky 1990, Boix 1999). Els entrevistadors han estat, o bé els membres de l'equip del COS, o bé alumnes de cursos de segon o tercer cicle de la Facultat de Filologia de la Universitat de Barcelona.

En el decurs de l'entrevista, abans de concloure, es recullen dades socioculturals de l'informant. Aquestes dades socioculturals proporcionen informació sobre els corresponents capitals econòmics, socials, culturals i informatius, sobretot els lingüístics. En el cas dels materials 2, sobre famílies lingüísticament mixtes, s'aplega una informació detallada sobre la socialització lingüística dels membres individuals de la parella, de la parella mateixa, i de la família amb fills. Especialment interessa conèixer no sols la competència declarada, sinó la relació amb la norma lingüística culta del català que tenen els informants, en la mesura que suposem que influeix en els trets de la seva parla, conjuntament amb la seva posició social (vegeu el precedent de Soler 1986).

2.2 *Tractament de les dades*

2.2.1 Enregistrament

Les entrevistes s'enregistren mitjançant gravadora de so o vídeo. Comptem, per tant, d'una banda, amb enregistraments en suport audiovisual i, de l'altra, amb gravacions només de veu. Quan es porta a terme l'entrevista s'explica a l'informant que es té la intenció d'incloure-la en un corpus i se'n garanteix l'anonimat en tot moment. Si l'entrevistat no té cap inconvenient que el material passi a formar part del Corpus, signa un document en el qual dona el seu consentiment.

2.2.2 Transcripció

Normalment se'n transcriu un fragment de, pel cap baix, mitja hora de durada i no es tenen en compte els cinc primers minuts de l'entrevista perquè poden ser de caràcter més formal. Aquesta transcripció segueix el sistema de transcripció establert pel CORCS (Oller *et alii* 2000, annex 1). Els transcriptors inicials, generalment els mateixos autors de les entrevistes, són estudiants dels cursos de Sociolingüística de segon cicle i de doctorat de Filologia Catalana, col·laboradors voluntaris o membres de l'equip de recerca del COS.

2.2.3 Selecció i classificació de les entrevistes

La història social i sociolingüística fornida pels informants al llarg de l'entrevista o sol·licitada explícitament quan aquesta acaba, permet construir una fitxa social de cada informant. Aquesta història i la mateixa experiència de les entrevistes en què es coneix personalment els entrevistats permeten matisar amb detall els trets socioculturals dels informants i relacionar-los amb els estrictament lingüístics. Així es pot avaluar fins a quin punt els trets socials i lingüístics del parlant entrevistat, responen al perfil social segons el qual havien estat seleccionats. Tot plegat permet:

a) descartar les entrevistes d'aquells informants que no tenen el català com a llengua familiar principal o compartida amb el castellà. Hem decidit excloure aquests parlants de la primera mostra perquè presumiblement els trets que mostren provenen d'un procés d'aprenentatge del català com a segona llengua. Evidentment, abans, s'han refusat les que presenten una qualitat de so dolenta.

b) classificar, amb més coneixement de causa els informants en les caselles de la graella social del COS. Hi ha força casos dubtosos en l'atribució dels informants en les graelles socials, forçosament simplificadores. Sobretot es plantegen problemes en dues situacions. D'una banda, aquells informants grans que han canviat de posició social al llarg de la vida, i de l'altra aquells informants depenents d'altri, com els estudiants i les mestresses de casa. En el primer cas s'ha optat per atribuir-los una adscripció social segons la que tenien en la socialització primària i secundària bàsica. En el segon cas, s'ha adscrit els entrevistats dependents al sector social de les persones de qui depenien (cònjuge o pares, segons escaigui). Les fitxes socials de cada informant permeten sempre, en tot cas, matisar aquestes dades a l'usuari del COS.

(c) excloure totes aquelles entrevistes en què la producció de parla tenia trets de formalitat i distanciament molt marcats en l'entrevista. Trets formals i distanciadors visibles són les respostes mínimes (gairebé telegràfiques) dels entrevistats, el fet que es basin en textos escrits (algun informant d'edat ofereix una història de vida preparada prèviament per escrit), i altres de més opacs, com les autocorreccions freqüents, que assenyalen un alt grau de vigilància metalingüística per part dels informants.

2.2.4 Digitalització

El so de l'enregistrament es digitalitza mitjançant el programa de tractament de veu *SoundEdit 16* (versió 2.1) i, posteriorment, s'estampa en CD-ROM amb el programa *Adaptect Toast*. La digitalització permet esborrar el soroll de fons i destacar

la veu dels informants i els entrevistadors. A més a més, també permet seleccionar fragments de so per tal d'aconseguir una major precisió en la transcripció.

2.2.5 Correcció

Tot i el seguiment d'unes convencions comunes en la primera transcripció, les transcripcions són heterogènies perquè han estat fetes per un grup divers de col·laboradors. Per això, un equip reduït de becaris d'investigació del COS s'encarrega de corregir els textos segons els criteris de transcripció del Corpus Oral de Registres, Conversa Col·loquial i Varietats Socials (CORCS) de la Universitat de Barcelona (vegeu l'annex 1). També s'encarreguen de la correcció de l'entrevista per tal que es correspongui al so amb la màxima fidelitat possible. Per salvaguardar la intimitat de l'informant s'esborren de l'enregistrament totes les referències a noms propis (especialment cognoms) que podrien posar en perill la privacitat de l'entrevistat i se substitueixen per un senyal sonor neutre. A la transcripció, aquestes referències es canvien per mots ficticis amb una estructura semblant, és a dir, s'intenta que tinguin el mateix nombre de síl·labes i es manté la posició de l'accent.

2.2.6 Paral·lelització

Està previst posar en relació la veu i la transcripció escrita de cada entrevista mitjançant d'un programa informàtic de paral·lelització. La paral·lelització permetrà visualitzar la representació gràfica d'una seqüència determinada de so —prèviament seleccionada— al mateix temps que s'escolta.

2.3 Possibilitats d'accés i explotació de les dades del COS

El Corpus compta amb dues bases de dades que aporten informació dels materials: una de parlants i una altra de textos. La primera, la base de dades de parlants, conté les dades següents:

- dades identificatives de la interacció i dels parlants (curs; autor; número d'enregistrament; codi d'identificació de la cinta, del text i de l'autor, i codi d'identificació de la interacció)
- informació sobre la classificació i la digitalització dels materials
- origen geogràfic dels entrevistats i dels seus pares (localitat i barri de naixement, llengua de la comunitat lingüística del lloc de naixement)

- residència actual dels parlants (barri i temps que fa que hi viu) i residències anteriors (barri i durada de l'estada)
- caracterització social dels informants (sexe, edat, estat civil, nombre de fills, ocupació, estudis, estudis de català, estatus socioeconòmic en què s'inscriu el parlant i grup social de l'entrevistat segons la classificació del COS), i
- caracterització lingüística dels parlants (llengua primera de provenença familiar, llengua d'ús oral habitual actual i llengua dominant actual).

La base de dades de textos recull les informacions següents:

- dades identificatives de la interacció (curs; autor; codi d'identificació de la cinta, el text i l'autor; número d'enregistrament i codi d'identificació de la interacció)
- durada i extensió de la interacció (nombre exacte de minuts que s'han enregistrat, nombre de mots del text no homogeneïtzat i nombre de mots del text homogeneïtzat)
- estat del tractament dels enregistraments (programa de tractament de textos amb què ha arribat la transcripció, fase actual del procés d'homogeneïtzació)
- dades de classificació de les transcripcions, i
- qualitat dels materials originals (qualitat del so i qualitat de la transcripció).

Paral·lelament, per a cada entrevista s'ha elaborat una fitxa que conté informació sobre el nivell econòmic, social, cultural i simbòlic de cada parlant.

EMILI BOIX-FUSTER, MIREIA GALINDO I F. XAVIER VILA
Centre Universitari de Sociolingüística i Comunicació,
Universitat de Barcelona

REFERÈNCIES BIBLIOGRÀFIQUES

- ALVAR 1972: *Cuestionario para el estudio coordinado de la norma lingüística culta. II: Morfosintaxis*, coord. Manuel Alvar, Madrid, CSIC.
- BADIA 1962: Antoni Maria Badia i Margarit, *Gramàtica Catalana*, Madrid, Gredos.
- BADIA 1995: Antoni Maria Badia i Margarit, *Gramàtica de la llengua catalana. Descriptiva, normativa, diatòpica, diastràtica*, Barcelona, Enciclopèdia Catalana.
- BOIX 1993: Emili Boix, «El contacte de llengües entre els joves de Barcelona al final dels vuitanta», a *Actes AILLC 9*, III, 283-300.
- BOIX 1996: Emili Boix, «Els materials de llengua oral del corpus de català contemporani de la UB (CUB)», a Payrató (1996b: 93-114).
- BOIX 1998: Emili Boix, «La variació social», *Caplletra*, 25, 135-162.
- BOIX 1999: Emili Boix, «L'entrevista semidirigida en sociolingüística», a *El contacte i la variació lingüística. Descripció i metodologia*, ed. Emili Boix et alii, Barcelona, PPU, 239-283.

- BOIX 2001: Emili Boix, «Varietats socials», a *Enciclopèdia de la Llengua Catalana*, ed. Francesc Vallverdú i J. Bañeres, Barcelona, Edicions 62.
- BOIX 2002: Emili Boix, «Barcelone 2000: un état de la question sociolinguistique», *Terminogramme. Revue de terminologie et aménagement linguistique*, 103-104, 213-243.
- BRIZ 1996: Antonio Briz, «El corpus de conversació col·loquial Val.es.co», a Payrató (1996b: 255-295).
- COLON 1973: Germà Colon, intervenció en el col·loqui posterior a «Morphosyntaxe» d'Antoni Maria Badia i Margarit, a *La linguistique catalane*, París, Klincksieck, 236.
- HAMMER & WILDAVSKY 1990: Dean Hammer i Aaron Wildavsky, «La entrevista semi-estructurada de final obert», *Historia y fuente oral*, 4, 23-57.
- HERNÁNDEZ 1998: C. Hernández García, *Algunas cuestiones más sobre el contacto de lenguas. Estudio de la interferencia lingüística del catalán en el español de Cataluña*, tesi doctoral inèdita, Departament de Filologia Hispànica, Universitat de Barcelona.
- LABOV 1966: William Labov, *The Social Stratification of English in New York City*, Washington DC, Center of Applied Linguistics.
- LABOV 1994: William Labov, *Principles of Linguistic Change. Internal Factors*, Oxford, Blackwell.
- LE PAGE & TABOURET-KELLER 1985: Robert Le Page i Andrée Tabouret-Keller, *Acts of Identity. Creole-based Approaches to Language and Ethnicity*, Cambridge, Cambridge University Press.
- LLISTERRI 1996: Joaquim Llisterri, «Els corpus lingüístics orals», a Payrató (1996b: 27-70).
- OLLER *et alii* 2000: Anna Oller *et alii*, «El COC del CUB: un corpus per a l'estudi de la conversa col·loquial», *ZfK*, 13, 58-91.
- PAYRATÓ 1996: Lluís Payrató, «Transcripció del discurs oral», a Payrató (1996b: 181-216).
- PAYRATÓ 1996b: *Corpus, corpora. Actes del 1e i 2n col·loquis lingüístics de la Universitat de Barcelona*, ed. Lluís Payrató *et alii*, Barcelona, PPU.
- PLA 1995: Joan Pla, «L'obertura de (ə) a Barcelona: el xava i altres varietats», a Turell (1995: 139-159).
- SÀNCHEZ 1996: C. Sánchez, «La selecció de dades lingüístiques: una perspectiva sociològica», a Payrató (1996b: 165-179).
- SOLÀ 1999: Joan Solà, «Intuïció», *Avui*, 16/9/99.
- SOLÀ 2003: Joan Solà, «The Cambridge Grammar of the English Language (3)», *Avui*, 13/3/2003.
- SOLER 1986: J. Soler, *Language Standardization and Norm Foundation in Catalonia*, tesi doctoral, Indiana University.
- SUBIRATS 1992: Marina Subirats, *Enquesta metropolitana 1986. Condicions de vida i hàbits de la població de l'àrea metropolitana de Barcelona 20. Transmissió i coneixement de la llengua catalana a l'àrea metropolitana de Barcelona*, Bellaterra, Institut d'Estudis Metropolitans.
- THIBAUT & VINCENT 1990: Pierre Thibault i Danielle Vincent, *Un corpus de français parlé. Montréal 1984. Historique, méthodes et perspectives de recherche*, Quebec, Université Laval.
- TURELL 1995: *La sociolingüística de la variació*, ed. Maria Teresa Turell, Barcelona, PPU.
- VIAPLANA 1997: Joaquim Viaplana, *Dialectologia*, València, Universitat de València.
- VILÀ & FARGAS 1999: *Normativa i ús de la llengua*, ed. M. Vilà i A. Fargas, Barcelona, Graó.

3. ANNEXOS

3.1 *Críteris de transcripció*

Font: versions 26/6/1995, 4/3/1997, 15/9/1997, 22/6/1998 del document de convencions (Departament de Filologia Catalana, Universitat de Barcelona) [A. Payrató 1996 i Oller *et alii* 2000].

ASPECTES PROSÒDICS

/	entonació ascendent
\	entonació descendent
—	manteniment
:	allargaments

PAUSES I ENCAVALCAMENTS

duració de les pauses

(.)	fins a 1 segon
(..)	fins a 2 segon
(...)	fins a 3 segon
(durada ^m)	més de tres segons

encavalcaments

encavalcaments [text afectat]

[text afectat]

ASPECTES VOCALS

{(fenomen) text afectat}fenòmens observables
simultanis a la producció de parla

tempo

(AC) tempo ràpid (accelerat)
(DAC) tempo lent (desaccelerat)

intensitat

(F) intensitat forta (forte)
(P) intensitat fluixa (piano)

altres

(EV) estrafer la veu
(CAN) text cantat
(??) fragment incert
x seqüències intel·ligibles o conflictives (un
signe per síl·laba)
@ riure (un signe per síl·laba)
(INH) inhalació
(EXH) exhalació

() Comentarís descriptius i pre-ferents: aspectes vocals i gestuals que es puguin descriure amb precisió. Ex.: (s'apaga la música), {(cop de porta) quin mal\}.

(()) Comentarís explicatius i secundaris: moviments, desplaçaments, gestos i sorolls que no es puguin delimitar en l'espai i el temps amb exactitud, i la naturalesa dels quals no es pugui determinar amb precisió. Ex.: ((se sent música de fons)), ((soroll de cops de taula i altres)).

COMENTARIS

3.2 *Fragments d'entrevistes*

Exemple de transcripció 1: CLL és una dona de 80 anys nascuda i socialitzada a Barcelona, al barri del Clot. Va començar a treballar als 8 anys, en una fàbrica del sector tèxtil, i posteriorment es va dedicar a fer usura. Actualment està jubilada i viu en una residència per a gent gran. No sap ni llegir ni escriure, no va anar mai a escola, i diu que no li fa cap falta saber-ne. Ocupa el dia mirant la televisió, les revistes del cor i passejant. El català és la seva llengua materna. Parla un català viu, molt ric en fraseologia, que podríem identificar amb el barceloní col·loquial de la classe treballadora de principis de segle. No manifesta cap tendència política, però reproduceix els tòpics que critiquen l'actitud dels qui manen.

CLL:	el meu pare deia que_	d'haver-se operat\	Pere_
	no li agradava el cine_	(.. 0.84) p(er)ò un dia va trobar_	per què no m'ho deia_
	i es quedava a casa_	una:: infermera_	que: jo hauria fet un tractament\
	(.. 0.53) (INH)	(.. 0.51) i la meva mare tota ne-	(.. 0.54) diu_
	i al quedar-se a casa pues_	gra_	si la meva dona se n'hagués enterat_
	m:_	i jo amb un mocador al cap_	terat_
	m:_	que m'arribava als peus_	em haguera mort\
	coses d'home_	(.. 0.77) i tenia:: tretze anys_	(... 1.24) i és igual morir d'una
	un dia va anar amb un veí_	(... 1.20) @@@ tota negra:_	forma_
	(.. 0.89) se'n anava por ahi_	i que no podies re_	com morir d'una atra\
	d'amagat de:: la mare_	ni mirar_	era molt tranquil\
	perquè la mare_	(.. 0.40) ni anar al cine un any_	molt:: alegre\
	si ho s'hagués enterat_	ni_	
	el mata_	re_	
	(.. 0.66) el mata_	una cosa gran eh/era massa\	
	i_	(.. 0.65) (INH)	
	(.. 0.46) va agafar: mall\	allavons pues la meva mare_	
	(.. 0.86) i des d'allavons_	(.. 0.98) va trobar aquesta_	
	pues la meva mare no es va enterar\	i diu_	
	(.. 0.80) fins que el metge_	i a:ra_	
	el metge li va dir que era la pendis\	però on vas tan negra_	
	hi havia una passa de pendis_	(.. 0.48) a:i el meu Pere_	
	{(F) i tothom tenia pendis_}	diu el teu Pere_	
	(.. 0.87) no miraven ni rayo equis_	si te les ha fet portar:_	
	ni_	(.. 0.63) diu si ha mort d'això_	
	tot xxxx_	de::_	
	{(AC) pendis pendis pendis_}	de mal de dona_	
	bueno_	(INH)	
	(.. 0.83) el van operar_	allavons la meva mare es va deses-	
	i el van trobar_	perar_	
	que estava tot_	i va anar a trobar el metge\	
	(.. 0.47) amb pus\	(.. 0.83) (INH)	
	i allavons_	i el metge li va dir_	
	no hi havia pencilina_	miri_	
	ni pensament\	li dic la veritat\	
	(INH)	(.. 0.32) perquè el seu marit m'ho	
	allavons_	va dir\	
	el meu pare va morir d'així_	(.. 0.79) jo dic_	