

**MANUAL DE TRANSCRIPCIÓ  
DEL DISCURS ORAL  
MATERIALS DE TREBALL**

Òscar Bladas



Barcelona, 2009

Primera edició, 2009

La reproducció total o parcial d'aquesta obra per qualsevol procediment, compresos la reprografia i el tractament informàtics, i la distribució d'exemplars mitjançant lloguer o préstec públics, resta rigorosament prohibida sense l'autorització escrita dels titulars del Copyright, i estarà sotmesa a les sancions establertes per la Llei.

*Aquesta publicació ha comptat amb dos ajuts del ministeri  
(Ministerio de Educación y Ciencia/HUM2005-01936 i Ministerio de Ciencia e  
Innovación/FFI2008-01230/FILO)*

© Òscar Bladas

Edició: Òscar Bladas  
Departament de Filologia Catalana de la UB  
(Secció de Lingüística)

Realització: GrEPAD (Grup d'Estudis de Pragmàtica i Anàlisi del Discurs): Lluís Payrató, Núria Alturo, Josep Besa, Neus Nogué, Òscar Bladas, Marta Payà.

Producció: PPU, S.A.  
Diputació, 213. 08011 Barcelona  
Tel. 93 451 65 70 – Fax 93 452 10 05  
www.ppu.es – ppu@ppu.es

ISBN: 978-84-477-1057-74

Dip. Legal: B-25136-2009

## TAULA DE CONTINGUTS

<i>Nota d'agraïment</i> .....	9
<i>Presentació</i> .....	11
<b>1. INTRODUCCIÓ</b>	
1.1 L'aproximació a la llengua oral .....	17
1.2 La transcripció com a metodologia .....	18
1.3 La transcripció discursiva del Corpus de Català Contemporani de la Universitat de Barcelona (CCCUB) .....	20
<b>2. L'ENREGISTRAMENT DE DADES ORALS</b>	
2.1 Aspectes metodològics de l'enregistrament .....	23
2.2 Aspectes ètics de l'enregistrament .....	27
2.3 Eines de transcripció .....	30
<b>3. LA TRANSCRIPCIÓ DE DADES ORALS</b>	
3.1 Introducció: aspectes generals de la transcripció .....	33
3.2 Principis metodològics de la transcripció .....	36
3.3 Objectius de la transcripció .....	42
3.4 La transcripció discursiva .....	44
3.5 Models de transcripció discursiva .....	46
3.6 La transcripció discursiva de la Universitat de Califòrnia Santa Barbara .....	51

## 4. LA TRANSCRIPCIÓ DISCURSIVA DEL CCCUB

4.1	Descripció general .....	53
4.2	Tractament previ de la transcripció .....	55
4.3	Presentació del context comunicatiu .....	59
4.4	Convencions de la transcripció discursiva del CCCUB .....	65
4.4.1	Unitat de transcripció: el grup tonal .....	65
4.4.2	Seqüències tonals terminals i truncaments .....	71
4.4.3	Aspectes prosòdics conflictius.....	75
4.4.3.1	Llargada dels grups tonals .....	76
4.4.3.2	Vocatus, connectors, marcadors discursius, interjeccions, incisos, tematitzacions, rematitzacions i mots cua .....	79
4.4.3.3	Anotació de les seqüències tonals terminals.....	85
4.4.4	Unitat lingüística: la paraula .....	90
4.4.5	Unitat discursiva: el torn de parla .....	91
4.4.6	Pauses .....	94
4.4.7	Encavalcaments .....	98
4.4.8	Fragments inintel·ligibles .....	104
4.4.9	Aspectes vocals no verbals .....	106
4.4.9.1	<i>Tempo</i> , intensitat i variacions tonals ....	106
4.4.9.2	Riure .....	110
4.4.9.3	Inhalacions i exhalacions .....	112
4.4.9.4	Sons paralingüístics: retroalimentadors i omplidors de parla .....	114
4.4.10	Interjeccions i onomatopeies .....	116
4.4.11	Altres fenòmens vocals: xiuxiueig, cant, etc. ....	118
4.4.12	Gestualitat .....	121
4.4.13	Comentaris del transcriptor .....	126
4.4.14	Regularitzacions .....	132
4.4.15	Regularitzacions del CCCUB .....	136
4.4.15.1	Ús de parèntesis .....	138
4.4.15.2	Fonètica .....	140
4.4.15.3	Morfologia .....	142

4.4.15.4 Lèxic .....	148
4.4.15.5 Regularització de manlleus .....	151
4.4.16 Convencions tipogràfiques .....	154
4.4.17 Errades de transcripció .....	156
4.5 Nivells de la transcripció discursiva del CCCUB .....	165
4.5.1 Transcripcions ampla i intermèdia .....	166
4.5.2 Transcripció estreta .....	171
4.5.3 Recursos complementaris .....	181
<b>5. TRANSCRIPCIÓ, REVISIÓ I EDICIÓ .....</b>	<b>185</b>
<b>CLOENDA .....</b>	<b>189</b>
<b>BIBLIOGRAFIA .....</b>	<b>191</b>
<b>TAULA DE CONVENCIONS .....</b>	<b>201</b>
<b>ÍNDIX TEMÀTIC .....</b>	<b>205</b>
<b><i>Annex: CD-ROM de pràctiques</i></b>	

## *Nota d'agraïment*

Aquest manual no és obra d'un sol autor. El qui signa aquestes línies només s'ha encarregat d'endreçar els documents que el departament de Filologia Catalana de la Universitat de Barcelona ha anat compilant durant anys per elaborar els diversos corpus discursius que formen part del Corpus de Català Contemporani de la Universitat de Barcelona (CCCUB). És, doncs, un manual que han anat escrivint de mica en mica professors, becaris de doctorat i de segon cicle i estudiants, cadascú amb els seus suggeriments —sovint fruit de moltes hores de transcripció—, fins a tenir un bon plec d'apunts que consultàvem, modificàvem i ampliàvem a mesura que s'anaven constituint nous corpus. Ja feia temps, però, que volíem ordenar aquesta documentació i donar-la a conèixer fora del departament com a eina de treball per estudiar la llengua oral des de disciplines com ara la pragmàtica, la sociolingüística i l'anàlisi del discurs.

Tanmateix, el manual que teniu a les mans no hauria estat possible sense la iniciativa de Lluís Payrató, veritable impulsor dels estudis de pragmàtica dins i fora de la Universitat de Barcelona. Sense la seva dedicació a l'estudi de la llengua i a la promoció de la recerca lingüística, ni aquest llibre ni cap dels corpus discursius publicats fins ara pel departament haurien vist mai la llum, i sobretot no haurien estat una font d'inspiració i estudi per a un bon nombre d'investigadors. Per tant, és a ell a qui devem, abans que a ningú, el que es pot considerar el primer manual de transcripció del discurs oral en català.

Tampoc no podem oblidar la feina de la resta de membres del Grup d'Estudis de Pragmàtica i Anàlisi del Discurs (Núria Alturo, Josep Besa, Neus Nogué i Marta Payà), dels professors del departament Emili Boix i Francesc Xavier Vila i dels col·laboradors que a l'inici de la constitució del CCCUB van contribuir a perfilar el model de transcripció que hem aplicat des

d'aleshores (Rosa Aponte, Antoni Bernadó, Maria Carbó, Maria Casals, Eva Casasús, Ignasi Clemente, Elisabet Curt, Mireia Galindo, Albert Soler i Marta Torres), amb un agraïment especial a l'Anna Oller, que metòdicament va posar per escrit les primeres versions de la proposta original. I tot plegat gràcies als materials que han aportat durant anys els estudiants de segon cicle i de doctorat i que han permès fornir els corpus que a poc a poc hem anat transcrivint i revisant.

I, finalment, per a l'elaboració d'aquest manual també hem d'agrair la inestimable col·laboració de Jaume Fitó, Montserrat Grima, Roser Marí, Maria Pilar Perea i —per una fugaç però lucidíssima llambregada a una primera versió— Joan Solà.

Com dèiem al principi, aquest manual no és obra d'un sol autor...

(. 0.13) per això\_  
{(amb sinceritat) mo::ltes gràcies\_  
a tots\}

Ò. B.

## PRESENTACIÓ

### *Sobre els perquès i el com de la transcripció*

En el desenvolupament de la ciència es produeixen punts d'inflexió que permeten el desenvolupament de nous paradigmes o programes de recerca. En aquests punts sol donar-se una confluència de factors —teòrics (amb nous plantejaments o replantejaments d'hipòtesis), metodològics (amb metodologies noves o refoses) o tècnics (sovint en forma d'avanços de tota mena)— que faciliten el canvi.

En la història de la lingüística s'han produït, per descomptat, moltes situacions d'aquesta mena, i el mateix es pot dir, més en particular, de la història de l'anàlisi del discurs i de la parla. Per exemple, igual que, en el seu moment, la possibilitat del recurs a la gravadora va obrir portes a la dialectologia (i tot seguit a la sociolingüística), més recentment el recurs al vídeo i a les noves tecnologies ha permès un salt qualitatiu en la recerca sobre el discurs oral i sobre la multimodalitat comunicativa.

Però els avenços tecnològics, per si sols, no ho explicarien tot, ni de bon tros. De fet, el que s'ha esdevingut durant els últims anys és que els avenços s'han aparellat amb un interès teòric i metodològic creixent per totes les qüestions referents a la transcripció del discurs oral, encara que en certa manera —en aparença— pugui semblar contradictori. L'aparent contradicció neix d'un tipus d'argumentació que concep o imagina la transcripció com una operació mecànica, i per tant susceptible d'acabar sent duta a terme de manera automàtica o automatitzada (i per consegüent, també, susceptible de ser resolta millor per una màquina que per un ser humà). Però ara per ara, en realitat, necessitem la interpretació d'una ment humana per transcriure. Si bé els progressos tècnics ens



estalvien temps i esforços en moltes operacions, la màquina “transcriptora”, “interpretadora” i “analitzadora” del discurs oral i de la comunicació humana en la seva globalitat continua sent una mena d’entelèquia, de desig idíl·lic i d’artefacte més propi d’un context de ciència-ficció que d’un de ciència *real* (l’actual o la d’un futur immediat). Només cal pensar en els anys que han hagut de transcórrer perquè es poguessin fer traduccions automatitzades (i encara de textos més aviat tècnics) i imaginar tot el que necessitaríem —des de la fonètica fins a la pragmàtica— per traduir o analitzar (o transcriure, simplement) un text oral de manera automàtica. Les aparences poden fer pensar a algú que tot plegat no és pas tan complicat, o que no som tan lluny d’aconseguir-ho... però l’estat de les fases actuals, en què ens trobem, demostra que sí.

El fet és que la transcripció, és a dir, en poques paraules, la reconversió o transposició de dades orals al codi escrit, continua sent no només convenient (en qualsevol ocasió), sinó imprescindible (en la majoria de les anàlisis). La raó última és que la transcripció real que solem dur a terme està molt lluny de ser una mena de simple operació de “transvasament” de dades —que és el que justament imagina qui no l’ha practicada. En realitat, la transcripció és una operació de selecció i tractament de les dades que implica anàlisi i interpretació. Per això, des de fa anys, sabem que la transcripció és una operació amb un vessant teòric molt més important que no pas el vessant tècnic que, també, sempre, inclou. En altres paraules, el fet que les repercussions pràctiques (en un sentit corrent del terme) de la transcripció siguin tan evidents i tan importants, no ha d’amagar que encara són més importants totes les decisions teòriques que s’hi aparellen, sense les quals no es pot dur a terme.<sup>1</sup>

La primera de les decisions ha de ser sempre *què* és el que transcrivim (o *descrivim*, si considerem que no cal transcriure-ho), i no pas *com* ho transcrivim (aquest és un

---

<sup>1</sup> Segurament caldria atribuir a Elionor Ochs (1979) de ser la primera a fer aquesta afirmació o, almenys, de ser la primera a l’hora d’argumentar aquesta qüestió amb contundència i habilitat.

aspecte secundari i que es pot resoldre de maneres diverses). I la decisió de què cal transcriure pertany al plantejament teòric de l'anàlisi, al *perquè* de la transcripció, que n'és el punt fonamental: transcrivim per descriure, per analitzar o estudiar, per presentar dades d'una manera més o menys simplificada i pertinent a uns propòsits... i la raó de ser de la transcripció és la que explica el caire de les decisions que s'han d'anar prenent al llarg del procés, un procés que, com sap prou bé qui l'ha practicat, pot arribar a ser molt llarg i costós, en especial, és clar, quan necessitem un grau de precisió o subtileza molt alt.

Atès el que s'ha exposat, no caldria, en absolut, emfatitzar la rellevància i la utilitat del llibre d'Oscar Bladas que presentem. És més que evident, tant com el buit que omplirà. Aquest llibre és, en part, fruit d'una recerca que el GrEPAD (Grup d'Estudis de Pragmàtica i d'Anàlisi del Discurs de la Universitat de Barcelona) va començar fa anys (almenys des de 1990) i que s'ha concretat, entre d'altres treballs, en corpus sobre la llengua corrent i els registres orals (COC, COR i CAP, subcorpus del CCCUB, Corpus de Català Contemporani de la Universitat de Barcelona).<sup>2</sup> Però també, i en la part fonamental, aquest llibre és fruit de l'especialització i la perícia del seu autor. Cap dels corpus esmentats no s'hauria pogut elaborar sense una reflexió, una pràctica i una experiència dilatada sobre la transcripció. Òscar Bladas personifica aquesta etapa i ha estat la persona clau en molts moments del procés, del qual presentem ara, precisament, el resultat més explícit, la culminació. El manual de

---

<sup>2</sup> Tots aquests treballs s'han beneficiat d'ajuts del Ministeri d'Educació i Ciència (l'últim, FFI2008-01230/FILO) i de la Generalitat de Catalunya (l'últim, 2009 SGR 521, per al grup consolidat GEV, Grup d'Estudi de la Variació de la Universitat de Barcelona). En aquests treballs, citats a continuació per ordre cronològic, s'ha fet servir el sistema de transcripció proposat en aquest manual: Payrató, Lluís i Núria Alturo (ed.) (2002): *Corpus oral de conversa col·loquial. Materials de treball*, Publicacions de la Universitat de Barcelona, Barcelona; Alturo, Núria; Bladas, Òscar; Payà, Marta i Lluís Payrató (ed.) (2004): *Corpus oral de registres. Materials de treball*, Publicacions de la Universitat de Barcelona, Barcelona, i Payrató, Lluís i Jaume Fitó (ed.) (2008): *Corpus audiovisual plurilingüe*, Publicacions de la Universitat de Barcelona, Barcelona.

transcripció que ha escrit és una eina molt ajustada i afinada que recull el seu saber i tots els resultats de la recerca que ha dut a terme des de fa molts anys en aquest camp, tant des d'un punt de vista teòric, aprofundint en el perquè de les decisions, com des d'un punt de vista pràctic i aplicat, per exemple fent nombrosos cursos de formació de transcriptors.

El manual està pensat com una eina d'aprenentatge per als no iniciats, però alhora com una eina de perfeccionament per als qui ja en coneixen la problemàtica i les tècniques. Per això inclou des dels aspectes més bàsics i elementals, amb les pràctiques corresponents (en exercicis en un CD-ROM molt ben concebuts per a l'aprenentatge), fins als més complexos, que tenen a veure amb qüestions com el tipus de sistema de transcripció que es farà servir en cada recerca, amb els subsegüents graus d'especialització, amb les notacions (gràfiques), la mena de comentaris o observacions addicionals per part del transcriptor, el reflex de l'entonació i dels aspectes no verbals (i la fidelitat amb què es descriuen), la manera de mesurar i representar les pauses, etc.

La transcripció del discurs oral té molt camí, encara, per endavant, però en aquest llibre veureu que s'hi avança, i molt. Els primers sistemes de transcripció fonètica, que van acabar amb un alfabet fonètic estàndard i internacional, daten de final del segle XIX. Els de transcripció del discurs oral, que no han desembocat encara en un sistema comú, unificat, daten de la segona meitat del segle XX, i fins a final d'aquest segle no apareix la primera obra de recull i anàlisi de sistemes de transcripció.<sup>3</sup> La diferència en el temps entre uns sistemes i els altres és considerable, però tinguem en compte també que ho és igualment la diferència en la complexitat —almenys en l'amplitud— del que s'ha de transcriure.

El llibre que presentem és el primer d'aquesta naturalesa en català i, segurament, en la majoria de llengües romàniques.

---

<sup>3</sup> Edwards, Jane A. i Martin D. Lampert (ed.) (1993): *Talking data: Transcription and coding in discourse research*, Lawrence Erlbaum, Hillsdale, Nova Jersey.

Als mèrits de la rigorositat i sistematicitat del seu autor, i de la seva capacitat didàctica, s'hi han d'afegir els de l'oportunitat, l'originalitat i la capacitat d'innovació. Benvingut sigui el llibre i tant de bo que tots aquests mèrits es tradueixin en el futur en un augment dels estudis sobre el discurs oral, quantitatiu i qualitatiu, així com de la pertinència dels sistemes de transcripció que s'hi fan servir.

Lluís Payrató  
Departament de Filologia Catalana  
Universitat de Barcelona

# 1. INTRODUCCIÓ

“Les noves tecnologies no resolen pas tots els problemes relacionats amb el maneig de les dades [...]. Per això la tècnica de la transcripció adquireix una importància decisiva per a la majoria dels estudis que podem endegar prenent com a base alguna de les modalitats de la llengua oral, siguin de la mena que siguin”, Lluís Payrató (2003: 208).

## 1.1 L'APROXIMACIÓ A LA LLENGUA ORAL

Una de les paradoxes amb què es troba l'investigador interessat en la llengua oral és que, abans de començar qualsevol anàlisi, ha de posar per escrit els mots que vol estudiar. Sense un suport gràfic que deixi constància de les paraules que els informants han pronunciat en un lloc i un moment determinats, i com ho han fet, li serà del tot impossible tirar endavant la recerca amb un mínim de rigor científic. Per això, encara que sembli contradictori, el primer que haurà de fer serà recollir en forma escrita, en paper o en un arxiu informàtic, les dades orals que prèviament haurà enregistrat.

Ens podríem plantejar si actualment cal reproduir per escrit la llengua oral, tenint en compte que els últims avenços tecnològics (magnetòfons i càmeres de vídeo digitals, programes informàtics de tractament de so i imatge, internet, etc.) ens permeten conservar, intercanviar i consultar dades orals amb molta facilitat. Per exemple, si enregistrem en vídeo una conversa col·loquial, caldrà *escriure-la* abans per poder analitzar-la? Tenint a la nostra disposició un enregistrament audiovisual, val la pena dedicar un temps i un esforç gens menyspreables a representar gràficament unes dades que podem recuperar en qualsevol moment? D'entrada, no sembla una operació necessària, ben al contrari. Tot estudi sobre

llengua oral, però, exigeix una anàlisi detinguda i sistemàtica dels mots que s'han fet servir, la manera en què s'han pronunciat i el context en què apareixen, i molt sovint caldrà fer cerques entre materials orals de certa extensió. És aleshores quan es posa de manifest la necessitat de traslladar al paper, transcriure, els mots enunciats oralment. Sense aquest text escrit, l'aproximació a la llengua oral esdevé impracticable.

A continuació se'ns planteja un altre problema, encara més delicat: quin sistema de transcripció farem servir?, i què hem de transcriure?, cal anotar un mot que el parlant ha deixat a mitges o en què s'embarbussa?, i un fragment difícil d'entendre, també haurem de mirar de transcriure'l? Totes aquestes dificultats metodològiques ja eren ben conegudes per la dialectologia clàssica (veg. Veny 1995, Viaplana 1996), però encara s'han posat més de manifest des dels anys 60 fins ara per l'interès creixent per la llengua oral des de disciplines com ara l'anàlisi de la conversa, l'etnografia de la comunicació, la pragmàtica i l'anàlisi del discurs, entre d'altres. Per això actualment la transcripció de dades orals no es concep tant com un procediment necessari, però mecànic i incòmode, que cal seguir abans de començar l'anàlisi pròpiament dita, sinó sobretot com un enfocament metodològic que acabarà condicionant poc o molt els resultats finals de la recerca. Com assenyala Tusón (1995: 81), "podríem dir que amb la transcripció comença l'anàlisi".

## **1.2 LA TRANSCRIPCIÓ COM A METODOLOGIA**

En l'àmbit de les ciències socials és ben conegut el concepte de la *paradoxa de l'observador*, definida com "la modificació (no volguda) de la conducta de la persona enregistrada causada per l'observador o analista en estudiar-lo" (Payrató 2003: 207), que té efectes determinants en mètodes d'obtenció de dades com són l'observació participant, l'enquesta i l'entrevista (semi)dirigida (veg. Durante 1997: 166, Mesthrie 2001: secció 9). Lògicament, la transcripció per si mateixa no

comporta canvis en la conducta de l'informant, però sí que afecta decisivament la manera en què percebrem la realitat que volem estudiar, fins i tot quan disposem d'imatges d'aquesta mateixa realitat. Com veurem al llarg d'aquest manual, resulta impossible traslladar al paper tot el que es diu i tot el que succeeix en qualsevol situació comunicativa, o bé perquè hi ha massa informació a compilar o perquè la transcripció resultant acaba sent un text massa complex i feixuc de llegir. És per aquest motiu que el transcriptor, com un observador més, sempre es veurà obligat a triar la informació lingüística i/o contextual que inclourà en el text i inevitablement influirà en la realitat que l'analista —no sempre el mateix transcriptor— estudiarà amb més detall. Com el càmera que només enregistra el que veu a través de l'objectiu de l'aparell, el transcriptor es converteix en un filtre que interpreta la realitat i que fins i tot la pot arribar a modificar, encara que sigui involuntàriament. Això no vol dir que la transcripció hagi de renunciar a l'objectivitat, és a dir, a reflectir amb el màxim de rigor i fidelitat la llengua oral i el context en què aquesta emergeix. En tot cas, convé tenir present que qualsevol transcripció només podrà ser una aproximació més o menys reeixida a la realitat oral que es vol estudiar, tal com succeeix amb les enquestes, les entrevistes i altres mètodes d'obtenció de dades orals.

A aquests condicionants metodològics cal afegir-hi que la transcripció és un procés que generalment demana un esforç important en temps i recursos. En alguns casos només necessitarem posar per escrit el que ha dit l'informant sense incloure-hi informació contextual i utilitzant les convencions ortogràfiques convencionals; penseu, per exemple, en la transcripció d'una entrevista destinada a un estudi sociològic o antropològic. En la majoria de casos, però, la transcripció serà el procediment per fixar sobre el paper unes dades que haurem enregistrat per mitjà de mètodes molt diferents (enregistraments subrepticis, entrevistes, etc.) i amb propòsits també molt diversos que sovint impliquen utilitzar grafies i símbols especials, segmentar la cadena fònica, alinear el text amb el so o la imatge, marcar aquest mateix text segons formats estàndards

internacionals, etc. Això comporta una feina que pot resultar llarga i costosa, sobretot quan es treballa amb llengua espontània i informal en medis no controlats (amb soroll de fons, parlants que intervenen alhora, etc.). Per aquest motiu la majoria d'autors que han teoritzat sobre la transcripció aconsellen que optimitzem al màxim l'esforç que representa cada text i que triem molt bé el tipus de transcripció que necessitarem d'acord amb l'objectiu del nostre estudi, preveient, ja d'entrada, la informació que haurà de recollir la transcripció, el sistema de notació i l'eina informàtica que farem servir per transcriure i també a quin tipus de lector anirà destinat el text resultant (si és un lector expert o no, per exemple), entre altres qüestions que tractarem al llarg d'aquestes pàgines.

### **1.3 LA TRANSCRIPCIÓ DISCURSIVA DEL CORPUS DE CATALÀ CONTEMPORANI DE LA UNIVERSITAT DE BARCELONA (CCCUB)**

El propòsit d'aquest manual és fer algunes recomanacions sobre transcripció de llengua oral i sobretot donar a conèixer el sistema de transcripció discursiva que utilitza el departament de Filologia Catalana de la Universitat de Barcelona en l'elaboració del Corpus de Català Contemporani de la Universitat de Barcelona (CCCUB) (veg. Payrató *et al.* 1996, Alturo *et al.* 2002; per consultar els materials en línia, veg. el web del dipòsit Recercat) i, més concretament, dels subcorpus destinats a l'estudi de la llengua en el seu context de producció, això és, el Corpus Oral de Conversa Col·loquial (COC) (veg. Payrató i Alturo 2002), el Corpus Oral de Registres (COR) (veg. Alturo *et al.* 2004), el Corpus de Varietats Socials (COS) (veg. Boix-Fuster *et al.* 2006) i el Corpus Audiovisual Plurilingüe (CAP) (veg. Payrató i Fitó 2008).

Aquest model de transcripció, proposat inicialment per Du Bois *et al.* (1991, 1993) per a l'anglès i adaptat al català per Payrató (1995), és un sistema dissenyat per reproduir sobre el paper la llengua oral en qualsevol de les seves modalitats



(formal/informal, espontània/preparada, etc.) juntament amb les circumstàncies comunicatives que l'envolten i alhora la condicionen. Això justifica l'adjectiu *discursiu,-va* amb què ens referirem a aquest sistema de transcripció al llarg del manual, perquè no s'ha de veure tan sols com un conjunt de convencions per posar per escrit uns mots pronunciats en veu alta, sinó sobretot com un model metodològic pensat per incorporar tota aquella informació contextual imprescindible per estudiar el que sovint s'ha anomenat la *llengua situada* (veg. Payrató 1998: 27). És amb aquest enfocament que el CCCUB ha compilat i traslladat al paper tota mena de situacions comunicatives, des de converses col·loquials fins a reunions de feina, entrevistes semidirigides o narracions d'experiències personals, entre d'altres.

El model que descrivim en aquest manual ha anat variant des de l'inici del CCCUB, a principis dels anys 90. Durant aquest temps la pràctica ens ha aconsellat modificar certs criteris de notació per aconseguir un sistema més eficaç i, per exemple, alguns símbols que s'utilitzaven en les primeres versions del model s'han reservat per a transcripcions més detallades o, a la inversa, s'han potenciat grafies que inicialment tenien una funció més marginal. Així doncs, les convencions que exposarem al llarg d'aquestes pàgines són sobretot el resultat de posar constantment a prova el sistema, que també ha anat incorporant molts dels suggeriments i observacions dels transcriptors que han col·laborat en l'elaboració dels subcorpus esmentats més amunt.

En els apartats següents tractem els aspectes més metodològics i tècnics de l'enregistrament i la transcripció de dades orals (veg. § 2), abordem algunes qüestions teòriques relacionades amb la transcripció (veg. § 3) i descrivim en detall el sistema de transcripció discursiva del CCCUB (veg. § 4, 5), que és d'on provenen la majoria d'exemples que analitzarem. Hi afegim també un índex temàtic per facilitar les consultes i un CD-ROM que conté deu pràctiques per familiaritzar-se amb el model de transcripció que aquí descrivim.



## **2. L'ENREGISTRAMENT DE DADES ORALS**

Si bé el propòsit d'aquest manual és donar a conèixer el sistema de transcripció discursiva del CCCUB i sobretot ensenyar a aplicar aquesta tècnica de transcripció, la importància que té l'enregistrament com a mètode d'obtenció de dades ens obliga a referir-nos-hi breument perquè condiciona molt tot el procés de transcripció posterior. Per això en aquest capítol comentem, primer, algunes qüestions metodològiques i ètiques que cal tenir presents a l'hora d'enregistrar una interacció verbal i transcriure-la (veg. § 2.1, 2.2) i, segon, recomanem algunes eines de transcripció (veg. § 2.3).

### **2.1 ASPECTES METODOLÒGICS DE L'ENREGISTRAMENT**

Actualment disposem d'excel·lents recursos tècnics per enregistrar una interacció verbal. A banda dels magnetòfons digitals, que asseguren una qualitat de so òptima i permeten transferir les dades a un ordinador fàcilment, també comptem amb càmeres de vídeo (semi)professionals fàcils de transportar i utilitzar que ens proporcionen una molt bona qualitat de so i imatge. Optarem per un sistema d'enregistrament o un altre segons l'objectiu de la transcripció i l'anàlisi posterior, però les possibilitats tecnològiques actuals fan recomanable que sempre que sigui possible obtinguem imatges de la interacció que volem transcriure, perquè la informació visual que aconseguim ens pot ser de gran utilitat fins i tot si només estem interessats en les dades sonores. Per exemple, si transcrivim una conversa col·loquial amb molts participants només a partir del so, sovint costarà determinar qui ocupa el torn de parla o saber a qui es dirigeix un parlant concret i, òbviament, resultarà impossible

endevinar quins gestos fan servir els informants mentre interactuen. Un enregistrament audiovisual, encara que sigui de baixa qualitat, sempre ens proporcionarà més informació i obtindrem com a resultat una transcripció molt més completa. Evidentment, el gran inconvenient que presenta la gravació audiovisual és que no resulta tan discreta com un enregistrament sonor i pot afectar la naturalitat de les dades.

Els primers enregistraments per a corpus discursius del CCCUB es van fer mitjançant magnetòfons no digitals en situacions poc controlades (un mercat, una casa particular amb el televisor encès, etc.) i sovint d'una manera subreptícia, cosa que no sempre garanteix una qualitat de so apropiada. Una selecció dels millors enregistraments, tant per la qualitat de so com pel tipus d'interacció que representaven, va permetre elaborar el COC i el COR. Per al COS, en canvi, les condicions d'enregistrament eren més controlades (normalment cases particulars o locals amb un micròfon o un magnetòfon molt a prop de l'informant) i això va assegurar unes gravacions sonores de qualitat molt més bona que no comportaven tantes dificultats per als transcriptors, perquè no havien de dedicar tant d'esforç a entendre què deien els informants i es podien concentrar més en altres aspectes de la transcripció. Recentment el CCCUB ha incorporat enregistraments audiovisuals mitjançant càmeres digitals, que garanteixen una qualitat de so i imatge excel·lent, sobretot si la gravació té lloc en un medi controlat. Aquest va ser el procediment del CAP, en què les entrevistes, enregistrades amb dues càmeres de vídeo digitals, es feien en una aula universitària destinada a seminaris.

Així doncs, la qualitat de l'enregistrament resultarà crucial de cara a la transcripció de les dades, perquè com més nítids siguin el so i la imatge menys temps haurà de destinar el transcriptor a esbrinar què diuen els participants en la interacció, especialment si n'hi ha molts i parlen alhora. Es poden donar alguns consells per aconseguir un enregistrament sonor o audiovisual de bona qualitat —de fet, la bibliografia ja en recull força (veg. Bruce 1987, Goodwin 1993, Tusón 1995, Duranti

1997: 455)—, però, en general, no deixen de ser prevencions de sentit comú que, ben aplicades, ens poden estalviar molta feina durant la transcripció. Seguint algunes de les recomanacions que fan Calsamiglia i Tusón (1999: 345) a l'hora d'enregistrar en el que anomenen *entorns naturals*, destaquem els consells següents:

1. Abans de l'enregistrament, comprovarem que la gravadora o la càmera funcionen correctament. Les interferències de l'aparell poden dificultar el procés de transcripció i arribar a fer desestimar la gravació.

2. Estudiarem bé la localització de l'enregistrament per garantir una qualitat de so acceptable. Evitarem al màxim el soroll de fons, ja sigui en un lloc tancat (apagant ràdios, televisors i altres aparells que puguin molestar; tancant les finestres per evitar el soroll del carrer, etc.) o en un lloc obert (buscant carrers poc transitats, acostant el micròfon a l'informant, etc.). Hem de tenir en compte que el soroll de fons sempre queda gravat amb un volum molt més alt del que nosaltres percebem *in situ*, de manera que, per exemple, el so d'una porta en tancar-se, encara que sigui fluix, en l'enregistrament pot semblar un cop molt fort. També hem de procurar acostar-nos tant com puguem a l'informant, perquè, si no, correm el risc que la veu d'algú altre més proper al micròfon o qualsevol soroll li tapin la veu. Si l'enregistrament és subreptici i el magnetòfon està amagat en algun lloc (una bossa, un racó d'un menjador, etc.), hem de preveure que qualsevol moviment de l'aparell o fregament amb la roba o qualsevol altre material pot distorsionar el so. Per això sempre que sigui possible és recomanable fer una prova pilot abans de gravar tant si l'enregistrament té lloc en un medi controlat com si no.

3. Un cop fet l'enregistrament, i abans de començar a transcriure, comprovarem que la qualitat del so de tota la gravació és prou bona per treballar-hi. Una reproducció íntegra

de l'enregistrament ens permetrà decidir si podrem transcriure'l amb garanties o si, per contra, hem d'enregistrar una altra vegada en unes condicions més favorables. El que hem d'evitar és que, un cop iniciat el procés de transcripció, hàgim de desestimar l'enregistrament perquè en un moment determinat perd qualitat de so o imatge. Òbviament, també ens assegurarem que el material que hem enregistrarat és l'adequat per a l'anàlisi que ens proposem. Per això és aconsellable fer una prospecció prèvia i determinar a partir d'una mostra representativa si les dades que hem recollit són adients per a la recerca o si cal compilar-ne de diferents.

4. Aprofitarem la memòria fresca després de la gravació endreçant les notes que hàgim pres i apuntant les dades rellevants que encara recordem immediatament després de l'enregistrament. Una bona manera de recuperar aquesta informació és repassar de seguida la gravació i anotar tot el que pugui ser pertinent per a l'anàlisi, cosa que també ens servirà per familiaritzar-nos amb la mostra i saber de quines parts consta. Calsamiglia i Tusón (1999: 347) recomanen fer un guió de l'enregistrament per trobar fàcilment els fragments que ens interessin més. Si no hem estat testimonis directes de la gravació, també haurem d'obtenir la informació contextual com més aviat millor per aprofitar que la persona encarregada de dur-la a terme encara pot recordar-ne molts detalls.

Aquesta tasca de compilació de dades contextuais és molt convenient fins i tot si disposem d'imatges, perquè, com hem apuntat més amunt (veg. § 1.2), tota aproximació al fet de parla sempre serà necessàriament parcial. La càmera només capta una part de la realitat i per això haurem de recollir qualsevol incidència que no hagi quedat enregistrada i que pugui tenir un efecte decisiu en la gravació (per exemple, que l'informant es mostri molt nerviós o insegur).

5. Farem almenys una còpia de seguretat de l'enregistrament.

Una vegada hem obtingut les dades, les transferirem a un editor de so o imatge per seleccionar els fragments de parla que transcriurem. La tria dependrà, evidentment, dels objectius de la transcripció i de l'anàlisi posterior. Podem optar per transcriure tot l'enregistrament o transcriure'n només els segments més rellevants per a la recerca. La primera opció exigeix un esforç en recursos i temps molt superior, sobretot si la interacció és informal (per exemple, una conversa col·loquial), però com a contrapartida permet disposar en tot moment de fragments de parla ja transcrits i la feina feta pot ser útil per a altres anàlisis i altres investigadors. La segona opció no comporta tant d'esforç de transcripció però sí una anàlisi aprofundida de les dades abans de decidir què transcrivim. Si aquest estudi previ no és prou acurat, ens haurem de plantejar més endavant si transcrivim fragments que no havíem tingut en compte o si descartem transcripcions ja acabades.

Finalment, si disposem d'un enregistrament audiovisual, també és aconsellable exportar el senyal d'àudio a un editor de so o un programa especialment dissenyat per a la transcripció (veg. § 2.3). Transcriure amb editors d'imatge acostuma a ser més lent —sobretot si no necessitem alinear text i imatge— i, per tant, estalviarem temps si fem una primera transcripció només a partir del so i després la completem amb els visionats que calguin. Hem de tenir present que per obtenir una transcripció força minuciosa com la que aquí descriurem, en què marquem els encavalcaments i la durada de les pauses, entre altres aspectes contextuais, ens farà falta un suport informàtic amb uns comandaments que ens permetin avançar i retrocedir amb agilitat al llarg del senyal acústic, perquè haurem d'escoltar diverses vegades un mateix fragment abans de donar-ne per bona la transcripció.

## **2.2 ASPECTES ÈTICS DE L'ENREGISTRAMENT**

Quan gravem en un medi no controlat, una de les primeres qüestions que sorgeixen és si hem d'avisar els

informants que seran enregistrats. D'una banda, una gravació subreptícia garanteix que els participants en el fet de parla actuaran amb absoluta espontaneïtat —per bé que probablement cada vegada hi ha menys inhibició davant de càmeres i micròfons— i resollem, en part, la paradoxa de l'observador (veg. § 1.2). D'altra banda, un enregistrament ocult sovint implica una qualitat de so i imatge deficient i sobretot contravé el dret dels informants de preservar la seva privacitat. Per tant, si optem per registrar d'amagat, haurem de ser conscients que podem vulnerar la intimitat dels que participen en la interacció verbal i que fins i tot ens poden reclamar responsabilitats per no haver-los advertit prèviament. Tot i que no deixa de ser problemàtica, una manera d'evitar conflictes és informar els parlants que han estat enregistrats un cop feta la gravació i, òbviament, en cas que no ens donin el consentiment per fer-la servir, destruir-la.

En estudis en què s'enregistra diverses vegades els mateixos informants hi ha la possibilitat de demanar-los permís per gravar, però sense avisar quan es grava, per obtenir mostres de parla més espontània. En les entrevistes també és recomanable dedicar els primers cinc minuts a tractar temes poc rellevants per a l'anàlisi perquè els informants s'acostumin a la presència del magnetòfon o la càmera, tal com es fa en el COS, en què no es transcriuen els primers minuts de l'enregistrament.

En tot cas, després de la gravació, sigui subreptícia o no, sempre demanarem permís per poder transcriure-la i analitzar-la. Per això haurem preparat amb antelació un document que els informants ompliran amb les seves dades personals i signaran per acreditar que ens donen el seu consentiment perquè utilitzem l'enregistrament amb finalitats estrictament acadèmiques. Com a mostra, reproduïm a sota el full amb què el CCCUB obté aquest permís.




 <p>UNIVERSITAT DE BARCELONA</p>	<p>Facultat de Filologia          Departament de Filologia Catalana          Gran Via de les Corts Catalanes, 585          E-08007 Barcelona</p>
<p>El sotasignat/La sotasignada, participant en un enregistrament que s'inclourà en el Corpus Oral del Departament de Filologia Catalana de la Universitat de Barcelona, dono el meu permís perquè els materials recollits puguin ser utilitzats amb finalitats associades amb la recerca acadèmica.</p>	
<p>Lloc i data: ..... .., ..... de .....</p>	
<p>Signatura:</p>	
<p>Nom i cognoms:</p>	
<p>DNI:</p>	

Figura 2.1 Full del CCCUB per demanar el permís dels informants

Per la seva banda, els investigadors també s'hauran de comprometre a destinar tant l'enregistrament com la transcripció a usos exclusivament acadèmics i garantir l'anonimat dels informants ometent totes aquelles informacions que puguin servir per identificar-los. Si cal, lliuraran a cada informant un document que certifiqui que els materials enregistrats s'utilitzaran únicament per a la recerca i que s'eliminaran totes les dades personals que hi apareguin.

Si només treballem amb so, ometrem tant de l'enregistrament com de la transcripció els cognoms dels informants i de terceres persones i totes aquelles dades (números de telèfon, noms d'empreses i establiments comercials, etc.) que puguin ajudar a identificar-los. En la gravació distorsionarem el fragment de parla en què apareixen dades personals mitjançant l'editor de so que utilitzem i en la transcripció les substituïrem per dades

fictícies (veg. § 4.4.13). Si treballem amb imatge, no té sentit amagar la identitat dels informants, però també tindrem cura d'evitar la difusió de cognoms, números de telèfon, adreces, etc. tant dels mateixos informants com de terceres persones.

## 2.3 EINES DE TRANSCRIPCIÓ

El mercat ofereix tot un ventall de programes informàtics per transcriure la llengua oral. Habitualment el CCCUB ha fet servir l'editor de so Sound Forge (veg. Sony Creative Software), amb què podem visualitzar l'amplitud del senyal acústic a partir d'arxius d'àudio en format WAV (veg. figura 2.2) i alhora mesurar les pauses i delimitar els encavalcaments amb precisió. També ens permet distorsionar el so per mantenir l'anonimat dels informants i aplicar filtres per eliminar el soroll de fons i millorar la qualitat del so. És un programa fàcil d'utilitzar que afavoreix una transcripció àgil, especialment si volem obtenir textos d'una certa extensió amb poca informació contextual. El principal inconvenient que presenta és que no ha estat dissenyat expressament per a tasques de transcripció, de manera que alhora hem de tenir obert un editor de textos per escriure el que anem escoltant i hem de passar d'un programa a l'altre contínuament.

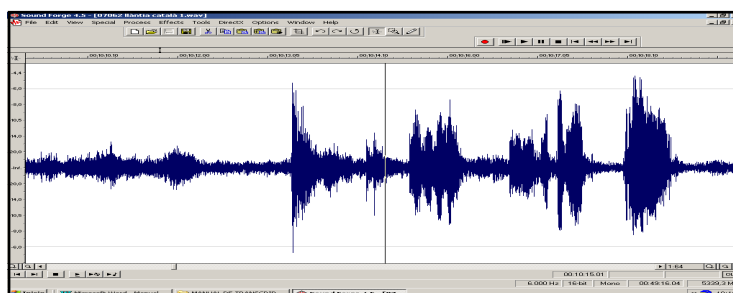


Figura 2.2 Imatge de l'editor de so Sound Forge

Pel que fa als programes pensats específicament per transcriure llengua oral, cal esmentar-ne quatre. D'entrada, el Voice Walker (veg. Universitat de Califòrnia Santa Barbara) i el

Soundsciber (veg. Michigan Corpus of Academic Spoken English), que, tot i que són força elementals i també demanen tenir un editor de textos obert, resulten idonis per a transcripcions en què no necessitem anotar els encavalcaments ni la durada de les pauses. Són més complets els programes WaveSurfer (veg. KTH Royal Institute of Technology) i Transcriber (veg. Sourceforge), que com el Sound Forge permeten visualitzar l'amplitud d'ona, però tenen l'avantatge que disposen de panells per escriure-hi el text i alinear-lo amb el so, cosa que els converteix en eines molt útils per a transcripcions que han d'incorporar molta informació contextual (durada de les pauses, encavalcaments, etc.) (veg. figura 2.3).

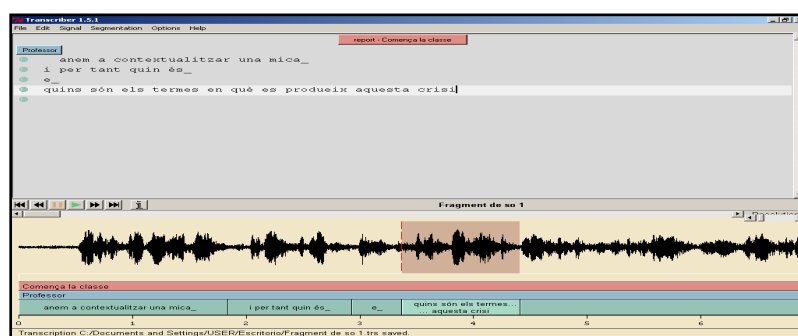


Figura 2.3 Imatge del programa de transcripció Transcriber

A més d'aquests programes, cal mencionar el SoundWriter (veg. Universitat de Califòrnia Santa Barbara), recomanable sobretot per a transcripcions en què segmentem el discurs en grups tonals (veg. § 4.4.1) i hem de delimitar amb molta precisió les pauses, els encavalcaments, etc. La seva aplicació principal, però, és alinear text i so, de manera que s'aconsella que ja es disposi d'una transcripció prèvia abans de fer-lo servir.

Per a més informació sobre editors de so i programes de transcripció, consulteu també el web de Joaquim Llisteri, en què s'ofereix una descripció detallada i actualitzada dels recursos informàtics amb què compten els investigadors per analitzar dades sonores.

D'altra banda, si analitzem imatges, també tenim a l'abast un bon grapat de programes per treballar. A part de l'Adobe Premiere (veg. Adobe Systems), que ens permet segmentar i editar les imatges que ens interessa transcriure, disposem del Transana (veg. Universitat de Wisconsin-Madison) i l'Elan (veg. Max Planck Institute for Psycholinguistics), que és l'editor d'imatges que ha emprat el CCCUB per elaborar el CAP (veg. figura 2.4). Tots dos programes estan especialment dissenyats per editar i transcriure dades audiovisuals i alinear-les amb el text i també permeten importar transcripcions d'altres aplicacions, com ara el Transcriber, cosa que pot ser molt útil per a textos d'una certa extensió i complexitat.

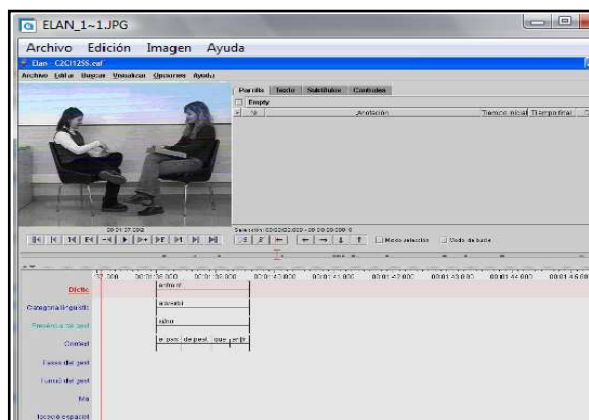


Figura 2.4 Imatge del programa Elan

Uns altres programes que s'utilitzen habitualment per transcriure dades audiovisuals són l'Exmaralda (veg. Universitat d'Hamburg) i el Clan (veg. Childes), tots dos amb sistemes de transcripció propis (veg. § 3.5). Per a transcripcions que demanen una notació fonètica i prosòdica particularment exhaustiva, s'han desenvolupat, entre d'altres, les aplicacions Praat (veg. Universitat d'Amsterdam) i Tobi (veg. Ohio State University).

### 3. LA TRANSCRIPCIÓ DE DADES ORALS

#### 3.1 INTRODUCCIÓ: ASPECTES GENERALS DE LA TRANSCRIPCIÓ

La bibliografia defineix la transcripció com el “procediment de trasllat o transposició a una forma escrita d’unes dades que originalment s’han produït a través del canal oral” (Payrató 2003: 208). Podem anomenar *transcripció*, doncs, qualsevol sistema que utilitzem per reflectir sobre el paper un producte lingüístic enunciat oralment, fins i tot quan prové de la lectura d’un text ja escrit, com ara un poema, un sermó o una declaració pública.

El discurs oral no té les mateixes característiques que el discurs escrit. L’entonació, la gestualitat, l’espontaneïtat i l’expressivitat, entre altres factors, modelen la parla per crear un producte lingüístic distint del que elaborem damunt un full de paper o en la pantalla de l’ordinador. Per tant, si ens interessa traslladar la llengua oral a un suport gràfic, necessitarem un mètode que ens permeti posar-la per escrit sense perdre-hi massa informació, perquè inevitablement el text mai podrà reproduir amb tota exactitud el que s’ha enunciat en veu alta. Són dos canals diferents i, doncs, dues maneres de comunicar diferents. Fixem-nos en el fragment de (1), extret d’una entrevista del COS.

- (1) LLU: als setze als quinze anys la meva mare va començar que jo havia d’aprendre a ballar sí i que jo havia d’aprendre a ballar com que no em va deixar tenir mai cap amiga mai la millor amiga la mare la millor amiga la mare escolti vaig sortir atontada atontada atontada en tot bueno que havia d’anar a ballar que havia d’anar a ballar  
(COS: LLU, 638-657, adaptat)

Un text com (1) mai correspondrà al que la informant LLU (una dona de 80 anys del barri del Clot de Barcelona) va enunciar en el moment de ser entrevistada, encara que tots dos fragments,

l'oral i l'escrit, continguin els mateixos mots. En passar d'un canal a l'altre hem de prescindir per força de la fonètica, l'entonació i la gestualitat que caracteritzen —més que no pas complementen— el discurs oral, així com de les vacil·lacions, les pauses i els sons orals no verbals propis de la parla espontània, entre altres aspectes (veg. Duranti 1997: 188). En conseqüència, més que intentar reflectir amb tota minuciositat damunt el paper el que hem escoltat prèviament, haurem de decidir quina informació ens interessa més (fonètica, prosòdica, morfològica, discursiva, etc.) i enfocar la transcripció de cara a aquesta mateixa tria.

Segons la informació que ens convingui recollir, optarem per un tipus de transcripció o un altre. Un dels sistemes més coneguts i universals és la transcripció fonètica mitjançant l'Alfabet Fonètic Internacional (AFI), creat per representar gràficament les característiques acústiques i articulatòries dels sons de la cadena fònica (veg. Bonet *et al.* 1997). Vegeu a continuació el fragment de (1) transcrit en AFI.

(2) LLU:  
.....  
.....  
  
.....  
.....  
  
.....  
.....  
  
.....  
.....  
  
.....  
.....  
.....

Altres mètodes de transcripció se centren més en el contingut semàntic i ideològic del discurs i adopten convencions més pròpies de les obres de teatre, els guions de cinema, etc., i segueixen les normes ortogràfiques de cada llengua en particular, tal com representem a (3).

- (3) LLU: als setze, als quinze anys, la meva mare va començar que jo havia d'aprendre a ballar, sí, i que jo havia d'aprendre a ballar. Com que no em va deixar tenir mai cap amiga, mai —la millor amiga, la mare; la millor amiga, la mare, escolti!—, vaig sortir atontada, atontada, atontada en tot. Bueno, que havia d'anar a ballar, que havia d'anar a ballar!

I altres models de transcripció estan pensats per reflectir com es va desgranant el discurs, de manera que també recullen aspectes típicament orals com són les pauses, els allargaments vocàlics i els encavalcaments entre parlants, i utilitzen grafies no estàndards i altres recursos gràfics, com ara parèntesis i claudàtors, per aproximar-se a la pronúncia real; compareu (3) amb la transcripció discursiva de (4).

- (4) 1 LLU: (*pausa*) aaals setze-- als quinze anys, la me(v)a mare va  
 2 començar, que jo havia d'aprendree, a ballar, (*pausa*) sí.  
 3 I que jo havia d'aprendre a ballar, com que no em va dixer  
 4 tenir mai cap amiga, (*pausa*) mai. (*pausa*) la millor amiga,  
 5 la mare, la millor amiga la mare, escolti, vai sortir, (*pausa*)  
 6 atontada. (*pausa*) atontara, atontara en tot. (*pausa*) Bueno,  
 7 (*pausa*) que havia d'anar a ballar, que havia d'anar a ballar.

A més, una mateixa transcripció pot presentar diverses versions segons el grau de detall amb què reproduceix el discurs oral. Pot ser més ampla, és a dir, oferir una visió més general del fragment de parla fent servir una notació mínima, com en el cas de (3) i (4), o bé més estreta, amb una descripció més completa i una notació més exhaustiva, com la transcripció de (5), adaptada del COS.

- |     |   |                              |    |                     |
|-----|---|------------------------------|----|---------------------|
| (5) | 1 | LLU: (.. 0.76) a::ls setze-- |    | dixar tenir mai cap |
|     | 2 | als quinze anys_             |    | amiga_              |
|     | 3 | la me(v)a mare va            | 9  | (... 1.09) mai\     |
|     |   | començar_                    | 10 | (. 0.26) la millor  |
|     | 4 | que jo havia                 |    | amiga_              |
|     |   | d'aprendre:_                 | 11 | la mare_            |
|     | 5 | a ballar_                    | 12 | la millor amiga la  |
|     | 6 | (... 1.10) sí\               |    | mare_               |
|     | 7 | i que jo havia               | 13 | escolti_            |
|     |   | d'aprendre a ballar_         | 14 | vai(g) sortir_      |
|     | 8 | com que no em va             | 15 | (.. 0.54) atontada\ |

16	(... 1.46) atontara_		d'anar a ballar_
17	atontara en tot\	20	que havia d'anar a
18	(.. 0.41) bueno_		ballar_
19	(.. 0.62) que havia		(COS: LLU, 638-657, adaptat)

Finalment, cal distingir l'anomenada *transliteració* del que és pròpiament la transcripció. La transliteració consisteix a posar per escrit el discurs oral mitjançant les normes ortogràfiques convencionals, com vèiem a (3), mentre que la transcripció reflecteix la llengua oral adaptant-hi, si cal, les convencions ortogràfiques i fent servir codis i símbols específics, com a (4) i (5) (vegeu com a mostra de corpus transliterats Briz *et al.* 2002, Albelda Marco 2005, Hidalgo i Sanmartín 2005). Tots dos procediments presenten avantatges i inconvenients. La transliteració obvia molts aspectes del discurs oral, com ara la prosòdia, però té l'avantatge que facilita les cerques dins el text i la difusió de materials. En canvi, la transcripció s'acosta d'una manera més fidedigna a la llengua parlada, però com a contrapartida ha de recórrer a solucions idiosincràtiques que poden dificultar tant la cerca com la difusió de dades.

### 3.2 PRINCIPIS METODOLÒGICS DE LA TRANSCRIPCIÓ

Edwards (1993, 2003), una autora que ha teoritzat a bastament en aquest camp, assenyala que en molts treballs de recerca la transcripció de dades orals rep poca atenció, perquè només s'hi fa referència al final com a annex i no es detallen els criteris que s'han aplicat, malgrat la repercussió que això pot tenir en el resultat final de l'anàlisi. Segons Edwards, tota transcripció ha de tenir en compte els principis metodològics que resumim a continuació:

1. *Principis d'elaboració de categories.* Les categories que s'utilitzin en una transcripció seran discriminables, és a dir, cada grafia correspondrà clarament a un sol fenomen, com a l'AFI, en què cada símbol representa un so o un tret fonètic determinat; altrament, el sistema esdevindria confús i inoperatiu. Les



categories també seran exhaustives, per poder reflectir el màxim de fenòmens lingüístics i/o discursius, però alhora es faran servir de manera contrastiva per recollir només aquells fenòmens rellevants per a l'anàlisi. Per exemple, si una transcripció està pensada per estudiar la gestió dels torns de parla, és lògic que inclogui sobretot símbols per anotar encavalcaments, pauses, truncaments, etc., més que no pas per indicar, posem per cas, elisions vocàliques.

2. *Principis de legibilitat.* La transcripció procurarà facilitar la lectura i, per tant, explotarà tots aquells recursos que ajudin a interpretar el text d'una manera àgil, com ara el subratllat i la negreta, entre altres recursos tipogràfics que n'accentuen la prominència visual. Edwards (1993) també recomana les estratègies següents:

a) *Proximitat d'elements relacionats.* Evitarem que dos o més elements que estiguin molt vinculats entre si quedin massa separats. Per exemple, la informació prosòdica, com ara un ascens tonal representat per /, l'adjuntarem al fragment de parla afectat (*ara plou/*) i eliminarem els espais en blanc o altres elements que quedin entremig (cf. *ara plou /*). També compactarem al màxim la informació i, per exemple, procurarem integrar els símbols prosòdics dins el text, més que no pas situar el text en una línia i els símbols en una altra.

b) *Separació visual.* Els elements de naturalesa diferent, com ara els comentaris del transcriptor, els marcarem d'una manera ben visible per evitar confusions amb el text de la transcripció, per exemple, posant-los en cursiva.

c) *Ordre temporal.* Sempre que sigui possible, respectarem l'ordre cronològic del discurs oral per reflectir adequadament com s'han anat encadenant els enunciats al llarg de l'esdeveniment comunicatiu i no entorpir la lectura del text.

d) *Marcatge mnemònic.* Triarem símbols que ens permetin reconèixer amb facilitat què volen significar. Per exemple, podem representar l'entonació descendent i ascendent mitjançant les grafies icòniques \ i / respectivament.

e) *Nombre mínim de símbols.* No utilitzarem més símbols dels que calguin per no enfarfegar massa la transcripció. Triarem aquests símbols en funció de l'anàlisi a què sotmetrem el text transcrit, de manera que prescindirem d'aquells que no aportin informació rellevant. Ehlich (1993) hi afegeix que la transcripció ha de ser fàcil de modificar, perquè a mesura que avança l'anàlisi puguem corregir errades que detectem en el text o incloure-hi més informació.

3. *Principis de tractament computacional.* La transcripció ha de facilitar les cerques informatitzades, perquè, si no, ens exposem a trobar-nos dos problemes. El primer és que no recuperem tota la informació disponible. Per exemple, si fem servir la grafia no estàndard *aiga*, serà impossible que després la puguem recuperar fent una cerca de partir de la forma estàndard *aigua*. Això ho podem resoldre escrivint la forma estàndard al costat de la grafia no estàndard o preparant una taula de correspondències (*aigua-aiga-aigo-auia*) abans de començar la cerca. El segon problema és l'invers, és a dir, recuperar més informació de la necessària, cosa que sol succeir si, per exemple, no exclouem de les cerques el text contingut en els comentaris del transcriptor. Ho solucionarem ometent de la transcripció tota aquella informació que pot invalidar les cerques o marcant-la amb un format diferent perquè el programa no la tingui en compte. Per a cerques automàtiques a partir de corpus, marcats o no, consulteu el web de Joaquim Llisterra, en què es descriuen els programes d'extracció de dades que més s'utilitzen actualment (WordSmith, Concordance, MonoConc Pro, etc.).

Edwards subratlla que seguir aquests principis metodològics també implica tenir en compte tot un seguit de qüestions que acabaran condicionant la nostra feina com a transcriptors, però també com a lectors de textos ja transcrits. Una transcripció ha de ser operativa tant a l'hora de traslladar les dades al paper com a l'hora d'extreure'n informació: si resulta massa completa per als nostres interessos com a analistes,

haurem esmerçat un temps i un esforç innecessaris; i a la inversa, si és massa ampla, haurem de repassar-la i acabar de completar-la més endavant, cosa que també comportarà un esforç extra. Per això abans de començar a transcriure és fonamental plantejar-nos quins són els nostres objectius d'anàlisi i com hi adequarem el sistema de transcripció que farem servir.

A continuació, i seguint també Edwards (1993, 2003), esmentem alguns aspectes conflictius especialment relacionats amb els models de transcripció discursiva com el que aquí ens ocupa:

1. *Disposició del text dins la pàgina.* En la bibliografia (veg. Ochs 1979; Edwards 1993, 2003) trobem diverses maneres de distribuir el text dins la pàgina, per exemple: en format vertical (6a), en columna (6b) o en partitura (6c). Cadascun d'aquests sistemes presenta avantatges i inconvenients. El format vertical reflecteix molt bé l'ordre cronològic del discurs, però no permet aparellar fàcilment els encavalcaments —a (6) representats per claudàtors. Per contra, els formats en columna i en partitura representen més gràficament els encavalcaments, però sembla que no faciliten tant la lectura.

(6a) 1 EEG: Li agrada molt la pintura i, els llibres de pintura...  
 2 PPM: Mhm. Que pinta, tam[bé...]  
 3 EEG: [No] no no no. Ui no. És un negat, completament, per pintar...  
 4 PPM: Ha! [ha! ha! ha!]  
 5 EEG: [Però li agrada x]  
 6 PPM: Ha! [ha! ha! ha!]  
 7 HHM: [Ha! ha!]  
 (COS: EEG, 448-461, adaptat)

(6b)	EEG:	
	1 Li agrada molt la pintura i,	
	2 els llibres de pintura...	
	PPM:	
	3 Mhm. Que pinta,	EEG:
	4 tam[bé...]	[No] no no no. Ui no. És un negat,
	5	completament, per pintar...
	PPM:	EEG:
	6 Ha! [ha! ha! ha!]	[Però li agrada x]

PPM:	HHM:
7 Ha! [ha! ha! ha!]	[Ha! ha!]
(COS: EEG, 448-461, adaptat)	

- (6c)
- 1 EEG: Li agrada molt la pintura i, els llibres de pintura...
  - 2 PPM: Mhm. Que pinta, tam[bé...]
  - 3 EEG: [No] no no no. Ui no. És un negat,  
completament, per pintar...
  - 4 PPM: Ha! [ha! ha! ha!]
  - 5 EEG: [Però li agrada x]
  - 6 PPM: Ha! [ha! ha! ha!]
  - 7 HHM: [Ha! ha!]
- (COS: EEG, 448-461, adaptat)

2. *Unitat de transcripció.* Una de les primeres qüestions que hem de considerar és com segmentarem el text de la transcripció. La solució tradicional, pròpia de les obres de teatre, els guions de cinema, etc., és distribuir el text per torns de parla i puntuar-lo d'acord amb criteris bàsicament textuais. Els analistes de la llengua oral, però, prefereixen utilitzar unitats delimitades per la prosòdia, com ara els grups tonals (veg. § 4.4.1).

Ara bé, si decidim segmentar el discurs en unitats entonatives, hem de tenir present que treballar amb prosòdia demana una certa pràctica i fa molt més costosa la transcripció, perquè caldrà escoltar moltes més vegades un mateix fragment de parla fins a determinar-ne les característiques tonals. A més, com apunta Edwards (1993: 21), no només és difícil percebre els diferents matisos entonatius del discurs oral, sinó també reflectir-los en el paper; i fins i tot si mesurem els correlats acústics de la cadena fònica, sempre haurem d'interpretar les dades que obtinguem. D'altra banda, tampoc no hi ha unanimitat de criteris pel que fa a la notació prosòdica, perquè els autors que assenyalen els moviments tonals sovint fan servir punts de referència diferents: el final del grup entonatiu només (amb símbols d'ascens, descens i manteniment), el fragment comprès entre l'accent més prominent i el final del grup tonal (amb símbols d'ascens-descens, descens-ascens, etc.) o tot el grup tonal. Així doncs, abans de transcriure,

haurem de decidir fins a quin punt necessitem segmentar el discurs en unitats prosòdiques i, en aquest cas, com les anotem. Lògicament, la decisió dependrà del tipus d'estudi que vulguem desenvolupar i del temps i els recursos (nombre de transcriptors, infraestructura informàtica, etc.) de què disposem.

3. *Durada i posició de les pauses.* En moltes transcripcions discursives és rellevant indicar la durada de les pauses que es produeixen al llarg del discurs oral, però a vegades resulta difícil mesurar-la, ja sigui per factors fonètics (cal comptar el temps d'implosió d'una consonant oclusiva que segueix un silenci?) o ambientals (sovint hi ha soroll de fons que no permet distingir clarament els límits de la pausa). A més, determinar la durada de les pauses de tota una interacció verbal comporta força temps, per la qual cosa haurem de valorar, abans de començar a transcriure, si ens cal recollir aquesta informació. Si és així, també haurem de decidir on col·loquem els símbols de pausa dins el text. Per exemple, situats al final de torn de parla, sembla que atorguem el silenci només al parlant que ha cedit el torn; al principi, al parlant que l'ocupa; i entre línies, a tots dos interlocutors.

4. *Formes lingüístiques no estàndards.* Per a determinades anàlisis no ens caldrà transcriure les solucions no estàndards que apareixen en el discurs oral i n'hi haurà prou fent servir la grafia convencional, de manera que si un informant diu, posem per cas, *vegiar*, escriurem *vigilar*. Sovint, però, sí que haurem de recollir aquestes formes i aleshores ens haurem de plantejar com les representem. Principalment, hi ha dues maneres de fer-ho, totes dues amb avantatges i inconvenients: transcriure-les en AFI o adaptar-hi l'ortografia —escrivint, per exemple, *vegiar*. L'AFI permet reflectir tot tipus de solucions no estàndards, però alenteix el procés de transcripció i dificulta la lectura del text, sobretot si es combina amb altres símbols; mentre que l'adaptació ortogràfica no sol entorpir gaire la lectura, però com a contrapartida exigeix establir des de bon principi uns criteris clars per reproduir amb

un mínim de coherència les formes no estàndards que trobem. Així, davant una forma verbal com *va portar-nos*, pronunciada [••••••••••], haurem de decidir com la transcrivim (cf. *va porta(r)-nse*, *va portà'nse*, *va portà'ns-e*) per representar-la sempre de la mateixa manera i després poder recuperar-la mitjançant cerques automàtiques. Sobre aquesta qüestió, veg. § 4.4.14.

Tenint en compte tots aquests condicionants, el CCCUB ha desenvolupat un sistema de transcripció discursiva de nivell intermedi-estret que proporciona informació força detallada tant del contingut lingüístic com del context (formes no estàndards adaptades, encavalcaments, durada de les pauses, etc.) i que alhora procura garantir una lectura relativament àgil del text. Com veurem més endavant (veg. § 4), no és un model de transcripció discursiva dissenyat per a un tipus d'anàlisi específica, sinó més aviat per oferir als investigadors de diverses disciplines uns materials ja endreçats per estudiar la llengua oral.

### 3.3 OBJECTIUS DE LA TRANSCRIPCIÓ

Cal insistir que sempre adaptarem la transcripció als aspectes lingüístics i/o discursius que més ens interessa posar en relleu. No necessita la mateixa notació una transcripció enfocada a l'anàlisi sintàctica de les tematitzacions, posem per cas, que una transcripció pensada per estudiar els torns de parla: mentre que la primera no caldrà que inclogui pauses ni encavalcaments, la segona sí que haurà d'assenyalar on es produeixen els silencis i quant duren i delimitar amb precisió els fragments de parla que se superposen. Per tant, abans de posar-nos a transcriure, haurem de decidir quin tipus de convencions ens convenen més, sobretot per no destinar un esforç excessiu a tot el procés de tractament de les dades.

Primerament ens haurem de plantejar quanta informació haurà de contenir el text transcrit. Si ens interessa una transcripció ampla, només haurem de triar aquells símbols que

creguem imprescindibles per a l'anàlisi posterior, però si necessitem una transcripció estreta, haurem de seleccionar curosament la informació que fixarem en paper, tenint en compte que com més detallada és una transcripció més esforç i temps demana. A més, haurem de pensar si transcrivim íntegrament tot el fet de parla o només aquells fragments que considerem més representatius per a l'anàlisi. I també haurem de preveure si adaptarem les solucions no estàndards que puguem trobar. Si no les adaptem, seguirem les normes ortogràfiques convencionals, però si decidim fer-ho, caldrà establir uns criteris d'adaptació mínims. Una altra possibilitat és transcriure en grafia estàndard i opcionalment acompanyar el text amb transcripció fonètica —Ochs (1979), per exemple, proposa fer servir l'AFI per transcriure només els sons orals no verbals, especialment dels nens.

Com assenyalen Calsamiglia i Tusón (1999: 344), també hem de tenir present si l'estudi a què destinem la transcripció serà de caràcter quantitatiu o qualitatiu. Una anàlisi quantitativa comporta sospesar seriosament si admetem formes lingüístiques no estàndards com ara *vegiar*, *portà'nse*, *allavons*, etc., que sempre resulten més difícils de recuperar a partir de cercadors automàtics. En canvi, una aproximació qualitativa al text accepta més fàcilment les solucions no estàndards, però amb l'inconvenient que les cerques seran més lentes, perquè haurem de preveure totes les variants d'una mateixa forma (cf. *llavors*, *allavors*, *allavons*, *allavòrens*, *allavontes*, *llons*, etc.). En aquest sentit, les eines informàtiques que utilitzarem per explotar la transcripció també condicionaran els criteris amb què treballarem. Una transcripció ampla, amb una notació mínima, sempre serà més manejable informàticament que una transcripció estreta, en què l'acumulació de formes no estàndards i convencions tipogràfiques (parèntesis, claudàtors, codis, símbols, etc.) pot entorpir força el tractament automàtic del text.

I, finalment, ens haurem de preguntar fins a quin punt interpretarem la seqüència oral que traslladem al paper. La majoria d'autors (veg. Tusón 1995: 81, Green *et al.* 1997, Briz *et*

al. 2002: 58, n. 2) ja adverteixen que és del tot impossible obtenir una transcripció completament neutra que no estigui sotmesa al filtre subjectiu del transcriptor. Tard o d'hora toparem amb fragments conflictius en què haurèm d'optar per una solució o una altra i assumir un cert risc d'error. Això és el que passa, per exemple, en fragments confusos i amb molt soroll de fons en què ens sembla sentir dos o més enunciats possibles, o en casos d'homofonia com *para!/pare!* en català oriental i els pronoms *li* i *l'hi*, amb grafies que representen mots i estructures sintàctiques ben diferents però que es poden confondre en contextos comunicatius concrets.

Malgrat tot, sí que podem decidir si utilitzem certes convencions que impliquen una anàlisi del transcriptor, encara que sigui mínima. Els enunciats irònics, posem per cas: si convenim marcar-los, obtindrem una transcripció més completa, però també més subjectiva, perquè hi haurà moments en què costarà saber si les paraules de l'informant volen ser iròniques realment o si tan sols és una interpretació que en fem nosaltres com a transcriptors. El model de transcripció discursiva del CCCUB, dissenyat per proporcionar material d'estudi a investigadors de disciplines diverses, opta per un nivell d'interpretació mínim perquè un mateix text serveixi per a molts tipus de recerca. Per contra, una transcripció amb una finalitat més específica, com ara l'estudi de l'humor en la conversa col·loquial, admetrà més convencions ja encaminades cap a l'anàlisi posterior del text, per exemple, un codi especial per als enunciats irònics.

### **3.4 LA TRANSCRIPCIÓ DISCURSIVA**

La principal característica de la transcripció discursiva és que incorpora molts elements del context, entès com a "conjunt de circumstàncies que envolten l'enunciació o producció d'un text, tant de tipus sociocultural, situacional, lingüístic o cognitiu" (Payrató 2003: 286). Això inclou des d'aspectes prosòdics com ara els grups entonatius i les inflexions tonals fins a les pauses,



els encavalcaments, la gestualitat o el marc físic i cognitiu en què es desenvolupa el fet de parla. El resultat és un text acompanyat de molta informació contextual per ajudar el lector a ubicar aquest mateix text en la situació comunicativa en què van ser produïts els enunciats originals. Tot plegat converteix la transcripció discursiva en un tipus de transcripció força completa i de gran utilitat per a disciplines com la pragmàtica, l'anàlisi del discurs i la sociolingüística, en què el context adquireix una importància teòrica cabdal.

Tanmateix, la transcripció discursiva té l'inconvenient que sovint ha d'integrar informació de caràcter massa divers i difícil de combinar. La notació de la gestualitat és un dels casos més paradigmàtics en aquest sentit: incloure en una transcripció discursiva la codificació pròpia dels sistemes més habituals de transcripció gestual (veg. Duranti 1997, Poyatos 1994, 2002; McNeill 1992, 2005; i els webs McNeill Lab i Max Planck Institute for Psycholinguistics) representa complicar tant el text de la transcripció que normalment els gestos s'acaben descrivint, més que no pas transcrivint (veg. § 4.4.12). Un altre obstacle amb què sol topar la transcripció discursiva és, com ja hem apuntat, l'adaptació de solucions lingüístiques no estàndards. Alguns sistemes de transcripció resolen el problema emprant les formes corresponents en llengua estàndard, sobretot si són models de transcripció pensats per a cerques automàtiques extenses (veg. § 3.5). Però aquells sistemes que adapten l'ortografia a les formes no convencionals es veuen obligats a recórrer a tota mena de recursos (parèntesis, subratllat, transcripció fonètica, etc.) que en excés poden carregar innecessàriament el text.

La llengua col·loquial és, de bon tros, la varietat lingüística que planteja més dificultats a qualsevol model de transcripció discursiva. Com assenyala Payrató (1995: 46), els motius són diversos, però la majoria es poden atribuir a l'espontaneïtat i la informalitat que caracteritzen aquesta varietat: hi ha més d'un canal obert alhora (el lingüístic i el gestual); l'elocució acostuma a ser ràpida; l'entonació presenta molts matisos i resulta difícil de representar; les vacil·lacions, els falsos inicis i els encavalcaments són constants; sovint es produeixen converses

paral·leles; es recorre contínuament al discurs exofòric, etc. Per això diversos autors aconsellen que a l'hora de recollir la informació contextual es parteixi d'un enfocament realista i es defugin plantejaments massa ambiciosos (veg. Payrató 1995: 52, Calsamiglia i Tusón 1999: 350). Primer, perquè el text resultant esdevé massa difícil de llegir i és poc adequat per a les cerques, i, segon, perquè exigeix un esforç en temps i recursos desproporcionat respecte a qualsevol tipus d'anàlisi. Val més determinar d'entrada els aspectes contextuais que més ens interessa destacar, sempre d'acord amb l'objectiu de l'estudi, i aplicar els criteris d'anotació que creguem més convenients.

### **3.5 MODELS DE TRANSCRIPCIÓ DISCURSIVA**

Com s'ha apuntat en més d'una ocasió (Payrató 1995, Calsamiglia i Tusón 1999), no disposem de cap sistema de transcripció discursiva universal equiparable a l'AFI, per la qual cosa cada disciplina, o fins i tot cada grup d'investigadors, ha tendit a transcriure la llengua oral a partir d'una metodologia pròpia i molt enfocada cap a un tipus d'anàlisi molt concreta. En part, això es deu al fet que els models proposats fins ara probablement són massa ambiciosos en voler compaginar valors tan desitjables com la neutralitat i la funcionalitat amb procediments que enriqueixen el text amb molta informació, un equilibri que, com hem vist en els apartats precedents, no és gens fàcil d'aconseguir.

A banda dels sistemes tradicionals de transcripció de la dialectologia (veg. Veny 1995, Veny i Pons 1998, Viaplana i Perea 2003) i la fonètica (veg. Bonet *et al.* 1997), que no solen recollir els aspectes contextuais de la interacció, els precedents més immediats de la transcripció discursiva tal com l'entenem actualment cal situar-los en el desenvolupament, als anys 60 i 70, de l'etnometodologia i l'anàlisi de la conversa, amb el treball pioner de Sacks *et al.* (1974) (veg. també Atkinson i Heritage 1984, Button i Lee 1987), i l'etnografia de la comunicació (veg. Gumperz i Berenz 1993), que van crear models adoptats

posteriorment per disciplines més pròpiament lingüístiques com són la pragmàtica i la sociolingüística (veg. Ochs 1979, Tannen 1984).

Tanmateix, no és fins a finals dels anys 70 i sobretot principis dels anys 90 (veg. Du Bois 1991) que apareixen els primers treballs que es plantegen la transcripció discursiva des d'una perspectiva més metodològica. Ochs (1979), que presenta un model dissenyat per a la transcripció del discurs infantil, segueix les convencions emprades per Sacks *et al.* (1974) en el camp de l'anàlisi de la conversa, però fa una primera proposta de caràcter generalista que serveixi per a diferents tipus d'anàlisi. Posa un èmfasi especial en el format de la transcripció, perquè considera que condiona molt la lectura del text (per exemple, recomana l'ús de columnes per a la transcripció de discurs infantil), i subratlla que el text no ha de contenir massa informació, perquè, si no, es fa massa difícil de seguir. Uns anys més tard Du Bois (1991: 77) estableix cinc principis bàsics perquè un sistema de transcripció discursiva resulti mínimament eficaç:

1. *Definició de categories.* Les categories que selecciona la transcripció han de ser rellevants per a l'anàlisi posterior i discriminables, és a dir, han de referir-se clarament a elements distints per evitar ambigüitats.

2. *Accessibilitat.* Les convencions que s'utilitzin han de resultar familiars al lector o fàcils d'aprendre per no entorpir la lectura del text.

3. *Robustesa.* El suport informàtic que utilitzem ha de facilitar la divulgació del text transcrit i alhora evitar que el document pateixi modificacions (per exemple, que es desgavellin els tabuladors i els salts de línia). Actualment aquests problemes es resolen, en part, amb el format PDF.

4. *Economia.* Les convencions del text han de reduir al màxim l'esforç que comporta tant l'elaboració com la lectura

d'una transcripció discursiva. Per això, entre altres estratègies, Du Bois recomana usar convencions curtes per a fenòmens molt freqüents, evitar comentaris llargs del transcriptor i utilitzar grafies fàcils d'identificar si s'insereixen dins un mot.

5. *Adaptabilitat*. El sistema de transcripció ha de poder oferir diferents nivells de descripció per acomodar-se a tot tipus d'anàlisis. Per tant, caldrà preveure, com a mínim, una versió ampla, una d'intermèdia i una d'estreta per al mateix sistema. Com veurem en l'apartat següent, aquests cinc principis constitueixen la base del model de transcripció discursiva proposat per Du Bois *et al.* (1991, 1993), que alhora és el que ha adoptat, amb variacions, el CCCUB.

O'Connell i Kowal (1994) són crítics amb les directrius exposades per Du Bois (1991) i altres autors (veg. també Romero, O'Connell i Kowal 2002). Consideren que, en general, no s'han tractat a fons els problemes metodològics que implica la transcripció discursiva i remarquen que tota proposta ha de buscar fonamentalment la utilitat per a la investigació més que no pas facilitar la lectura del text. Per això no creuen que calgui un sistema de transcripció discursiva universal, sinó que el model que s'utilitzi, sigui quin sigui, s'adapti al màxim a l'especificitat de cada recerca. Tot i així, també recomanen seguir alguns criteris bàsics per garantir una transcripció mínimament sòlida; entre d'altres:

a) Mantenir l'ús convencional de les grafies que s'utilitzin —incloent-hi els signes de puntuació— per evitar ambigüitats.

b) Assegurar la integritat dels mots, és a dir, evitar que continguin convencions gràfiques que n'alterin la forma i dificultin les cerques automàtiques.

c) Descriure, més que no pas transcriure, els fenòmens verbals no vocals, com ara el riure.

d) Garantir mesuraments exactes, sobretot de les pauses.

e) Transcriure únicament aquells elements rellevants per a l'anàlisi.

A la pràctica, però, els models que s'han estès més són els que tenen en compte els objectius potencials de la transcripció i ofereixen versions més amples o més estretes que s'adapten als interessos particulars de cada disciplina, cosa que permet que una mateixa transcripció sigui útil a un ventall més ampli d'investigadors (veg. Payrató 1995: 57).

És en aquest context, i gràcies a l'eclosió de la lingüística de corpus als anys 90, que han anat apareixent diversos models de transcripció més o menys exitosos. Un dels sistemes estàndards més estesos és el Text Encoding Initiative (TEI) (veg. Text Encoding Initiative Consortium), que va sorgir als Estats Units a finals dels anys 80 amb la voluntat de dotar els investigadors i sobretot les indústries de la llengua d'una eina comuna per marcar i tractar informàticament els materials dels corpus (vegeu-ne una mostra a Briz *et al.* 2002: 59-63). Aquest model, que està dissenyat per facilitar les cerques i l'intercanvi de dades a nivell internacional, s'ha fet servir en la transcripció dels textos orals de corpus molt extensos, com ara el British National Corpus i l'International Corpus of English, i també ha inspirat sistemes de notació més específics com el del Network of European Reference Corpora, desenvolupat per al corpus Cobuild (veg. Llisterrí 1999). Tanmateix, com assenyalen Briz *et al.* (2002: 59) i Albelda Marco (2005), el sistema TEI té l'inconvenient que és força interpretatiu, no facilita gaire la lectura i està pensat sobretot per a textos escrits o com a màxim per a fets de parla formals enregistrats en contextos molt controlats, com ara entrevistes dels mitjans de comunicació, i, per tant, no resulta prou adequat per transcriure interaccions més informals i complexes com és una conversa col·loquial, sobretot si es necessita una transcripció discursiva estreta.

Actualment, en el marc del projecte inicialment europeu i ara internacional EAGLES/ISLE (veg. International Standards for Language Engineering), s'està desenvolupant el format de notació estàndard CES/XCES (veg. XCES), que intenta resoldre els problemes que planteja la proposta TEI. Alhora el Grup de Llengua Oral del projecte EAGLES ha fet importants recomanacions sobre la constitució i aplicació de sistemes

estàndards per a la notació i el marcatge de llengua oral (veg. Gibbon *et al.* 1997) i el Grup d'Interacció Natural i Multimodalitat de l'ISLE hi ha fet aportacions recents. Fins ara, però, el projecte europeu MATE/NITE és el que s'ha centrat més en la llengua oral i ha desenvolupat eines dissenyades específicament per treballar amb dades audiovisuals (veg. Univesirsitat d'Edinburg). En anglès, els corpus més extensos que incorporen informació discursiva i prosòdica fent servir sistemes de notació i marcatge estàndards són el Lancaster/IBM Spoken English Corpus, el British National Corpus i sobretot el London-Lund Corpus of Spoken English. També s'han emprat formats estàndards en la transcripció i marcatge del material oral integrat al Corpus de Referència del Español Actual. Per a més informació sobre sistemes de notació estàndards, veg. Lehmberg i Wörner (2008) i el web de Joaquim Llisterra.

Un sistema de transcripció molt estès, aplicat tant a l'anglès com a altres llengües, és el model Chat, vinculat al projecte Childes (veg. Childes), iniciat en el camp de la psicolingüística per a la transcripció i intercanvi de dades sobre llenguatge infantil. L'èxit d'aquest sistema ha fet que s'utilitzi en molts altres corpus, com ara el corpus d'italià oral Lablita i especialment el projecte C-Oral-Rom, que ha donat com a resultat un corpus oral que inclou mostres de diverses llengües romàniques (veg. Universitat de Florència). Igualment ha tingut molt bona acollida el sistema Hiat (veg. Ehlich 1993), que ha esdevingut el model de transcripció adoptat per l'eina informàtica Exmaralda (veg. Universitat d'Hamburg), d'ús també molt general.

Cal esmentar també altres sistemes de transcripció menys estesos. Per exemple, els corpus d'àmbit francòfon utilitzen sovint el model desenvolupat per Blanche-Benveniste i Jeanjean (1987) i el grup GARS (veg. Bilger 2002), com ara el Corpus de Référence du Français Parlé, i des d'una perspectiva més sociolingüística s'ha fet la proposta de Thibault i Vicent (1990) (vegeu-ne una mostra a Payrató 1995: 51) i Vicent *et al.* (1995), aplicada sobretot al francès quebequès. En castellà destaca el sistema de transcripció discursiva del grup

Valesco, de la Universitat de València i la Universitat d'Alacant. El seu precedent més immediat és el model de Sacks *et al.* (1974) i té com a objectiu principal obtenir una transcripció semiestreta sense perjudicar gaire la lectura del text aprofitant al màxim les convencions de l'ortografia. Com el sistema desenvolupat per Du Bois *et al.* (1991, 1993), adopta com a unitat bàsica el grup entonatiu i assenyala les inflexions tonals; a més, adapta ortogràficament les formes no estàndards i pot oferir versions més amples o més estretes d'un mateix text (veg. Briz 1996, Briz *et al.* 2002, Hidalgo i Sanmartín 2005: 26). Finalment, vegeu també el sistema de transcripció ampla que utilitzen Tusón (1995) i Calsamiglia *et al.* (1997) per analitzar una conversa col·loquial i un debat televisiu respectivament.

### **3.6 LA TRANSCRIPCIÓ DISCURSIVA DE LA UNIVERSITAT DE CALIFÒRNIA SANTA BARBARA**

En el marc del Santa Barbara Corpus of Spoken American English, del departament de lingüística de la Universitat de Califòrnia Santa Barbara (veg. Santa Barbara Corpus of Spoken American English), l'equip d'investigació de John Du Bois, Stephan Schuetze-Coburn, Danae Paolino i Susanna Cumming, entre altres autors, ha desenvolupat des dels anys 90 un sistema de transcripció discursiva relativament fàcil d'aprendre i aplicar que alhora permet representar el discurs oral amb un format de lectura força àgil. Aquest model, que té com a unitat bàsica el grup entonatiu, es fonamenta en els cinc principis de Du Bois (1991) que hem esmentat en l'apartat anterior, encaminats essencialment a proporcionar transcripcions de nivell intermedi i estret amb una notació que no resulti difícil d'interpretar (veg. Universitat de Califòrnia Santa Barbara).

Un avantatge important d'aquest sistema de transcripció discursiva és que presenta la flexibilitat necessària per adaptar-se a molts tipus d'anàlisis. Pot servir tant per a una transcripció ampla que inclogui simplement el text i alguns aspectes vocals

no verbals (per exemple, el riure i les inhalacions) com per a una transcripció intermèdia (amb anotació d'encavalcaments, durada de les pauses, etc.) o una transcripció més estreta amb informació prosòdica detallada. Òbviament, com tot sistema de transcripció, també té inconvenients. Probablement, el més destacat és que demana una certa pràctica abans de saber utilitzar-lo amb fluïdesa, ja que es tracta d'un model basat en la identificació i caracterització prosòdica de grups tonals, no sempre fàcils de delimitar. És, doncs, un sistema força interpretatiu pel que fa a la segmentació del discurs oral, cosa que, a més, obliga a sotmetre la transcripció a una revisió en profunditat abans de donar-la per definitiva.

Malgrat tot, els múltiples avantatges que comporta (adaptabilitat, lectura àgil, etc.) van fer decidir els membres del CCCUB a adoptar-lo com a model de transcripció discursiva des de l'inici del corpus a principis dels anys 90, fent-hi les adaptacions convenients. En la resta del manual expliquem amb més detall el funcionament d'aquest sistema de transcripció del discurs oral. Recordem al lector que en el CD-ROM adjunt hi trobarà exercicis i les solucions corresponents per familiaritzar-se amb les convencions d'aquest model.



## 4. LA TRANSCRIPCIÓ DISCURSIVA DEL CCCUB

### 4.1 DESCRIPCIÓ GENERAL

Seguint la proposta de Du Bois *et al.* (1991, 1993) per a l'anglès, adaptada al català per Payrató (1995), el sistema de transcripció discursiva del CCCUB vol ser, essencialment, una eina pràctica per a l'estudi de la llengua oral des de disciplines ben heterogènies (l'etnografia de la comunicació, l'anàlisi del discurs, la sociolingüística, la pragmàtica, etc.) però amb un interès comú pel context de producció de l'enunciat. Per això busca un equilibri realista entre, d'una banda, una descripció del context relativament completa i, d'una altra, unes convencions clares i prou flexibles perquè es puguin adaptar a un ampli ventall d'anàlisis. Amb aquest objectiu, el sistema preveu tres modalitats de representació que es poden acomodar amb facilitat a la idiosincràsia de cada estudi. A continuació descrivim breument cadascuna d'aquestes modalitats, que aquí presentem per separat per motius expositius, però que més aviat s'han de concebre com a tres versions d'un mateix model que caldrà ajustar a l'objectiu de cada recerca.

a) *Transcripció ampla.* Nivell de representació que prescindeix de força convencions relatives a l'entonació, així com de codis que, si no són rellevants per a l'anàlisi, poden entorpir la lectura del text. No sol adaptar ortogràficament les formes lingüístiques no estàndards i simplifica la notació dels aspectes vocals i gestuals. Això la converteix en una modalitat de transcripció molt apropiada, per exemple, per a l'anàlisi semàntica o ideològica del discurs (veg. § 4.5.1).

b) *Transcripció intermèdia.* Aquesta modalitat recull informació prosòdica —incloent-hi allargaments, codis entonatius, etc.—, encavalcaments i pauses; adapta

ortogràficament les solucions no estàndards i descriu amb més detall els aspectes contextuais rellevants. Com veurem més endavant, és un nivell de transcripció apte tant per a l'estudi de fenòmens discursius i conversacionals (per exemple, la gestió dels torns de parla) com per a la recerca més pròpiament lingüística (veg. § 4.4).

c) *Transcripció estreta*. Aquesta versió utilitza totes les convencions de la transcripció de nivell intermedi però incorpora una notació prosòdica més exhaustiva, cosa que la fa especialment útil per a estudis que han de tenir molt en compte aspectes entonatius (veg. § 4.5.2).

En aquest manual descriurem principalment una transcripció de nivell intermedi-estret, que és la que fa servir el CCCUB per defecte. Considerem que és la modalitat més neutra de totes tres —amb el benentès que tota transcripció és necessàriament interpretativa i selectiva (veg. § 3.3)— i més fàcil d'adaptar a diferents tipus d'anàlisis. Com veurem al llarg d'aquest capítol, és un tipus de transcripció que conté força informació discursiva i contextual i que alhora aprofita certs elements de la notació prosòdica de la transcripció estreta (per exemple, la durada de les pauses), però sense entrar en la minuciositat pròpia d'aquesta darrera modalitat. També és la versió que permet un tractament informàtic més àgil del text sense prescindir de massa informació, per bé que en alguns casos caldrà tractar les dades abans de fer algunes cerques, sobretot per recuperar solucions lingüístiques no estàndards.

Totes tres modalitats tenen com a unitat bàsica el grup tonal, que és de naturalesa prosòdica (veg. § 4.4.1). En les versions intermèdia i estreta, a més, s'anoten amb símbols específics les inflexions melòdiques al final de cada grup entonatiu, i en la transcripció estreta també s'assenyalen els accents primaris i secundaris i els grups tonals menors, entre altres elements. A diferència de la proposta de Du Bois *et al.* (1991, 1993), però, el sistema de transcripció discursiva del CCCUB prescindeix d'aquells aspectes que exigeixen al transcriptor una interpretació semàntica o pragmàtica més

aprofundida del text. Ens referim, per exemple, a la notació funcional dels grups entonatius (veg. § 4.5.2) i a la codificació de citacions i fragments parentètics, elements no sempre ben definits que requereixen una transcripció molt més analítica i especialitzada. És per això que aquí reservarem les convencions de caràcter més interpretatiu per a la modalitat estreta.

En els apartats següents presentem més detalladament el sistema de transcripció discursiva del CCCUB. Primer tractem qüestions relacionades amb el format del text (veg. § 4.2) i la descripció de la situació comunicativa (veg. § 4.3); en segon lloc, desglossem les convencions de la modalitat intermèdia-estreta (veg. § 4.4) i a continuació, i per aquest ordre, les de la modalitat ampla (veg. § 4.5.1), intermèdia (veg. § 4.5.1) i estreta (veg. § 5.4.2). Per acabar, al capítol 5 donem alguns consells generals sobre transcripció, revisió i edició del text.

## **4.2 TRACTAMENT PREVI DE LA TRANSCRIPCIÓ**

Tot model de transcripció discursiva ha de plantejar-se des de bon començament com disposarà dins la pàgina el text de la transcripció i tota la informació complementària que l'acompanya. Cal preveure, per exemple, si el text tindrà un format vertical o apaïsat, si es justificarà al marge dret o a l'esquerre o si es distribuirà en columnes. Ochs (1979) assenyala que aquestes qüestions, lluny de ser banals, poden influir decisivament tant en la lectura com en la comprensió del text.

Ja hem apuntat més amunt que el sistema de transcripció discursiva del CCCUB adopta com a unitat bàsica el grup tonal, de caràcter prosòdic, i no unitats pròpiament textuais com són la frase o el paràgraf. Per reflectir l'ordre temporal amb què es va desenvolupant el discurs, els grups tonals se situen en columna, un sota l'altre, de manera que el text resultant afavoreix una lectura vertical —semblant a la d'un diari— més que no pas horitzontal. Dues columnes més acompanyen la dels grups

entonatius, totes dues situades a l'esquerra: una, la més externa, per enumerar els grups tonals; i l'altra, al mig, per indicar el parlant que intervé en cada moment. Vegeu com a mostra la figura 4.1, que reproduïx un fragment del COR (EMP6).

499	CRA:	Escola d'Idiomes_	
500		digui'm\	
501	C05:	(. 0.14) hola\	<b>Grups tonals i informació complementària</b>
502		bon dia\	
503	CRA:	[bon dia\]	
504	C05:	[mira\] ((se sent música de fons fins al final de la interacció))	
505		(. 0.21) trucava per fer una pregunta\	
506	CRA:	digue'm\	
507	C05:	el cursos de català teniu\	
508	CRA:	(INH)	
509		això hauries de trucar al departament de català eh\	
510	C05:	a ver- --	
511		a veure\	
512		em dones el telèfon si us plau\	
513	CRA:	sì\	
514		un momentet\	
515	C05:	(. 0.99) sì\ ((la música se sent ara amb més intensitat))	
516	CRA:	(. 0.45) et passo una trucada eh\ ((ho diu a una persona del servei d'assessorament lingüístic que treballa a la mateixa sala))	
517	SAL:	vale\	
518		(CRA marca el número de telèfon)	
519		(se senten els senyals de la trucada)	
520	SAL:	sì\	
521		((es talla la comunicació perquè ja han passat la trucada))	

**Inicials dels parlants**

**Numeració de grups tonals**

Figura 4.1 Distribució del text en la transcripció discursiva del CCCUB

La numeració és contínua i, per tant, no es reinicia en cada pàgina ni en cadascuna de les diverses interaccions que pot

incloure una mateixa transcripció. El fragment de la figura 4.1, per exemple, és un extracte d'una transcripció més llarga que comprèn diverses converses telefòniques: la numeració no torna a començar en cadascuna d'aquestes converses, sinó que es manté des del principi de tot fins a l'últim grup tonal de la darrera interacció, per evitar confusions a l'hora de localitzar un fragment o citar-lo. Convé observar que els números no es refereixen a les línies de la pàgina, sinó als grups tonals i/o les informacions complementàries que s'hi afegeixen, per la qual cosa podem trobar grups tonals llargs que ocupen més d'una línia però només tenen un número associat (veg. figura 4.1: 504, 516, 521).

No cal repetir les inicials dels parlants davant de cada grup tonal, sinó que només apareixeran un sol cop al principi de cada intervenció, així aconseguim que la transcripció resulti menys densa visualment. Una altra opció és repetir-les davant de tots els grups tonals en un color diferent per a cada informant; d'aquesta manera ajudem a identificar els torns de parla, sobretot si la interacció compta amb molts participants (vegeu-ne una mostra en les transcripcions de conversa col·loquial en el CD-ROM annex a Payrató i Alturo 2002).

Aquí aconsellem separar les columnes mitjançant tabuladors, perquè els espais manuals no garanteixen una alineació correcta del text. Si visualitzem les marques no imprimibles en el processador de textos que utilitzem, l'aspecte de la transcripció serà com el de (1), en què només hi ha espais entre mots, i entre mots i alguns símbols del sistema de transcripció. Tipogràficament, cada grup tonal correspon a un paràgraf i, per tant, acabarà amb un salt de paràgraf.

- (1) 499→CRA: →Escola.d'Ildiomes\_¶  
500 →diguim¶  
501→C05: →(.0.14)·hola¶  
502 →bon·dia¶  
503→CRA: →[bon·dia]¶  
504→C05: →[mira\]·((se·sent·música·de·fons·fins·al·final·de  
→la·interacció))¶  
505 →(.0.21)·trucava·per·fer·una·pregunta¶

506→CRA: →digue'm\¶

507→C05: →el-cursos-de-català·teniu\¶

És important evitar els espais i els tabuladors innecessaris, perquè poden interferir en les cerques, siguin automàtiques o manuals. Imaginem que en la transcripció de la figura 4.1 busquem tots els grups tonals de C05 que comencen per pausa —representem les pauses amb parèntesis i punts a l'interior segons la durada, veg. § 4.4.6— a partir de l'ordre de cerca de (2a). Si hi ha un espai entre el tabulador i el parèntesi, com a (2b), el programa no detectarà totes les ocurrences del text que busquem.

(2a) C05:→(.

(2b) 501→C05:→(. 0.14) hola\¶

Només deixarem espais en els contextos que especifiquem a continuació (pel que fa a les convencions de transcripció que esmentem, veg. § 4.4):

a) Entre pauses i text: *(. 0.45)·text*.

b) Entre claudàtors, que indiquen encavalcament, i text extern: *[text]·text*; però si el claudàtor se situa enmig d'un mot, no deixarem espai, cf. *[te]xt*.

c) Entre claus de codi i text extern: *{text}·text*; però si la clau se situa enmig d'un mot, no deixarem espai, cf. *{te}xt*.

d) Entre el parèntesi de tancament de codi i text: *{(L2)·text}*; entre codis, però, no deixarem espai, cf. *{(AC)(L2)·text}*.

e) Entre @, símbol de riulla, i text: *text·@@·text*.

f) Entre els parèntesis dels comentaris del transcriptor, simples o dobles, i text: *((a ANA))·text\((sorolls de fons))*.

g) Entre les ics dels fragments intel·ligibles i el text: *xX(0.80)Xx·text*.

Recomanem revisar regularment la transcripció per eliminar espais, tabuladors i altres elements tipogràfics sobers i comprovar-ho un altre cop al final de tot el procés de

transcripció. Per a altres convencions tipogràfiques, veg. § 4.4.16.

### **4.3 PRESENTACIÓ DEL CONTEXT COMUNICATIU**

Com hem apuntat a § 2.1, una tasca obligada abans de començar a transcriure és compilar i ordenar tota la informació relativa al context comunicatiu que resulti útil tant per al transcriptor com per al lector. No cal dir que no existeix un sistema únic per organitzar i presentar aquestes dades, perquè, en general, dependrà dels objectius de la transcripció i del volum d'informació disponible. Tot i així, en el cas que s'estigui elaborant un corpus mínimament extens, serà convenient crear una base de dades que contingui tota la informació que hàgim recollit de cada interacció verbal. Abans que res, però, establim un sistema per etiquetar els textos que formaran part del corpus, s'hagin de transcriure o no. Aquest sistema ha de ser prou clar perquè puguem trobar ràpidament l'enregistrament o la transcripció que ens interessa. N'és una bona mostra la notació del CAP (veg. Payrató i Fitó 2008: annex 2), en què les majúscules i les xifres que formen el nom de cada transcripció permeten identificar fàcilment qualsevol enregistrament del corpus. Per exemple, l'etiqueta C1TN01CC correspon a una entrevista feta en català (C1) i amb un procés d'elicitació experimental (T) i una finalitat narrativa (N) a la informant 01, que té el català com a llengua familiar (CC).

D'altra banda, la informació que contindrà la base de dades en relació amb cadascun dels enregistraments pot ser molt variada. El COC, el COR i el COS compten amb dues bases de dades diferents: una per a la informació sobre les interaccions verbals enregistrades i una altra per a les dades dels participants en cadascuna d'aquestes interaccions. La primera conté, entre d'altres, la data i el lloc d'enregistrament; el nombre de parlants i les inicials amb què apareixen en la transcripció; la qualitat del so de la gravació; si els informants sabien que eren enregistrats o no, o si l'enregistrament prové

dels mitjans de comunicació; la durada de l'enregistrament i el nombre de paraules de la transcripció, i les característiques comunicatives del fet de parla (to formal/informal; mode planificat/espontani; camp genèric/específic; tenor interactiu, instructiu, argumentatiu, etc.; veg. Payrató 1988: 51). La segona base de dades inclou data i lloc de naixement dels informants, sexe, professió, llengua d'origen familiar, llengua de domini actual, etc. (vegeu-ne una mostra a Payrató i Alturo 2002: 20).

De cara al lector de les transcripcions, caldrà seleccionar la informació contextual més rellevant i buscar la manera més adequada de presentar-la. En aquest cas, cada subcorpus del CCCUB opta per un procediment específic d'acord amb els seus objectius particulars:

1. *COC i COR*. Aquests dos subcorpus organitzen la informació contextual a partir del model d'anàlisi del context comunicatiu conegut amb l'acrònim anglès de Speaking, desenvolupat per Dell Hymes i John J. Gumperz en el marc de l'etnografia de la comunicació (veg. Boix i Vila 1998, Payrató 2003: 60). Aquest mètode permet aproximar-se al context en què té lloc la interacció verbal mitjançant la descripció dels vuit aspectes comunicatius següents, cadascun dels quals correspon a una lletra de l'acrònim:

a) *Situació*. Descripció del moment i l'espai físic en què es produeix el fet de parla, incloent-hi tota la informació rellevant que hi estigui relacionada, per exemple, si els participants es trobem amb periodicitat i si ho fan al mateix lloc.

b) *Participants*. Nombre d'interlocutors que intervenen en la interacció —encara que no ho facin d'una manera activa, com ara el públic d'una obra de teatre— i tota la informació disponible de cadascun d'ells que ajudi a contextualitzar la transcripció (edat, sexe, professió, tipus de relació entre ells, etc.), garantint-ne sempre l'anonimat (veg. § 2.2).

c) *Finalitats ('ends' en anglès)*. El propòsit pel qual es desenvolupa la interacció, que pot tenir un fi interactiu (per



exemple, un sopar d'amics en un restaurant), explicatiu (una visita guiada en un museu), argumentatiu (un debat per televisió), etc.

d) *Actes*. Les principals fases en què s'estructura la interacció. Per exemple, una conversa prototípica té una obertura, un bloc central i un tancament (veg. Tusón 1995: 49).

e) *Clau* ('key' en anglès). El to que adopta la interacció, és a dir, si és de caràcter formal, informal, irònic (per exemple, un pregó de carnestoltes), etc.

f) *Instruments*. Llengües i varietats geogràfiques, socials i funcionals que es fan servir al llarg de la interacció, a més dels canals no orals (llengua escrita, llengua de signes, etc.) amb què es complementin, com ara els apunts i els gràfics que empra un conferenciant.

g) *Normes*. Les convencions socials que guien el desenvolupament de la interacció. Per exemple, habitualment no s'espera que els assistents a una conferència interrompin el ponent per fer-li preguntes, sinó que n'hi facin un cop ha acabat l'exposició, després dels aplaudiments. En una transcripció neutra, però, no caldrà analitzar les normes de la interacció, perquè això correspondrà als estudis que se'n facin posteriorment des de disciplines com ara l'etnografia de la comunicació i l'anàlisi de la conversa; vegeu-ne exemples a Tusón (1995) per a la conversa col·loquial i Calsamiglia *et al.* (1997) per a un debat televisiu.

h) *Gènere*. Fet de parla prototípic (conte, conversa col·loquial, roda de premsa, etc.) a què pot ser assignada la interacció. Com en el cas anterior, i de cara a una transcripció neutra, tampoc no caldrà estudiar en profunditat el text, perquè ja s'encarregaran de fer-ho disciplines com la lingüística textual i l'anàlisi del discurs, entre d'altres; vegeu-ne una mostra de caràcter comparatiu entre llengua oral i llengua escrita a Castellà (2004). Tot i així, sí que és convenient tenir en compte certes variables funcionals (text monològic/dialògic, públic/privat, literari/no literari, etc.) tant per classificar les mostres del corpus (vegeu com a exemple els criteris adoptats al COR a Alturo *et al.* 2004) com per presentar-les als usuaris d'una manera ordenada.

A partir de la informació obtinguda mitjançant el model etnogràfic Speaking, cada transcripció del COC i el COR ofereix al lector un text de presentació per ajudar-lo a situar-se en el context en què es desenvolupa el fet de parla. A (3a) donem com a exemple una descripció contextual del COC i a (3b) una del COR. En el cas del COC, a Payrató i Alturo (2002) el text es complementa amb una taula que conté la data de la interacció i informació sobre els parlants (edat, sexe, lloc de naixement, professió, grup social, estudis, etc.), que molt sovint no saben que són enregistrats.

- (3a) [Conversa] enregistrada un diumenge de febrer al domicili particular de dos dels participants, a Esparreguera (Baix Llobregat). Hi intervenen dues parelles joves (entre 23 i 29 anys), amigues, que descansen després d'haver assistit a un dinar de grup en una altra localitat. La conversa és interrompuda un moment per MEL (59 anys), que saluda les dues parelles i se'n va. El tema que predomina és el dels casaments, que és tractat a partir de qüestions com ara la tria del lloc de la celebració, el casament de persones conegudes, el tipus de casament (civil o per l'església), l'actitud dels sacerdots, el vestuari, el comiat de solter, l'habitatge de la parella, el desenvolupament de la cerimònia i del convit, etc. (Payrató i Alturo 2002: conversa 02)
- (3b) Judici que té lloc en una sala del Jutjat Número 9 de 1a Instància i Instrucció de Barcelona, el dia 20 de febrer de 1998. Es tracta d'un judici de faltes, en què es vol determinar si l'acusat va incórrer en una falta de furt el mes de setembre de 1997. Hi intervenen el jutge (JUG); dos testimonis (TE1 i TE2) i la fiscal (FSL). JUG és un home de 45 anys, catalanoparlant i nascut a Barcelona. FSL és una dona d'uns seixanta anys catalanoparlant. L'acte es desenvolupa en una sala força petita, on hi ha la taula del jutge i la de l'agent judicial i, perpendiculars a aquestes dues, les taules de la fiscal i l'advocat defensor. Davant de les taules hi ha el banc on seuen els acusats i els bancs per al públic que vol assistir als judicis. En aquest cas, però, no hi ha assistents. Com a símbols, al fons de la sala, al darrere del jutge, hi ha la bandera catalana i l'espanyola i una fotografia del rei Joan Carles I. Els testimonis van entrant a la sala a mesura que l'agent judicial, a instàncies del jutge, els crida. L'acusat no ha comparegut al judici i, per tant, l'advocat defensor, que sí que hi assisteix, no pot intervenir. (Alturo *et al.* 2004: CIV1, judici de primera instància)

Com es pot veure, cada text de presentació s'adapta al tipus de transcripció discursiva que acompanya. La transcripció del COC pertany a un corpus homogeni pel que fa al gènere (conversa col·loquial) i, per tant, el text posa més èmfasi a explicar al lector els temes que es tractaran i els fragments que poden resultar difícils d'interpretar —com ara la salutació de MEL, que interromp la conversa— que no pas a descriure el fet de parla en si. En canvi, la presentació de (3b), corresponent a una mostra inclosa en un corpus de gèneres discursius diversos, comprèn la data i el lloc de l'esdeveniment comunicatiu i algunes dades personals dels interlocutors, però sobretot fa una descripció tant de l'escena física com del desenvolupament de l'acte.

A més a més, en tots dos corpus s'adjunten taules que indiquen la durada de cada enregistrament i el nombre de mots de la transcripció respectiva, una informació fonamental per a les anàlisis estadístiques que se'n derivin.

2. COS. En aquest cas l'objecte d'estudi no són els gèneres discursius, sinó la varietat de català pròpia d'un sector de la població molt determinat (homes i dones de classes treballadores socialitzats a l'àrea metropolitana de Barcelona, veg. Boix-Fuster *et al.* 2006: 22). Per tant, la presentació inicial no se centra en el gènere discursiu que s'utilitza per obtenir les dades (l'entrevista semidirigida de final obert), sinó en la informació sociolingüística més rellevant (els capitals) de cada informant. Vegeu (4) com a mostra.

(4) **CAPITALS**

**Econòmic: baix**

ESF va començar a treballar als setze anys en una parada de llegums, a la plaça d'Esplugues. Més tard, va entrar a treballar en una fàbrica, on realitzava una feina totalment mecànica. Als 23 anys, quan es va casar, va anar a viure en un pis de lloguer de l'Hospitalet. Va deixar de treballar a fora de casa per dedicar-se a les feines domèstiques. L'any 1991, el seu marit es va quedar sense feina durant sis mesos; ara treballa de mecànic. Actualment, encara viuen a l'Hospitalet.

**Cultural: baix**

Va anar a l'escola pública fins a vuitè d'EGB, però no es va treure el graduat escolar (de sisè a vuitè va suspendre tots els cursos). Després

de l'escola, va entrar en una acadèmia del barri on feien secretariat, però tampoc va acabar aquests estudis. Fa tres anys es va apuntar a l'escola d'adults del barri, on va fer un curs de confecció i després un d'alfabetització. Actualment està fent el pregraduat. Li agrada anar al cinema, i de la televisió li agraden els concursos i les pel·lícules que no són violentes; també es mira els documentals (de qualsevol tipus), i assegura que no li agrada gens Tele5, prefereix TV3. No llegeix revistes, i de la ràdio li agraden els programes musicals en què no es parla.

**Social: mitjà-baix**

Marxen alguns caps de setmana (llargs) a la muntanya, amb la família. Van a passar les vacances en un càmping. Es relacionen amb els amics de sempre, que són una parella de Cornellà. ESF també es relaciona amb gent de l'escola d'adults.

**Simbòlic: baix**

El seu capital simbòlic és baix.

**Altres observacions**

L'entrevistada parla el català amb tota la seva família, excepte amb el seu marit, que, tot i ser barceloní, parla el castellà.

(Boix-Fuster *et al.* 2006: entrevista 05, ESF)

En un quadre adjunt també s'indica la durada dels enregistraments i el nombre de paraules que conté cada transcripció.

3. *CAP*. Aquest corpus consta d'un total de 36 entrevistes semidirigides amb què s'obtenen dades per a l'anàlisi de la multimodalitat i les tipologies textuais en català, castellà i anglès. L'entrevista, doncs, tampoc no és l'objecte d'estudi, sinó el mètode d'elicitació (veg. Payrató i Fitó 2008: 11) i, per tant, no caldrà descriure el context comunicatiu abans de cada transcripció. N'hi haurà prou explicant com s'utilitza aquesta metodologia en un document a part que també inclogui altres informacions d'interès (per exemple, les dades dels informants), que al *CAP* es proporcionen en una taula prèvia a les mostres.

En comptes d'un text de presentació, una altra possibilitat és incloure a l'inici de cada transcripció una fitxa amb un format ja establert que compili la informació més destacada del fet de parla (data i lloc d'enregistrament, informació sobre els parlants,

contingut de la transcripció, etc.). En són un bon exemple les fitxes que encapçalen les transcripcions discursives del Corpus Oral de Registres de la Universitat de Lleida (veg. Universitat de Lleida). Aquest corpus, que consta de 40 enregistraments, ha estat dissenyat per estudiar la llengua oral a partir dels factors comunicatius de mode, camp i tenor. Per aquest motiu cada transcripció va precedida d'una etiqueta identificativa i una fitxa que descriu el fet de parla a partir de les tres variables funcionals i altres dades (durada de l'enregistrament i argument del text) i observacions. Els estàndards de notació i marcatge internacionals també preveuen formats ja establerts per descriure el context comunicatiu (veg. Lehmborg i Wörner 2008).

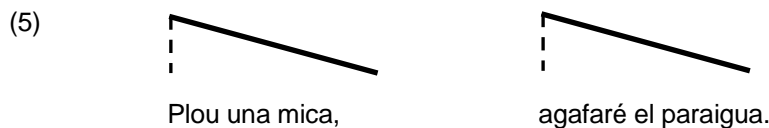
#### **4.4 CONVENCIIONS DE LA TRANSCRIPCIÓ DISCURSIVA DEL CCCUB**

Un cop hem vist la distribució del text dins la pàgina i alguns procediments per descriure el context comunicatiu en què s'emmarca el fet de parla, desglossem a continuació les convencions del sistema de transcripció discursiva del CCCUB corresponents a un nivell intermedi-estret, que és la modalitat de transcripció que adoptarem per defecte. Per a les convencions pròpies d'una transcripció més ampla o més estreta, veg. § 4.5.

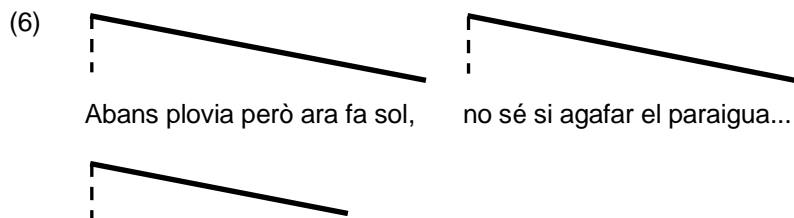
##### **4.4.1 Unitat de transcripció: el grup tonal**

D'acord amb la proposta de Du Bois *et al.* (1991, 1993), el sistema de transcripció discursiva del CCCUB adopta com a unitat bàsica el grup tonal o grup entonatiu —també anomenat *grup d'entonació*, *frase tonal* i *sintagma d'entonació*, veg. Prieto (2002a: 37)—, que aquí definim com el fragment de parla enunciat sota una única corba entonativa i delimitat pels trets prosòdics següents: compleció de la corba melòdica, ressitució, pausa, anacrusi i allargament. A continuació els descrivim per separat:

a) *Corba d'entonació completa.* La corba d'entonació “uneix els diferents tons amb què s'ha pronunciat cada segment. Els punts més alts de la corba melòdica es corresponen amb els accents principals de les seqüències discursives” (Payrató 1988: 91). En un grup tonal prototípic la corba entonativa presenta un inici i una terminació ben definits que donen unitat i coherència a tot el fragment. Ho representem esquemàticament a (5), en què cadascuna de les oracions queda delimitada per l'augment inicial del to i una caiguda gradual, de manera que la part final resulta més greu que la part inicial. Cadascuna d'aquestes oracions correspondrà a un grup tonal independent.



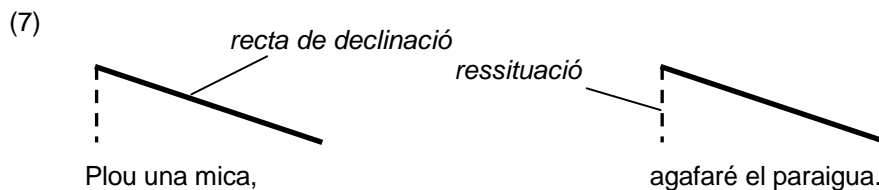
Els límits entonatius de la corba melòdica no han de coincidir necessàriament amb els límits sintàctics dels mots que conté. Els dos grups tonals de (5) són grups entonatius sintàcticament travats perquè en tots dos casos l'inici i la terminació de la corba entonativa es corresponen amb el començament i el final d'una oració simple, però és freqüent que no hi hagi tanta correspondència entre contorn prosòdic i sintaxi. Podem trobar fàcilment un grup tonal que inclogui més d'una oració, com en el primer grup entonatiu de (6), o a la inversa, una mateixa estructura sintàctica fragmentada en dos o més grups tonals, com mostren els altres dos segments de (6), en què les dues oracions subordinades dependents del verb *saber* s'enuncien en unitats prosòdiques diferents.



o si deixar-lo a casa.

Més endavant veurem que el final de la corba melòdica presenta una gran varietat d'inflexions que aquí denominarem *seqüències tonals terminals* i que donen compleció al grup entonatiu i alhora li adjudiquen determinats valors pragmàtics (veg. § 4.4.2).

b) *Ressituació o reajustament del to ('reset' en anglès)*. La corba melòdica, independentment dels pics accentuals, tendeix a declinar a mesura que avança en el temps, per la qual cosa la part terminal presenta un to més baix que la part inicial. Aquest descens entonatiu, o *recta de declinació* (veg. (7)), comporta que a l'inici del grup tonal següent el to es torni a col·locar en el nivell habitual de la tessitura del parlant, fenomen que es coneix com a *ressituació*, veg. (7).



La ressitució és un dels indicadors més clars per detectar si comença un grup tonal nou. El reajustament implica, a més d'una pujada del to, una intensitat més forta, i això es tradueix en un cop de veu més o menys perceptible que pot ajudar molt el transcriptor a delimitar aquestes unitats prosòdiques.

c) *Pausa*. L'absència de producció verbal també és un dels indicis més útils per identificar grups tonals, perquè la represa de la parla implica generalment un grup entonatiu nou. Tanmateix, la manca de pausa no significa necessàriament que estiguem davant d'un únic grup tonal. Per exemple, en un ritme d'elocució ràpid és molt habitual que s'encadenin dos o més grups entonatius sense cap pausa entremig, o que resulti mínima, per la qual cosa haurem de recórrer constantment a altres fenòmens prosòdics (el contorn melòdic i sobretot la ressitució) per detectar els límits entonatius de la cadena fònica. A més, també és possible que en un grup tonal hi hagi

una micropausa interna, com a (8) —en què representem la pausa mitjançant (.)—, per bé que l'experiència ens demostra que es tracta d'un cas força esporàdic.

- (8) 481 PPP: (.) s'ha de treure aquella pasta del (.) de la nevera  
(COC: 07, adaptat)

*d) Anacrusi.* Juntament amb la ressituació del to i una intensitat més elevada, l'inici de grup tonal acostuma a presentar una pronúncia accelerada de les primeres síl·labes. L'anacrusi, però, no és tan determinant com els indicadors anteriors i tindrà un caràcter secundari.

*e) Allargament.* En la part terminal del grup entonatiu sol produir-se un alentiment del ritme d'elocució i, en conseqüència, un allargament de les síl·labes finals. Com l'anacrusi, no és un indicador determinant i per això també tindrà un paper secundari en la delimitació de grups tonals. En la transcripció del CCCUB representarem els allargaments amb dos punts (:) darrere la vocal o la consonant allargada, sense espai entremig i d'acord amb la graduació següent: allargament breu (:), mitjà (::) i llarg (:::). Si dura més, repetirem tantes vegades com calgui la mateixa grafia (veg. també § 4.4.10, 4.4.17).

Un grup entonatiu prototípic presentarà tots cinc trets prosòdics alhora: corba melòdica completa, ressituació, pausa, anacrusi i allargament final. Ara bé, a la pràctica no sempre és fàcil distingir-los tots cinc i això fa que al llarg de la transcripció puguin sorgir dubtes a l'hora de segmentar la cadena fònica en grups tonals.

El criteri general del sistema de transcripció discursiva del CCCUB és segmentar un fragment de parla en dos grups entonatius quan hi identifiquem almenys dos dels indicadors principals, això és, compleció de corba melòdica, ressituació i pausa. Per tant, dividirem en dos grups tonals un fragment de parla que contingui, posem per cas, una pausa seguida de ressituació, però el mantindrem com una sola unitat entonativa si presenta només una pausa molt breu a l'interior. Altres vegades trobem una lleugera ressituació del to combinada amb un



allargament vocàlic a la meitat del grup tonal, tot i que és un fenomen més propi de final de grup entonatiu: en aquest cas probablement no tindrem dues unitats prosòdiques, sinó una de sola, perquè l'allargament no és un factor tan determinant com la compleció entonativa, la ressituació o una pausa.

En alguns casos pot ser que fins i tot costi detectar com a mínim dos dels indicadors principals. Per això aconsellem que sempre s'analitzin tots cinc indicis d'una manera relativa, atenent les peculiaritats de cada idiolecte. Així, per exemple, podem trobar un fragment de parla amb una caiguda de la corba melòdica seguida d'una ressituació poc definida: si la caiguda tonal és considerable en relació amb la tessitura del parlant, molt probablement haurem de segmentar el fragment en dues unitats prosòdiques. Reprendrem la qüestió més endavant en relació amb les seqüències entonatives terminals (veg. § 4.4.2).

De moment cal observar que el resultat final de la transcripció consistirà en una concatenació de peces entonatives delimitades fonamentalment per trets prosòdics que es van succeint una darrere l'altra. Així, el fragment oral de (9), procedent d'un missatge deixat en un contestador automàtic (veg. COR: PRIV2, 155-165), s'organitzarà en blocs entonatius de durada variable que haurem delimitat gràcies a les pistes prosòdiques que ens proporciona.

- (9) hola sóc Toni de Viatges XX sisplau quan vulgueu ja podeu passar per aquí a recollir la documentació o sigui a partir de avui a les cinc de la tarda ja podeu passar quan vulgueu gràcies i adéu

Si a (9) hi afegim les pauses curtes, (.); les pauses mitjanes, representades per (..), i els allargaments, com a (10a), de seguida veiem que el mateix fragment de parla es divideix, com a mínim, en cinc unitats entonatives que provisionalment delimitem mitjançant ||. Si, a més, hi marquem les síl·labes en què es produeix una ressituació del to —aquí subratllades per raons expositives—, com a (10b), comprovem que el mateix fragment encara es pot tallar en peces entonatives més petites, cadascuna de les quals constitueix un únic grup tonal.

(10a) hola sóc Toni de Viatges XX (.) || sisplau quan vulgueu ja podeu passar per aquí: || a recollir la documentació o sigui a partir de avui a les cinc de la tarda (..) || ja podeu passar quan vulgueu gràcies (..) || i adéu

(10b) hola || sóc Toni || de Viatges XX (.) || sisplau || quan vulgueu ja podeu passar per aquí: || a recollir la documentació || o sigui || a partir de avui a les cinc de la tarda (..) || ja podeu passar quan vulgueu || gràcies (..) || i adéu

De fet, com hem vist a § 4.2, la transcripció tindrà un format vertical i, per tant, la disposició definitiva de (9) serà molt semblant a la de (11). Aquest format afavoreix un text poc atapeït que agilita la lectura, per bé que té l'inconvenient que representa crear documents molt més extensos que els sistemes de transcripció de format horitzontal.

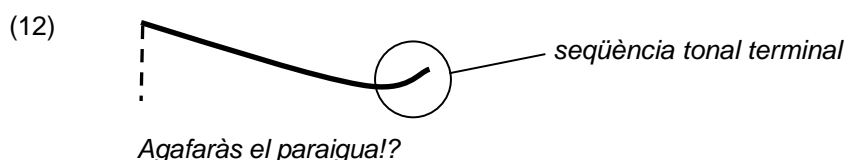
(11) 155 TOY: hola  
156 sóc Toni  
157 de Viatges XX  
158 (.) sisplau  
159 quan vulgueu ja podeu passar per aquí:  
160 a recollir la documentació  
161 o sigui  
162 a partir de avui a les cinc de la tarda  
163 (..) ja podeu passar quan vulgueu  
164 gràcies  
165 (..) i adéu  
(COR: PRIV2, adaptat)

Convé insistir que el resultat d'incorporar informació entonativa a la transcripció és una distribució del discurs basada en criteris principalment prosòdics, més que no pas sintàctics o semàntics. Per tant, els grups tonals no hauran de coincidir obligatòriament amb estructures sintàctiques com són el sintagma i l'oració ni tenir sentit complet en termes semàntics (veg. Payrató 1988: 106). Si bé és cert que generalment trobem una certa correspondència entre límits entonatius i límits sintàctics, semàntics i fins i tot pragmàtics, com en el cas dels vocatius (Prieto 2002a: 40), res no impedeix, per exemple, que un grup tonal contingui només una part d'un sintagma o que, al contrari, ajunti dues oracions sintàcticament independents. A

(11: 162-163) podem observar que el sintagma preposicional *a partir de avui a les cinc de la tarda* i el sintagma verbal *ja podeu passar quan vulgueu* no formen una unitat entonativa, malgrat que sintàcticament estan del tot relacionats, o que les dues fórmules *gràcies* i *adéu* no constitueixen un sol grup tonal tot i estar lligades per la conjunció coordinant.

#### 4.4.2 Seqüències tonals terminals i truncaments

Un cop definides les unitats prosòdiques amb què treballarem, reprenem el concepte apuntat més amunt de *seqüència tonal terminal*, que aquí emprarem per referir-nos a les inflexions de la corba melòdica al final del grup entonatiu. Prieto (2002a: 51) assenyala que l'última síl·laba tònica del grup tonal funciona com a àncora per a moviments melòdics que, d'una banda, permeten delimitar-lo i, d'una altra, li confereixen determinats valors pragmàtics (declaració, interrogació, continuïtat, etc.). Aquestes inflexions del to es produeixen sobretot en les síl·labes àtones situades entre l'última síl·laba tònica i el final de la seqüència, tot i que, si l'últim mot és agut, es concentren en la darrera síl·laba accentuada. Així, per exemple, en una pregunta eco com *Agafaràs el paraigua!?*, que pragmàticament pot expressar sorpresa i desacord alhora (cf. *Agafaràs el paraigua!?, però si fa sol!*), tindrem un moviment típicament ascendent en la terminació de la corba entonativa, tal com representem esquemàticament a (12). La inflexió del to que té lloc al final d'aquest enunciat no només ens permet saber que la corba melòdica és completa —i, doncs, que tot el fragment constitueix un únic grup tonal—, sinó que també ens aporta informació pragmàtica sobre com l'hem d'interpretar en el context de producció.



De seqüències tonals terminals n'hi ha de diverses menes (veg. Prieto 2002a i 2002b per a una descripció detallada), però de cara a una transcripció intermèdia-estreta només en distingirem tres tipus bàsics, cadascun representat per un símbol diferent: ascens (/), manteniment (⎵) i descens (\). Anotarem el símbol al final del grup tonal sense deixar espai amb el text precedent.


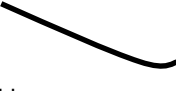

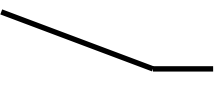



La distribució dels tres tipus d'inflexió melòdica pot resultar extremadament variada i dependre molt del context en què es pronuncia l'enunciat, però cada inflexió compta amb uns contextos d'aparició prototípics, com ara els que recollim a continuació:

a) *Ascens*. La terminació tonal ascendent és molt característica de les oracions interrogatives totals (*portes paraigua?*) i parcials (*qui porta el paraigua?*) més enfàtiques, les interrogatives imperatives (*vols deixar el paraigua!?*) i les preguntes eco com la de (12).

b) *Manteniment*. Les inflexions mantingudes són habituals en oracions declaratives situades enmig d'una seqüència discursiva (*aleshores vaig agafar el paraigua i...*), oracions dubitatives (*no sé si agafar el paraigua...*), precés (*va, home, deixa'm el paraigua...*), enumeracions (*agafar el paraigua, l'impermeable, les botes...*) i oracions deixades en suspens (*si plou i et deixes el paraigua...*).

c) *Descens*. El moviment descendent és més propi d'oracions declaratives, sobretot si se situen al final d'una seqüència discursiva (*...i és per tot això que he agafat el paraigua*) o tenen un matís categòric o imperatiu (*no discuteixis, agafarem el paraigua*); exclamatives (*ostres, m'he deixat el paraigua!*); interrogatives imperatives i preguntes eco amb un to molt categòric (*a veure, qui ha agafat el paraigua!?*; *com que no vol agafar el paraigua!?*); imperatives (*agafa el paraigua!*) i oracions amb algun element focalitzat (*el PARAIGUA!*, *m'he deixat*).

En realitat, cadascun dels tres símbols entonatius que fem servir en la transcripció de nivell intermedi-estret engloba més d'un tipus de seqüència tonal terminal. Com mostra esquemàticament la taula 4.1, el símbol d'ascens (/) pot equivaler a un ascens prototípic o un lleuger ascens; el de manteniment (⎵), a un manteniment prototípic, un manteniment-ascens o un manteniment-descens; i el símbol de descens (\), a un descens prototípic o un lleuger descens. Per reflectir amb més precisió aquests moviments tonals, hi ha l'opció d'utilitzar els símbols corresponents a una transcripció estreta (veg. § 4.5.2), recollits a la columna de la dreta de la taula 4.1.

Seqüència tonal terminal	Símbol en una transcripció intermèdia-estreta	Símbol en una transcripció estreta
 Ascens	/	↑
 Lleuger ascens	/	//
 Manteniment-ascens	-	/
 Manteniment	-	→
 Manteniment-descens	-	\
 Lleuger descens	\	//
 Descens	\	↓

Taula 4.1 Representació de les seqüències tonals terminals

En cas de dubte, optarem pel símbol més neutre o esperable en cada context, això és, l'ascens en els enunciats interrogatius, el manteniment en enunciats declaratius enmig d'una seqüència discursiva i el descens en els declaratius, sobretot si apareixen al final d'una enumeració o una explicació, i els imperatius (veg. § 4.4.3.3).

Si reprenem l'exemple de (11) i hi incorporem les seqüències tonals terminals corresponents, obtindrem la

transcripció de (13), en què no hi ha inflexions ascendents, però sí moviments descendents, sobretot al final de seqüència discursiva —vegeu els grups tonals (g. t.) 160, 163, 165—, i moviments de manteniment, que expressen continuïtat del discurs.

- (13) 155 TOY: hola\_  
156 sóc Toni\_  
157 de Viatges XX\_  
158 (.) sisplau\  
159 quan vulgueu ja podeu passar per aquí:\_  
160 a recollir la documentació\  
161 o sigui\_  
162 a partir de avui a les cinc de la tarda\_  
163 (..) ja podeu passar quan vulgueu\  
164 gràcies\_  
165 (..) i adéu\  
(COR: PRIV2, adaptat)

On apareixen exemples prototípics de moviments ascendents és en el missatge recollit a (14), en què la finalitat clarament exhortativa del discurs —per demanar a l'interlocutor que acudeixi a una cita— afavoreix que s'utilitzin seqüències tonals terminals ascendents en el primer vocatiu (g. t. 944) i en les dues preguntes dels grups tonals 946 i 947.

- (14) 944 AAN: Xavi/  
945 (...) Xavi\  
946 que estàs aquí/  
947 (...) que véns/  
948 (..) fa un quart d'hora que t'espero al carrer\  
949 (...) fins després\  
(COR: PRIV2, adaptat)

Sobre l'anotació de les inflexions tonals terminals, veg. també § 4.4.3.3, 4.4.17.

D'altra banda, al llarg del discurs oral també trobarem grups entonatius truncats, és a dir, grups tonals que el parlant deixa inacabats per raons diverses (embarbussament, autocorrecció, interrupció de l'interlocutor o d'algun element

extern, etc.). Indicarem el truncament amb el símbol --, sense deixar espai amb el text precedent. Si, a més, s'interromp un mot, representarem amb el símbol - -- el truncament de mot i de grup tonal alhora. En l'exemple de (15), extret d'una discussió per telèfon entre dos adolescents, apareixen dos truncaments de grup tonal (g. t. 542, 545) i dos truncaments de mot i grup tonal (g. t. 539, 544).

- (15) 536 GRR: però a veure Adrià\  
 537 vale\  
 538 vale que et va tocar els pebrots\  
 539 però c- --  
 540 però és que estic fins als ous\_  
 541 (INH)  
 542 de--  
 543 de fer el {(F)(L2) pardo\_}  
 544 i que la Gina se- --  
 545 sempre em tingui a a--  
 546 (INH)  
 547 a la seva disponibilitat\_  
 (COR: PRIV3, adaptat)

Per acabar, cal tenir present que delimitar grups entonatius i atribuir-los seqüències tonals terminals demana una certa pràctica, per la qual cosa és habitual que al principi hi hagi alguns dubtes i vacil·lacions. En l'apartat següent fem algunes recomanacions sobre aspectes prosòdics conflictius per facilitar la tasca, sempre laboriosa, de transcriure a partir de criteris prosòdics.

#### 4.4.3 Aspectes prosòdics conflictius

Un model de transcripció discursiva basat en unitats prosòdiques permet representar el discurs oral d'una manera més acurada que no pas un sistema que opta per unitats típicament textuales com són la frase i el paràgraf. Com a contrapartida, però, implica assumir la dificultat de traslladar a un suport escrit elements inherentment orals com els grups



entonatius i les inflexions tonals. A continuació tractem alguns aspectes relacionats amb la prosòdia que poden resultar problemàtics per al transcriptor no familiaritzat amb el sistema de transcripció discursiva del CCCUB i assenyallem els criteris que ha seguit aquest model de transcripció en cada cas.

#### 4.4.3.1 *Llargada dels grups tonals*

Una de les primeres qüestions que sorgeixen quan segmentem el discurs oral en grups entonatius és quants mots pot contenir cadascuna d'aquestes unitats prosòdiques. En realitat, no hi ha cap mínim ni cap màxim establerts, perquè la llargada dels grups tonals dependrà de molts factors comunicatius (velocitat d'elocució, idiolecte del parlant, etc.). Per a l'anglès s'ha descrit una mitjana de cinc mots en grups tonals d'una durada de dos segons (veg. Chafe 1979: 164). Pel que fa al català, no disposem d'estudis específics sobre aquesta qüestió, però les transcripcions discursives del CCCUB semblen indicar que la mitjana se situaria al voltant dels quatre mots per grup tonal en interaccions informals (per exemple, en la conversa 01 del COC la mitjana és de 4 mots/grup tonal; en l'entrevista a MMS del COS, de 4,1; i en la transcripció C1CN01CC del CAP, de 3,6) i entre quatre i cinc mots en interaccions més formals (per exemple, en la sessió parlamentària recollida al COR: POL2 la mitjana és de 4,5 mots/grup tonal). Òbviament, la qüestió demana una anàlisi més aprofundida i, per tant, la mitjana que aquí donem tan sols s'ha de considerar una estimació orientativa.

Malgrat tot, sembla que en general els grups tonals llargs (d'una línia sencera) són força rars, i moltes vegades s'expliquen per un ritme d'elocució ràpid. Compareu la transcripció de (16a), procedent del discurs que pronuncia en un míting polític una oradora experimentada i amb un ritme d'elocució habitualment molt ràpid, amb el fragment de (16b), extret de la presentació d'un llibre de poesia d'un parlant amb un ritme d'elocució molt més lent.

- (16a) 838 PPR: (..) fa vint anys que va morir i tenim un Parlament\_  
 839 que no pot legislar de re\  
 840 pràcticament\  
 841 (..) enviant parlamentaris\_  
 842 (..) però amb misèries competitives\  
 843 (..) fa vint anys que va morir i encara tenim totes les seves  
 administracions\  
 844 (..) fa vint anys que va morir\_  
 845 (..) i encara tenim {(L2) Guardia Civil y Policía Nacional} a  
 Catalunya\  
 846 (..) fa vint anys que va morir\_  
 (COR: POL1, adaptat)
- (16b) 298 JSJ: (..) bona tarda a tothom\_  
 299 (inhalació)  
 300 (..) e\_  
 301 moltes gràcies\_  
 302 (..) al\_  
 303 amic Castillo\_  
 304 n'estic\_  
 305 (..) estic content\_  
 306 (inhalació)  
 307 de ser aquí amb vosaltres\_  
 (COR: CULT4, adaptat)

Un idiolecte àgil, amb un ritme d'elocució ràpid com el de l'oradora de (16a), generalment produeix grups tonals llargs, mentre que un idiolecte amb més pauses, un ritme d'elocució lent i amb més omplidors de parla (veg. § 4.4.9.4), com el de l'informant de (16b), sol reflectir-se en una transcripció amb grups tonals més curts.

Hi ha altres factors que influeixen en l'extensió dels grups tonals. Habitualment una formalitat alta afavoreix grups entonatius relativament llargs, mentre que un discurs més informal comporta grups tonals més breus, tal com semblen apuntar les mitjanes que hem esmentat més amunt. També hi té a veure el grau d'espontaneïtat de la situació comunicativa: si el discurs ha estat preparat amb antelació, si s'utilitza un suport escrit o si es tracta d'un text recurrent, com ara l'explicació d'un guia turístic, és més probable que els grups tonals siguin força

llargs; com al fragment de (17a), extret d'un sermó —un discurs oral amb suport escrit. En canvi, en interaccions espontànies els grups tonals tendeixen a ser més curts; cf. (17b), procedent d'una conversa col·loquial.

(17a) 59 FMS: (...) i quantes persones han volgut aparentar\_  
60 (...) el que no tenen ni són\  
61 (...) i això és molt greu de vegades\  
62 (...) perquè hi ha persones que\_  
63 (.) aparenten ser bones\  
64 (...) tenen títols de bondat\  
65 (..) poden ser sacerdots\_  
66 poden ser jutges\_  
67 poden ser pares de família\_  
68 (..) poden ser persones que per estament social  
i professió\_  
69 (.) els cal ser bones persones\_  
70 (..) (inhalació)  
71 però es veu que a dins el cor\_  
72 (..) no hi tenen\_  
73 (..) el que cal\  
(COR: REL1, adaptat)

(17b) 470 HIL: (..) una de Olesa\_  
471 una modista\_  
472 que treballa--  
473 una\_  
474 que dissenya\_  
475 uns vestits\_ ((de núvia))  
476 (..) i treballa per una atra\_  
477 per una modista famosa no/  
478 que treballa\  
479 NAA: mhm\  
480 HIL: (..) i\_  
481 com que és d'Olesa\_  
482 la coneix\_  
483 i tal\_  
484 li va dir que si ella li li faria\_  
485 i va dir que sí\  
(COC: 02, adaptat)

Així doncs, a l'hora de segmentar el discurs oral en grups entonatius no només haurem de tenir en compte els criteris

pròpiament prosòdics descrits en els apartats anteriors, sinó també condicionants comunicatius com ara l'idiòlecte del parlant i les característiques funcionals de la interacció verbal (grau de formalitat, espontaneïtat del discurs, etc.), perquè ens ajudaran a preveure quin tipus de grup tonal predominarà al llarg de la transcripció.

Finalment, en aquells casos en què dubtem si hem de dividir un fragment de parla en dos o més grups tonals, recomanem que es tendeixi a segmentar. Ja hem assenyalat més amunt (veg. § 4.4.1) que pararem una atenció especial en la ressitució del to, que acostuma a ser l'indicador més clar, així com en les inflexions entonatives i els allargaments que tenen lloc al final de la corba melòdica. Tot i així, si encara dubtem, segmentarem el fragment en qüestió com a opció per defecte; veg. també § 4.4.17 per a algunes errades freqüents en la identificació de grups tonals.

#### 4.4.3.2 *Vocatus, connectors, marcadors discursius, interjeccions, incisos, tematitzacions, rematitzacions i mots cua*

Al llarg del discurs toparem amb elements lingüístics força heterogenis des d'un punt de vista formal (interjeccions, vocatus, mots cua, etc.) que a vegades no sabrem si constitueixen un grup tonal independent o si formen part de grups entonatius més grans. A continuació assenyalem els criteris que segueix el sistema de transcripció discursiva del CCCUB respecte a aquests elements sovint força problemàtics, per bé que no s'han de considerar regles d'aplicació estricta, sinó més aviat orientacions generals per al transcriptor poc avesat al sistema de transcripció que aquí exposem. Primer ens referirem als elements que habitualment transcriurem com a grups entonatius únics (vocatus, connectors, marcadors discursius, interjeccions i incisos) i més endavant als que tendirem a inserir en grups tonals més grans (tematitzacions, rematitzacions i mots cua).

a) *Vocatus*. Malgrat que la corba melòdica dels vocatius varia sensiblement segons la seva posició dins l'enunciat (veg. Prieto 2002b: 412), l'autonomia prosòdica d'aquests fàtics afavoreix que en el nostre model de transcripció discursiva els considerem preferiblement un grup tonal independent, sovint acabat amb un descens entonatiu, veg. (18a: 609, 18b: 816). Tanmateix, trobarem excepcions, perquè a vegades constitueixen grups tonals menors que només marcarem, si cal, en una transcripció estreta, cf. (18c).

(18a) 609 BEB: (...) Anna\  
 610 tu: has de prendre cafè descafeïnat no/  
 (COC: 06, adaptat)

(18b) 815 ANI: (...) truco a la Marta/  
 816 Beli\  
 817 BEB: (..) vale\  
 (COC: 06, adaptat)

(18c) 45 BEB: (...) Anna vols cafè\  
 (COC: 06, adaptat)

b) *Connectors*. D'acord amb Cuenca (2002, 2006), podem distingir dues menes de connectors: les conjuncions, que uneixen dos segments d'una mateixa oració i expressen relacions sintàctiques de coordinació (*i*, *o*, *però*) o subordinació (*que*, *si*), i els connectors parentètics, que relacionen semànticament dues oracions i presenten independència entonativa (*ara bé*, *així mateix*, *en tot cas*, *en canvi*, *és a dir*, etc.). Des del punt de vista de la transcripció discursiva, això es tradueix en contextos d'aparició força diferents: mentre que les conjuncions acostumen a formar part d'un grup tonal més gran —sobretot *o*, *que*, *si*; *i*, *però* mostren més variació—, els connectors parentètics generalment constitueixen un grup entonatiu únic. En donem alguns exemples en el fragment oral argumentatiu de (19):

(19) 1 I05: bueno\  
 2 do:ncs\_

3	jo estic a favor de:_	36	una vegada arribes
4	viure al--		allà penses_
5	bueno\	37	oh:\
6	a favor\	38	quina tranquil·litat\
7	a veure\	39	(..) jo prefereixo el
8	jo prefereixo viure en		ca- --
	el ca:mp_	40	e:l poble_
9	primer perquè::_	41	diguéssim\
10	no hi ha tant trànsit	42	mil vegades\
	com a Barcelona\	43	però després:s_
11	(.) perquè--	44	doncs_
12	fins i tot--	45	també hi han els
13	o sigui_		avantatges de la
14	l'altre dia van estar		ciutat\
	parlant d'això_	46	que és la
15	(..) i: m_		universita:t_
16	(..) i per exemple_	47	i hospita:ls_
17	a Barcelona mateix	48	i:_
	quan:_	49	(..) i:_
18	quan:_	50	bueno\
19	estàs en el metro_	51	(.) transport públic\
20	encara que no tinguis	52	per exemple\
	pressa_	53	que allà_
21	corres\	54	doncs_
22	(.) vas així més ràpid\	55	és lo que deia abans\
23	i això no ho aguanto\	56	(.) de l'estació_
	[...]	57	(.) fins a casa meva_
24	en canvi_	58	tinc ben bé cinc
25	vale\		quilòmetres\
26	és un:_	59	i com no vaig_
27	pal_	60	si no vaig en:_
28	doncs_	61	cotxe_
29	el haver de venir cada	62	que no en tinc\
	dia:_	63	(..) he d'anar en bus\
30	que é:s_	64	i ja són més
31	en total és una hora		despeses\
	i mitja_	65	(.) i bueno\
32	anar_	66	(..) però:_
33	i una hora i mitja	67	de preferir_
	tornar\	68	el poble\
34	cada dia\		(CAP: C1TA05CC, adaptat)
35	però bueno\		

Com podem veure, el transcriptor ha inclòs la majoria de conjuncions dins un grup tonal més gran (*perquè*, 9; *i*, 15, 16, 23,

33, 47, 59, 64, 65; *encara que*, 20; *que*, 30, 46, 53, 62; *però*, 43; *si*, 60) —excepte en els grups tonals 48 i 49 (*i*) i 66 (*però*)— i ha situat els connectors parentètics en un grup tonal independent (*doncs*, 2, 28, 44; *o sigui*, 13; *en canvi*, 24; *diguéssim*, 41; *per exemple*, 52), en algun cas combinats amb una conjunció (*primer perquè*, 9; *i per exemple*, 16).

c) *Marcadors discursius i interjeccions*. També solen formar grups tonals independents els anomenats *marcadors discursius*, que a diferència dels connectors tenen una funció més organitzativa (veg. González 2004) (a (19), *bueno*, 5, 50; a *veure*, 7), així com la majoria d'interjeccions pròpies i impròpies (a (19), *oh*, 37; *vale*, 25) i les combinacions entre aquests elements i conjuncions (*però bueno*, 35) o connectors parentètics (*doncs bé*, *doncs bueno*, etc.).

d) *Incisos*. Malgrat que els incisos mostren una tipologia formal molt variada, majoritàriament es caracteritzen per una autonomia prosòdica (veg. Prieto 2002b, Payà 2003) que reflectirem mitjançant grups tonals independents. A l'exemple de (20), extret d'una roda de premsa, tenim com a exemples l'adverbi oracional *lògicament* (g. t. 509) i l'oració *subratllo intentarem* (g. t. 511).

(20) 505 AMG: (.) un\_  
 506 (.) model\_  
 507 (.) aeroportuari\_  
 508 (..) que és\_  
 509 lògicament\_  
 510 (..) el que intentarem\_  
 511 (..) subratllo intentarem\_  
 512 (..) e::\_  
 513 negociar\_  
 (COR: POL4, adaptat)

e) *Tematitzacions i rematitzacions*. Els elements tematitzats adopten contorns prosòdics propis (amb ascens tonal si són a l'esquerra i descens si són a la dreta, veg. Prieto 2002b: 410); tanmateix, el fet que normalment constitueixin grups tonals menors (veg. § 4.5.2) afavorirà que els incloquem en grups entonatius més grans, tant si apareixen a l'esquerra, veg. (21a),

com a la dreta de l'enunciat, veg. (21b: 215). En alguns casos, però, sí que els correspondrà un grup tonal independent perquè hi pot haver una pausa relativament llarga entre el constituent desplaçat i la resta de l'enunciat, com a (21c).

(21a) 1051 MMM: am aquest senyor li han d'haver donat la la: jubilació  
anticipada\  
(COC: 07, adaptat)

(21b) 215 PPP: els jesuïtes també en fan de reformes\  
216 i grosses\  
(COC: 07, adaptat)

(21c) 466 TBB: (..) i quants anys s'hi va estar\_  
467 (.) a la fàbrica\_  
(COS: LLU, adaptat)

Aplicarem el mateix criteri respecte a les rematitzacions (veg. Vallduví 2002), com a (22a: 1185), i sovint caldrà marcar el constituent que rep el volum de veu més prominent amb el codi d'intensitat forta  $\{(F) \text{ text afectat}\}$  (veg. § 4.4.9.1), com es mostra a (22b: 330)

(22a) 1185 FAN: més comunicació hi ha d'haver\  
1186 (riure)  
1187 HIL: (.) sí\  
(COC: 02, adaptat)

(22b) 327 GRR: ara me la picarà ((la cadena))  
328 perquè l'Adrià sempre està pel terra\_  
329 (inhalació)  
330 i es  $\{(F) \text{ pica}\}$  la cadena\  
(COR: PRIV3, adaptat)

f) *Mots cua*. Anomenem *mots cua* totes aquelles interjeccions de caràcter marcadament fàtic (*eh!*, *eh?*, *oi?*, *no?*, *sí?*, etc.; veg. Cuenca 2002: 3225) que se situen al final de l'enunciat per emfasitzar la rellevància d'una informació, veg. (23a: 20); demanar confirmació (23b, 23c: 142) o simplement comprovar que el canal continua obert (23d: 5), entre altres funcions pragmàtiques. No sempre és clar si formen un grup



tonal independent o si pertanyen a un grup entonatiu més ampli. Tanmateix, el fet que es resil·labifiquin amb el mot anterior (cf. *és molt fort eh!* [••••••••••••••••]) i acostumin a formar grups tonals menors (veg. Prieto 2002b: 433) semblen arguments prou sòlids perquè tendim a inserir aquestes partícules en un grup tonal més gran, excepte en aquells casos en què com a mínim hi ha una pausa breu entre l'enunciat i el mot cua, cf. (23d: 4-5).

(23a) 20 ONI: jo en vull una mica més eh\ ((de sopa))  
 21 MAT: (..) és pasta pitufos\  
 (COC: 04, adaptat)

(23b) 1150 RRR: aquest és el que comprem no/ ((es refereix a un anunci de cereals que acaben de fer per televisió))  
 (COC: 08, adaptat)

(23c) 1462 PEI: va agafant bon color oi/ ((es refereix a un plat que cuinen en un programa de televisió))  
 1463 RRR: s:ì\  
 (COC: 08, adaptat)

(23d) 2 TOR: pues bueno\_  
 3 el pobre i el ric\_  
 4 es veia molt la diferència\_  
 5 (..) no/  
 (COS: TOR, adaptat)

El mateix criteri ens servirà per als mots cua que apareguin dins l'enunciat, com a (24a: 256). Ara bé, en el cas de *no?* és recomanable segmentar l'enunciat en dos o més grups entonatius si considerem que es pot produir ambigüitat en la interpretació del text. A (24b: 1691), per exemple, si no dividim el grup tonal, hi ha la possibilitat que el lector interpreti el *no* com un adverbí de negació (cf. *i els del Canàries molt bé no es porten*) i no pas com una interjecció.

(24a) 256 ALL: a tots s'ha de pagar eh per anar-hi\ ((als castells))  
 257 JON: hombre\  
 (COC: 10, adaptat)

(24b) 1690 PEI: és que encara pot guanyar el Barcelona\ ((la lliga de futbol))

1691 RRR: i els del Canàries molt bé no/  
1692 es porten/  
(COC: 10, adaptat)

#### 4.4.3.3 Anotació de les seqüències tonals terminals

Un altre aspecte que pot causar dificultats al transcriptor és l'anotació de les seqüències tonals terminals, perquè en alguns casos no resulta fàcil determinar quin tipus d'inflexió melòdica té lloc al final del grup entonatiu i com l'hem de representar. Els dubtes són especialment freqüents quan la seqüència tonal és lleugerament ascendent o lleugerament descendent i costa decidir si el símbol que emprarem ha de ser de manteniment, ascens o descens. Com hem apuntat més amunt (veg. § 4.4.2), recomanem que en cas de dubte s'opti per la solució més neutra o esperable segons si l'enunciat és interrogatiu, declaratiu o imperatiu.

Alhora algunes de les seqüències tonals terminals més conflictives estan vinculades als elements discursius que hem descrit en l'apartat anterior (vocatius, connectors, etc.). A continuació en reprenem alguns en relació amb les inflexions tonals al final de grup entonatiu i també analitzem la prosòdia de dos tipus de partícules discursives molt pròpies de la llengua oral, els retroalimentadors i els omplidors de parla.

a) *Vocatius*. Depenent de la funció pragmàtica que adquireixin, els vocatius presentaran una seqüència tonal ascendent, sobretot en enunciats interrogatius emfàtics, veg. (25a: 520); de manteniment, com en preguntes de caràcter més neutre com la de (25b: 1096-1097); o descendent, com ara en ordres (veg. 25c: 728) (veg. també Prieto 2002b: 459).

(25a) 520 ANI: (..) que marxés ja Anna/  
521 NIA: (.) no:\n(COC: 06, adaptat)  
(25b) 1096 BEB: {(F) Marina\_  
1097 però l'Anna on ha anat\

1098       (.) perquè té el biquini fet eh\  
(COC: 06, adaptat)

(25c) 726   CME: {(F) vingui:\  
      727        aqui\  
      728        mare\  
(COC: 07, adaptat)

*b) Connectors.* Tant les conjuncions com els connectors parentètics solen expressar continuïtat, per la qual cosa generalment aniran acompanyats d'una inflexió de manteniment. També preferirem el símbol de moviment tonal mantingut en aquells casos en què el connector adquireix un valor interjectiu, com ara el *però bueno* de (26: 110).

(26) 104   I06:   és una cosa que hi és\_  
      105        i costa d'acceptar\_  
      106        i\_  
      107        a veure\_  
      108        de vegades es fa molt difícil\_  
      109        (..) però\_  
      110        (..) però bueno\_  
              [...]  
      113        és lo que hi ha\  
(CAP: C1TE06CC, adaptat)

Cal tenir en compte, però, que alguns connectors poden situar-se al final de l'enunciat, com en el cas dels adversatius *però*, *per (ai)xò* (27a) i *i tot* (27b). Aquestes formes, a més d'inserir-se preferentment en un grup tonal més gran, incorporaran una seqüència tonal descendent.

(27a) 2213   RAS: l'Amparo Moreno hi té gràcia eh per (ai)xò\  
(COC: 03, adaptat)

(27b) 822   LIA:   i va pujar en helicòpter i tot\  
(COC: 09, adaptat)

*c) Interjeccions.* La independència pragmàtica i prosòdica de les interjeccions la tendirem a reflectir mitjançant una seqüència tonal terminal descendent, sobretot en contextos

marcadament exclamatiu, com ara el *rebec* de (28a: 159). Tot i així, és molt habitual que una mateixa interjecció presenti inflexions tonals diferents segons el context discursiu en què apareix i la funció pragmàtica que desenvolupa. N'és un exemple ben conegut la interjecció impròpia *home* (veg. Cuenca 2002: 3224, Cuenca i Torres Vilatarsana 2008), que podem trobar tant amb entonació final descendent —veg. (28b: 1599), en què expressa desacord— com amb un to mantingut —veg. (28c: 46), en què funciona més aviat com a omplidor de parla.

(28a) 159 GRR: (..) hòstia\  
160 però com se t'acudeix deixar-l'hi\  
(COR: PRIV3, adaptat)

(28b) 1597 PUY: a la Natàlia\_  
1598 no sé què collons li trobes Miquel\  
1599 home\  
(COC: 05, adaptat)

(28c) 45 ECC: no fas cap activita::t\_ ((el cap de setmana))  
46 I04: (..) home\_  
47 a mi m'agrada\_  
(CAP: C1CE04CC, adaptat)

Aquesta variabilitat prosòdica comporta que, a l'espera d'estudis més aprofundits sobre la qüestió (veg. Cuenca 2002: 3206), aquí no puguem establir correspondències sistemàtiques entre interjeccions, siguin pròpies o impròpies, i contorns entonatius concrets i, per tant, aconsellem seguir els criteris generals d'anotació prosòdica exposats a § 4.4.2.

De moment donem algunes orientacions sobre els patrons que solen preferir dos tipus d'interjeccions molt específiques: els retroalimentadors (*backchannels* en anglès), que comuniquen a l'interlocutor que el canal continua obert (veg. § 4.4.9.4), i els omplidors de parla, que permeten conservar el torn dins la conversa (veg. § 4.4.9.4). Els primers, constituïts per formes vocals no verbals (*aha*, *mhm*) i verbals (*ja*, *sí*), els marcarem per defecte amb entonació final descendent, com a (29a: 59), i els segons, que també comprenen formes vocals no verbals (*e*, *a*,

*m*, veg. (29b)) i formes pròpiament verbals (*llavors*, *doncs*, veg. (29c: 1)), els anotarem amb manteniment tonal.

(29a) 56	I06:	perquè ma mare	(29c) 1	I01:	(...) doncs:_
		no_	2		la immigració a
57		vull dir_			Europa és_
58		encara estava	3		(s'escura la gola)
		desmaia:da_	4		(..) és:_
59	ECC:	mhm\	5		un tema que:_
	(CAP:	C1CN06CC, adaptat)		(CAP:	C1TE1CC, adaptat)

(29b) 120 JCV: (.) e:\_  
 121 a nivell de: m:\_  
 122 m:\_  
 123 de: m:\_  
 124 (..) de cultura de\_  
 [...]  
 127 quines coses de la  
 Pirelli són les que:\_  
 128 (..) m:\_  
 129 les que més t'han:  
 sobtat\  
 (COR: EMP1, adaptat)

d) *Mots cua*. Acompanyarem de seqüència tonal ascendent els mots *cua* que demanen confirmació, incloent-hi la interjecció pròpia *oi?*, veg. (30a: 108), i les interjeccions impròpies *no?*, veg. (30b: 477); *saps?*, veg. (30c: 753), *sí?* i *d'acord?* (veg. Prieto 2002b: 436), així com totes les interjeccions que en general presentin un contorn melòdic interrogatiu, com en el cas d'*eh?* quan demana que es repeteixi l'enunciat precedent, com fa a (30d: 523), i *ah sí?*, que expressa sorpresa, veg. (30e: 946).

(30a) 107	NAA:	(..) ai\ 108 (..) sembla mentida oi/ (COC: 02, adaptat)	(COC: 02, adaptat)	(30c) 751	HER:	nosatres volíem la-- el xamfrà_ 752 753 saps/ (COS: HER)
(30b) 476	HIL:	(..) i treballa per una atra_ 477 per una modista famosa no/ (30d) 522	PPP:	està bé aquesta paraula\ (30e) 946	HER:	està bé aquesta paraula\ (30f) 947

523	CME: (.) eh/	(30e)	944	NAA: (..) moltes
524	PPP: (..) si està ben dita_			vegades sí\
525	això de lirona\		945	quan entren els nuvis s'aplaude:ix\
	(COC: 07, adaptat)		946	SUS: (..) ah sí::/\
				(COC: 02, adaptat)

En canvi, anotarem amb el símbol d'entonació descendent la interjecció *eh* quan no té un caràcter pròpiament confirmatori i subratlla més aviat la rellevància d'una informació, com a (31: 1098).

- (31) 1097 BEB: però l'Anna on ha anat\  
1098 (.) perquè té el biquini fet eh\  
(COC: 06, adaptat)

e) *Incisos*. Un dels trets més característics dels incisos és que presenten un descens tant del to com de la intensitat (veg. Prieto 2002b: 409, Payà 2003); per tant, hi adjuntarem el símbol de manteniment, com a (32a: 169-170), i en els casos més pronunciats el de caiguda, cf. (32b: 316).

(32a)	167	FTR: jo conec algun cas	(32b)	315	PEI: diuen que ahir_
		que:_		316	els entesos\ diuen que ahir
	168	TOR: (..) eh/		317	van jugar bé
	169	FTR: (.) un pagès_			(COC: 08, adaptat)
	170	per exemple_			
	171	que té:_			
	172	(..) molts diners_			
	173	molts diners_			
	174	i va amb una bicicleta_			
		(COS: TOR, adaptat)			

f) *Tematitzacions*. Si el constituent desplaçat a l'esquerra forma grup tonal propi, el grup entonatiu acabarà amb seqüència de manteniment, veg. (33a: 1390). Si està desplaçat a la dreta, presentarà descens tonal, sobretot en enunciats declaratius, com a (33b: 59) (veg. Prieto 2002b: 411).

- (33a) 1390 CME: i a l'Anna Mari\_

1391            els hi podem regalar\ ((les neules))  
(COC: 07, adaptat)

(33b) 58    RRR: és possible que tornin a venir\  
59            p(e)rò de moment no en necessiten d'aparcament\  
(COC: 08, adaptat)

#### 4.4.4 Unitat lingüística: la paraula

La unitat lingüística bàsica de la transcripció serà la paraula, que sempre que sigui possible haurà de conservar la seva forma íntegra per facilitar la lectura del text. Així, en casos com (34a: 161), en què ens veiem obligats a inserir dins el mot *alberg* un claudàtor numerat per representar un encavalcament (veg. § 4.4.7), optarem per la solució de (34a: 161) més que no pas la de (34b: 161), perquè altera menys la paraula afectada pel símbol d'encavalcament:

(34a) 160    ONI: segueixes amunt\_  
161            i trobaràs un atre al[berg juvenil\ 2]  
162    COC: [2 ah:\  
(COC: 04, adaptat)

(34b) 160    ONI: segueixes amunt\_  
161            i trobaràs un atre al[2berg juvenil\  
162    COC: [2 ah:\  
(COC: 04, adaptat)

Si una paraula queda truncada, ho indicarem mitjançant un guionet darrere l'últim so que s'hagi pronunciat, veg. (35a: 192). El mateix símbol ortogràfic ens servirà per assenyalar que el parlant s'embarbussa amb una paraula, encara que no quedi truncada, com en l'article *el* de (35b: 1690), o que completa un mot que al principi semblava que quedaria inacabat, com el *després* de (35c: 260) —en aquest últim cas deixarem espai darrere el guionet.

(35a) 192    COC: què ha fet la ma- la iaia al forn\_  
193            més bo:\  
194            una cuixeta/

(COC: 04)

(35b) 1689 LAU: (..) so- son pare el va\_  
1690 el- el va deixar re- --  
[...]  
1694 (.. 0.67) el va deixar retratat son pare\  
(COC: 01, adaptat)

(35c) 258 DDC: a part d'uns inèdits\_  
259 que també em va prometre que llegiria\_  
260 i des- prés\_  
261 farà un repàs de la seva\_ ((obra poètica))  
(COR: CULT4, adaptat)

Pel que fa a les convencions tipogràfiques que utilitzarem en el text de la transcripció (majúscules, transcripció de xifres, etc.), veg. § 4.4.16.

#### 4.4.5 Unitat discursiva: el torn de parla

Com hem vist a § 4.2, la transcripció discursiva del CCCUB s'estructura en grups entonatius disposats verticalment i distribuïts en torns de parla, que poden tenir una extensió molt variable (veg. Tusón 1995: 44). En una interacció monològica, la transcripció només constarà d'un sol torn de parla, mentre que en una interacció dialogada les intervencions seran múltiples i sovint s'encavalcaran les unes amb les altres. Situem les inicials dels parlants al començament de cada torn i no caldrà repetir-les davant de cadascun dels grups tonals que formen part de la mateixa intervenció, com a (36a); n'hi ha prou amb una sola vegada al principi de cada torn, cf. (36b).

(36a) 1908	IDA:	un moment\ ((sobre el dinar de Nadal))	1914	IDA:	iaia [Rose:r_]}
1909	IDA:	{{(F) el dia de Nadal_	1915	JJJ:	[sí home sí]
1910	IDA:	(..) o ho fem_	1916	JJJ:	això ja [{{(??) ho he dit}}
1911	IDA:	aquí casa meva_	1917	IDA:	[Carmen] Mi[quel]
1912	IDA:	(..) iaia Roser_	1918	JJJ:	[això ja] està clar ja\ 1919
1913	IDA:	home no:\	IDA:	m:_	1920
				IDA:	Meritxell o Eulàlia\ 1920



1921 IDA: i vosatres\ (COC: 09, adaptat)	1914 iaia [Rose:r_]}
(36b) 1908 IDA: un moment\ ((sobre el dinar de Nadal))	1915 JJJ: [sí home sÍ]
1909 {(F) el dia de Nadal_	1916 això ja [{{(??) ho he dit}}]
1910 (..) o ho fem_	1917 IDA: [Carmen] Mi[quel]
1911 aquí casa meva_	1918 JJJ: [això ja] està clar ja\
1912 (..) iaia Roser_	1919 IDA: m:_
1913 home no:\	1920 Meritxell o Eulàlia\ 1921 i vosatres\ (COC: 09, adaptat)

No hi ha cap límit en el nombre de lletres que triem per identificar els parlants, per bé que el sistema de transcripció discursiva del CCCUB normalment en fa servir tres. A més, s'intenta que la combinació de lletres s'adapti als patrons de sil·labificació del català i, per tant, es prefereix una inicial com IDA, com a (36), que no pas XZF o WJQ, per exemple. Cal recordar que les inicials dels parlants sempre aniran seguides de dos punts.

Per referir-nos a tots els participants en la interacció verbal o a tot un conjunt indeterminat, emprarem el símbol T??, com a (37a: 134), en què fa referència als assistents a un casament. El mateix símbol ens pot ser útil per indicar que tots els participants parlen alhora i resulta difícil entendre què diu cadascun, veg. (37b), en què les ics representen fragments de parla inintel·ligibles (veg. § 4.4.8) i les arroves, rialles (veg. § 4.4.9.2).

(37a) 130 CPP: posem-nos drets\ ((els assistents s'aixequen))	(37b) 695 SER: no me'n recordo\ 696 JJJ: (.) ah\ 697 (..)
131 amb l'Evangelí_	698 T??: @RR@
132 (..) el Senyo:r_	699 T??: xXXx
133 (.) sigui amb vosaltres\ 134 T??: (.) i amb el vostre esperit\ ((tothom se senya))	700 (...) 701 T??: xXXx 702 (..) 703 SER: vaig comprar això an allà_
(COR: REL2, adaptat)	(COC: 09, adaptat)

En aquells casos en què sabem què es diu però no identifiquem el parlant que intervé, atribuirem el torn de parla a ???, com en el diàleg de (38a) entre dos assistents a una subhasta de peix. I si creiem que probablement es tracta d'un parlant determinat però no n'estem del tot segurs, tenim l'opció de substituir l'última lletra de les inicials del parlant corresponent per ?, com a (38b), en què AAC, un dels participants en la mateixa subhasta de peix, és substituït per AA?

(38a)	453	???:	eh::\	(38b)	465	AA?:	no\
	454		hola:\		466		a mi no m'ho
	455	???:	hombre:\				demaneu\
	456	T??:	xXXx		467		demaneu-li amb
	457	???:	que has anat a la				ell\
			Penya\		468	T??:	xXXx
	458	???:	acabo d'arribar tu\		469	AA?:	jo faig de: matxaca
		(COR: EMP3, adaptat)					només\
						(COR: EMP3, adaptat)	

D'altra banda, el comentari ((a INICIALS)) ens servirà per assenyalar a quin dels diversos interlocutors es dirigeix en un moment determinat el participant que ocupa el torn de parla, cosa que ens pot ser molt útil en transcripcions que no van acompanyades d'imatge. Sempre situarem aquest comentari a l'inici de grup entonatiu i només podrà anar precedit de pauses, com a (39a: 1417). També el podrem utilitzar per a destinataris no humans, veg. (39b: 1242).

(39a)	1413	MER:	(.) el di[lluns\	1418		comença que
	1414		que hi van ana:r]			serà fred\
	1415	AMM:	[els dilluns\	1419	MER:	(.) sí\
	1416		que hi vaig a]nar\	1420		me- --
	1417	TON:	(.) ((a AMM))	1421		mengi\
			mare\		(COC: 03, adaptat)	
(39b)	1242	NAT:	((al gos)) Tresqui\			
	1243		Tresqui\			
	1244		té té té\			
		(COC: 03, adaptat)				

#### 4.4.6 Pauses

Les pauses tenen un paper molt important en tota interacció verbal (veg. Tusón 1995: 44, Calsamiglia i Tusón 1999: 20) i sovint també caldrà recollir-les en la transcripció. Generalment, les anotarem entre parèntesis, per bé que amb variacions segons el nivell de transcripció amb què treballem. En una modalitat ampla n'hi haurà prou representant les pauses més llargues o rellevants amb el símbol (*pausa*), com a (40a: 346), mentre que en una modalitat intermèdia i sobretot en una d'estreta apuntarem totes les pauses amb una durada mínima de 0.10 segons, que haurem mesurat fent servir les eines informàtiques adequades (veg. § 2.3), com a (40b: 343, 344, 346).

(40a) 343 FAN: tu de blanc no et cases.  
344 NAA: per què no?  
345 FAN: perquè no.  
346 HIL: (pausa) (riure)  
347 NAA: de color cru.  
348 {{rient) de color cru?}  
(COC: 02, adaptat)

(40b) 343 FAN: (. 0.25) tu de blanc no et cases\  
344 NAA: (.. 0.78) {(P) per què no\_}  
345 FAN: {(P) perquè no}  
346 HIL: (... 1.05) [@@@]  
347 NAA: [de color cru\  
348 {( @) de color cru/  
(COC: 02)

Per facilitar les cerques i la identificació de les pauses, davant de la xifra que n'indica la durada hi afegirem els punts que corresponen segons la taula 4.2. Al fragment de (41), extret d'un ple d'ajuntament, hi ha exemples de cada tipus de pausa: molt breu (g. t. 1108), breu (g. t. 1107, 1111, 1112, 1113, 1115), mitjana (g. t. 1117, 1118) i llarga (g. t. 1103, 1105).

Pausa	Durada	Notació
-------	--------	---------

Molt breu	de 0.10 a 0.29 segons	(. durada)
Breu	de 0.30 a 0.99 segons	(.. durada)
Mitjana	d'1 a 2.99 segons	(... durada)
Llarga	igual o superior a 3 segons	(.... durada)

Taula 4.2 Durada i notació de les pauses

- (41) 1103 (... 3.02) ((remor de fulls i algú que xiuxiueja))  
1104 SOS: la primera pregunta\_  
1105 (... 3.77)  
1106 SOS: boques d'incendi\ ((llegint))  
1107 (.. 0.94) té la policia municipal\_  
1108 (. 0.13) un plànol\_  
1109 on siguin senyalitzades\_  
1110 les boques d'incendi\  
1111 (.. 0.76) quantes n'hi ha al nostre municipi\  
1112 (.. 0.30) on són ubicades\  
1113 (.. 0.60) qui té cura:\_  
1114 del seu manteniment i revisió\  
1115 (.. 0.79) amb quina periodicitat\_  
1116 es fa aquesta revisió\  
1117 (... 1.03) són homologades les rosques amb les que  
utilitzen els bombers\  
1118 (... 1.35) ((a SOO)) el regidor ha de donar la contesta\  
1119 de governació\  
(COR: POL3, adaptat)

Per defecte, les pauses es col·locaran al principi de grup tonal —precedint qualsevol altre codi—, mai al final. Per tant, sempre les atribuirem al parlant que intervé immediatament després de la pausa i evitarem transcripcions com la de (42a), que corregirem per (42b).

- (42a) 986 ALK: l'has trobat/ (.. 0.68)  
987 DIU: eh/  
988 ALK: l'has trobat aquella altra gravació/ (.. 0.78)  
989 DIU: no no\  
990 és que no l'he buscada\  
(42b) 986 ALK: l'has trobat/  
987 DIU: (.. 0.68) eh/  
988 ALK: l'has trobat aquella altra gravació/  
989 DIU: (.. 0.78) no no\  
990 és que no l'he buscada\

(COR: CULT3, adaptat)

Com hem assenyalat més amunt (veg. § 4.4.1), també podem trobar pauses a l'interior de grup entonatiu, si bé es tracta d'un fenomen força excepcional. Aquí ho exemplifiquem amb l'enunciat de (8), reprès com a (43).

- (43) 480 MMM: haig d'anar a tirar el bròquil jo\  
481 PPP: (. 0.22) s'ha de treure aquella pasta del (. 0.18)  
de la neve[ra\  
482 MMM: [de] la nevera\  
(COC: 07)

Si es produeix una pausa dins un grup tonal i l'interlocutor l'aprofita per ocupar el torn de parla, com succeeix a (44a: 221-222) —en què CME diu *no* així que MMM fa una pausa interna al grup tonal 221—, tenim l'opció d'interrompre el grup tonal que conté la pausa i reprendre'l posteriorment amb el símbol & i sense numeració de línia, com es mostra a (44b).

- (44a) 221 MMM: (. 0.24) tots els vells que estaven aquí han anat a  
Montgat (. 0.13) per fer això de la reforma\  
222 CME: no\  
223 ja:\_  
224 però vull dir que:\_  
(COC: 07, adaptat)

- (44b) 221 MMM: (. 0.24) tots els vells que estaven aquí han anat a  
Montgat (. 0.13)  
222 CME: no\  
MMM: & per fer això de la reforma\  
223 CME: ja:\_  
224 però vull dir que:\_  
(COC: 07, adaptat)

A la pràctica, però, el fragment de parla inserit per l'interlocutor rarament coincideix exactament amb la pausa interna del grup tonal anterior. Normalment el que trobem és un encavalcament entre un segment de parla posterior a la pausa i el segment inserit per l'interlocutor que ocupa el torn puntualment, tal com reflecteix (45), en què els claudàtors representen

l'encavalcament. Per tant, tot i que és una possibilitat prevista pel model, és molt poc probable que hàgim d'utilitzar el símbol &, que reservarem per a transcripcions estretes (veg. § 4.5.2).

- (45) 221 MMM: (. 0.24) tots els vells que estaven aquí han anat a  
Montgat (. 0.13) [per fer] això de la reforma\  
222 CME: [no:\]  
223 ja:\_  
224 però vull dir que:\_  
(COC: 07, adaptat)

A vegades les pauses mitjanes i llargues s'han d'atribuir al conjunt dels parlants més que no pas a un de concret, com a (46: 288, 290), en què RRR i PEI van fent pauses relativament llargues durant la conversa perquè estan mirant la televisió. En aquests casos les pauses apareixeran en una línia independent numerada, així com tots els silencis d'una durada de 3 segons o més, que per convenció també tindran línia pròpia, veg. (46: 294). En el grup tonal posterior a la pausa sempre indicarem el parlant que intervé, encara que sigui el mateix que ja ocupava el torn de parla. Per exemple, vegeu que a (46: 289, 291, 295) es repeteixen les inicials RRR malgrat que aquest interlocutor ja disposava de torn de parla abans de la pausa.

- (46) 287 RRR: això també ho feia el Cruyff per això\  
288 (... 1.33)  
289 RRR: està ben buit eh:\ ((es refereix al Camp Nou))  
290 (... 1.42)  
291 RRR: això és avui no:/  
292 PEI: (. 0.43) {(??) esclar\  
293 RRR: {( @) ah\  
294 (.... 5.36)  
295 RRR: què els hi passa a aquests de cop i volta\  
((referint-se als jugadors del Barça))  
296 estaven la mar de bé i ara de cop i volta\_  
(COC: 08, adaptat)

Si la pausa dura un minut o més, ho indicarem amb quatre punts davant la xifra i un apòstrof entre els minuts i els segons, per exemple, (... 2'78), i també la situarem en una línia independent numerada.

Finalment, en una transcripció estreta emprarem el símbol (0) per assenyalar que l'inici d'un torn de parla s'enllaça immediatament amb el final del torn precedent, sense pausa perceptible. Tal com es representa a (47), col·locarem aquest símbol a l'inici del segon torn (veg. també § 4.5.2).

- (47) 918 ANA: vols venir a dinar/  
919 PAU: (0) sí\

#### 4.4.7 Encavalcaments

Per delimitar els encavalcaments que es produeixen en interaccions dialògiques, inserirem claudàtors d'obertura i tancament allà on comencen i acaben els fragments de parla superposats, que disposarem verticalment —no en columna o en partitura, veg. § 3.2. D'encavalcaments, n'hi ha de dos tipus: els simples, que relacionen segments orals corresponents a dos parlants i que habitualment no requereixen números per ser aparellats, com a (48a: 273-274, 277-278), i els múltiples, que agrupen segments enunciats alhora per més de dos interlocutors i que són numerats per facilitar la lectura del text, veg. (48b: 515-519).

- (48a) 271 LLU: (... 1.14) i la meva mare:\_  
272 era dobladora\_  
273 TBB: (.. 0.86) [era què/]  
274 LLU: [de filats\  
275 dobladora\_  
276 (.. 0.42) de filats\_  
277 sap què són [filats\_]\  
278 TBB: [sí\  
(COS: LLU, adaptat)
- (48b) 515 COC: el Levi[tane és per 1] quan [2 tens--]  
516 MAT: [1 clar\  
517 ANA: [2 qui ho va comprar això\  
518 ONI: [2 bueno\  
519 si et prens antibiòtic] una pastilla d'aquestes no

et [3 farà re:\  
 520 no et farà mal\]  
 521 ANA: [3 qui la va comprar\]  
 (COC: 04, adaptat)

Si convé, però, els encavalcaments simples també incorporaran números. A (49), per exemple, se'n podria prescindir, perquè només hi ha dos parlants que intervenen alhora, però el transcriptor ha considerat oportú afegir-hi xifres per ajudar el lector a fer correspondre uns encavalcaments amb els altres.

(49) 450 ELE: i saps de quina botiga és/ ((es refereix a un regal))  
 451 de la Natura\  
 452 (.. 0.43) [1 d'allà\  
 453 SER: [1 ah sí:\]  
 454 ELE: [2 xXXx]  
 455 SER: [2 és maca eh\  
 456 [3 el que a mi em van regalar l'any passat també és  
 de Natura\  
 457 l'any passat els regalos de xx tots eren amb l'envoltori  
 Natura\  
 458 ELE: [3 xXXx]  
 (COC: 09, adaptat)

Un mateix número pot agrupar més de dos fragments encavalcats, com a (50a), en què tres parlants primer pronuncien alhora la mateixa interjecció (g. t. 2133-2135) i després es tornen a encavalcar en els grups tonals 2137-2139. Com es pot observar, en el segon grup d'encavalcaments es reinicia la numeració, però també seria possible representar el mateix fragment seguint la seqüència numèrica, com a (50b: 2137-2139).

(50a) 2130 NAT: no\ 2131 perquè: jo mataré l'Arnau i ja està\ 2132 AMM: oi_ 2133 [1 oh::\] 2134 MER: [1 oh::\] 2135 RAS: [1 oh::\] 2136 AMM: (.. 0.36) [això\]	2137 MER: [pocas]:o:l:[ta\ 1] 2138 AMM: [1 @] 2139 ARN: [1 xx]xxx (COC: 03, adaptat)	(50b) 2130 NAT: no\ 2131 perquè: jo mataré l'Arnau i ja està\ 2132 AMM: oi_ 2133 [1 oh::\] 2134 MER: [1 oh::\] 2135 RAS: [1 oh::\] 2136 AMM: (.. 0.36) [això\]
---	---	---



2132	AMM:	oi_	2137	MER:	[2 pocas]:o:l:[ta\ 3]
2133		[1 oh:::]	2138	AMM:	[3 @]
2134	MER:	[1 oh:::]	2139	ARN:	[3 xx]xxx
2135	RAS:	[1 oh:::]		(COC: 03, adaptat)	
2136	AMM:	(. 0.36) [2 això]			

Els números apareixeran per defecte al principi de l'encavalcament, entre el claudàtor d'obertura i el text —i amb espai només entre la xifra i el text—, com es pot comprovar a (48b), (49) i (50). Si l'encavalcament comença a l'interior de mot, col·locarem el número al final del fragment encavalcat, entre el símbol d'entonació i el claudàtor de tancament, per facilitar la lectura; vegeu com a exemples (50: 2137) i (51a: 1910). En cas que no sigui possible cap d'aquestes dues opcions, inserirem el número dins el mot, sense espai, com a (51a: 1912), o el col·locarem en una posició que no destorbi gaire la lectura, com a (51b: 1597), en què el transcriptor va optar per situar-lo darrere el símbol d'entonació.

- (51a) 1907 VIE: i vaig pensar\  
1908 hosti\  
1909 tot això\_  
1910 com a mínim per aixecar-[ho\_ 1]  
1911 MAG: {(??) [1 es] te--}  
1912 et fa[2rà un bo]:ny\  
1913 VIE: [2 ja\  
(COC: 10, adaptat)
- (51b) 1592 MAG: ai\  
1593 goita que és maco\ ((referint-se a un regal))  
1594 VIE: i surt cart\  
1595 MAG: bos[se:s\ 1]  
1596 ALL: [1 però bueno\  
1597 són de bosses de lavan[da\ 2  
1598 per fer o]lor\  
1599 MAG: [2 ah::\  
(COC: 10, adaptat)

La transcripció de fets de parla en què els interlocutors sovint intervenen alhora, com ara una conversa col·loquial, sol comportar un nombre considerable d'encavalcaments. En

aquest cas, per ajudar a interpretar el text, podem inserir més d'un número en cada fragment encavalcat. Per exemple, a (52a: 1582-1583) el final de les intervencions superposades de JJJ i IDA coincideix amb l'inici d'un grup tonal enunciat per un parlant no identificat, i per representar-ho d'una manera més entenedora s'han repetit els números abans dels claudàtors de tancament (g. t. 1582), amb un espai abans de la xifra. A (52b) i (52c) succeeix una cosa semblant: a (52b: 1742-1743) l'encavalcament afecta l'interior d'un fragment de parla que ja està superposat a un fragment anterior, i a (52c) el grup tonal 958 s'encavalca parcialment amb el dos grups tonals posteriors. Repetint el número al final, indiquem més clarament on s'acaba cada encavalcament.

(52a) 1579 JJJ: [1 xXXx] ((sobre el dinar de Nadal))  
 1580 IDA: [1 jo el que vull és que\  
 1581 si coincidiu vosatres\_  
 1582 ho fem a [2 casa nostra\ 2] 1]  
 1583 ????: [2 xxxxx] la seva àvia\  
 1584 IDA: sents/  
 (COC: 09, adaptat)

(52b) 1739 IDA: (. 0.18) et vaig posa[:r aquell bolso que vas comprar]  
 per l'àvia\_  
 1740 ????: [xxx aquest pàjaro\  
 1741 ????: (. 0.53) ah és veritat [1 xXXx\  
 1742 IDA: [1 està a [2 dins d'un 2] calaix\ 1]  
 1743 ????: [2 home:\  
 1744 IDA: (. 0.17) a casa teva\  
 (COC: 09, adaptat)

(52c) 952 HIL: cuidado\  
 953 quan entres al restaurant sí\ ((s'aplaudeixen els nuvis))  
 954 NAA: (.. 0.55) sí::\  
 955 HIL: ah\  
 956 SUS: (.. 0.39) ah: al restaurant\  
 957 NAA: @@  
 958 SUS: [1 @@@@]  
 959 NAA: [1 [2 a l'esglé1]sia no\ 2]  
 960 HIL: [2 jo estava pensant] l'església\_  
 (COC: 09, adaptat)

Sempre procurarem respectar l'ordre cronològic del discurs, de manera que en els fragments encavalcats els grups tonals més llargs seguiran els grups tonals més curts, mai a la inversa. Per tant, en l'exemple següent la transcripció correcta serà (53a: 357-358) i no pas (53b: 357-358), perquè altrament no reflectiríem l'ordre temporal del discurs.

(53a) 354 FAN: i:\_  
 355 i per què portes els cascos posats/  
 356 (... 2.64)  
 357 FAN: [no ho entenc\  
 358 HIL: [perquè està es]coltant la ràdio\  
 (COC: 02, adaptat)

(53b) 354 FAN: i:\_  
 355 i per què portes els cascos posats/  
 356 (... 2.64)  
 357 HIL: [perquè està es]coltant la ràdio\  
 358 FAN: [no ho entenc\  
 ]

Malgrat tot, no sempre podrem respectar aquest ordre temporal. N'és un bon exemple el segment de (54a: 1081-1082), en què l'encavalcament té lloc entre un fragment intern del grup tonal 1081 i tot el 1082, i el mateix passa en els grups entonatius 1084 i 1085. Per resoldre-ho, podem deixar el primer grup tonal inacabat i reprendre'l a sota amb el símbol & (veg. § 4.4.6), com a (54b). Ara bé, tot i que aquest procediment garanteix l'ordre cronològic del discurs, també obliga a deixar grups tonals inacabats i fins a cert punt pot dificultar la lectura del text. Una altra possibilitat és alterar l'ordre natural del discurs però enumerar els encavalcaments per aparellar-los més fàcilment, com a (54c). Aquesta darrera opció serà la solució per defecte en una transcripció intermèdia-estreta.

(54a) 1080 EUU: {( @) és que sou més burros\  
 1081 MJJ: si és lo [que pen]ses\_  
 1082 EUU: [vull dir\_]\  
 1083 MJJ: estic tan content\  
 1084 perquè [si tu] ho penses\_  
 ]

1085 EUU: [clar\  
1086 MJJ: estic segur que no és així\  
(COC: 05, adaptat)

(54b) 1080 EUU: {( @) és que sou més burros\  
1081 MJJ: si és lo [que pen]  
1082 EUU: [vull dir\_  
MJJ: & ses\_  
1083 estic tan content\  
1084 perquè [si tu]  
1085 EUU: [clar\  
MJJ: & ho penses\_  
1086 estic segur que no és així\  
(COC: 05, adaptat)

(54c) 1080 EUU: {( @) és que sou més burros\  
1081 MJJ: si és lo [1 que pen]ses\_  
1082 EUU: [1 vull dir\_  
1083 MJJ: estic tan content\  
1084 perquè [2 si tu] ho penses\_  
1085 EUU: [2 clar\  
1086 MJJ: estic segur que no és així\  
(COC: 05, adaptat)

Finalment, farem servir el símbol [xXXx] per representar un fragment de parla encavalcat que conté un nombre indeterminat de síl·labes intel·ligibles, veg. (55a: 828), i, si cal, anotarem la durada del segment en segons amb la variant del mateix símbol [xX(durada)Xx], veg. (55b: 401). L'encavalcament pot ser total, com el de (55a: 828), o parcial, com a (55b: 401), en què és possible identificar dues síl·labes no encavalcades. Vegeu l'apartat següent per a la notació de fragments intel·ligibles.

(55a) 825 MER: ((a TON)) te'n 827 tots [arrossegant\  
recordes de la 828 ???: [xXXx]  
pel·lícula de l'ET/ 829 OSE: @  
826 OSE: vam arribar tard al 830 ???: oh:\  
cine\ (COC: 03, adaptat)

(55b) 400 AMM: aquell dia [que vam anar a Roses\  
401 NAT: [xX(0.91)Xx]xx

402 AMM: que vam seure allà\  
403 a dinar\  
404 te'n recordes\  
(COC: 03)

#### 4.4.8 Fragments intel·ligibles

Sovint trobarem fragments de parla difícils d'entendre, fins i tot quan l'enregistrament té una bona qualitat de so, ja sigui perquè l'informant abaixa el to de veu o perquè no vocalitza bé, parla massa de pressa, etc. Si, a més, el so de la gravació és deficient (amb soroll de fons, cops al micròfon, desplaçament dels interlocutors, etc.), en alguns moments serà molt més complicat reconèixer què diuen els informants. En aquests casos intentarem esbrinar almenys quantes síl·labes conté cada fragment intel·ligible i les representarem amb una ics cadascuna, com a (56a: 380), en què el transcriptor no ha pogut entendre el mot de tres síl·labes que tanca el grup entonatiu. No cal que les síl·labes estiguin juntes. Si ens fa la impressió que pertanyen a mots diferents, les distribuïrem tal com les sentim, com a (56b: 1,3), en què s'intueix que s'enuncien dos mots diferents en cada grup tonal, tot i ser intel·ligibles. A vegades ni tan sols podrem distingir el nombre de síl·labes que es pronuncien, per la qual cosa utilitzarem el símbol *xXXx*, que representa un fragment de parla intel·ligible de durada indeterminada. Si volem indicar-ne la durada en segons, emprarem la variant *xX(durada)Xx*, com a (56c: 48)

(56a) 380 ANA: puc menjar un xxx\  
381 COC: {(??) bueno\  
(COC: 04)

(56b) 1 ??? : xxx xxxx\  
2 BEB: va:\  
3 ??? : xx xxx\

4 MAM: @@@ @  
5 al menjador tothom\

(COC: 06)

(56c) 48 COC: xX(1.03)Xx\  
49 JOJ: no la llencis aquesta cansalada\  
50 dóna-la a ta mare\  
(COC: 04)

Tractarem els fragments intel·ligibles com a mots i, per tant, els combinarem amb altres convencions: claudàtors (57a: 424), terminacions tonals (57a: 425), truncaments (57b: 326), etc. Excepcionalment, si un grup tonal acaba amb un fragment intel·ligible i una inflexió terminal poc clara, deixarem el grup entonatiu sense el símbol prosòdic, com a (57c: 1752).

(57a) 422 LAU: (.. 0.39) escolta\  
423 però:\_  
424 MMA: xx[xx]  
425 LAU: [a mi no] em vas dir  
que el Pere et va xx\  
(COC: 01, adaptat)

(57c) 1751 REP: el clero\_  
1752 LAU: {(P) xx[x]}  
1753 REP: [saps] què és el  
clero\  
(COC: 01)

(57b) 325 ANI: (. 0.22) jo no me'l  
podia allisar\ ((el  
cabell))  
326 (.. 0.95) {(AC) però  
ara xx--}  
327 com que el porto  
curt ara\_  
(COC: 06, adaptat)

En cas que no acabem d'estar segurs del que diu l'informant, marcarem el fragment dubtós amb el codi *{(??) text afectat}*, veg. (58a: 836). Si es combina amb altres codis, anirà per defecte en primera posició, com a (58b: 679), en què apareix amb el codi d'elocució accelerada (veg. 4.4.9.1).

(58a) 836 MAM: {(??) al final em} posaré una mica més de cafè\  
837 ????: (.. 0.62) x  
838 BEB: (.. 0.96) ja t'ho porto\  
(COC: 06)

(58b) 676 ESF: els missioners aquests\_  
677 que van als puestos\_

678 (.. 0.52) jo no trobo que lis i solucionen gaire cosa\  
679 (.. 0.95) {(??)(AC) llavontes\_}  
680 jo no crec en aquestes coses\  
(COS: ESF)

En una primera versió de la transcripció, no cal dedicar gaire temps als fragments intel·ligibles; val més continuar transcrivint per no aturar-se massa en un punt concret. A mesura que anem avançant, ens anirem acostumant a l'idiòlecte de cada informant i als temes que es tracten al llarg de la interacció verbal, cosa que ens donarà bones pistes per desxifrar els fragments més problemàtics en revisions posteriors.

#### **4.4.9 Aspectes vocals no verbals**

En tot fet de parla són abundants els elements vocals que no tenen un significat pròpiament lingüístic però que, com les pauses i els encavalcaments, també proporcionen molta informació tant als participants en l'esdeveniment comunicatiu com a l'analista del discurs oral. Dins d'aquests aspectes hi incloem des de correlats prosòdics com ara el *tempo* i la intensitat (veg. § 4.4.9.1) fins al riure (veg. § 4.4.9.2), les inhalacions i exhalacions (veg. § 4.4.9.3) i els sons paralingüístics (veg. § 4.4.9.4).

##### *4.4.9.1 'Tempo', intensitat i variacions tonals*

Abans de començar a transcriure és molt recomanable familiaritzar-se amb l'idiòlecte de cada informant. Aspectes com la velocitat d'elocució o la intensitat són correlats prosòdics molt idiosincràtics que només es poden mesurar de manera relativa. Per exemple, si un informant acostuma a parlar amb un ritme d'elocució ràpid, no caldrà que assenyalem contínuament que produeix grups tonals a una velocitat elevada —ho podem indicar en la fitxa sobre els informants que complementa la transcripció. Per tant, per evitar ser redundants

i enfarfegar massa el text, només anotarem aquells aspectes que siguin rellevants en relació amb l'idiòlecte de cada parlant i les circumstàncies comunicatives de la interacció.

Si el *tempo* o ritme d'elocució és accelerat, delimitarem el text amb el codi *{{(AC) text afectat}}*, com a l'exemple (59a), procedent d'una subhasta de peix; i si és desacelerat, amb el codi *{{(DC) text afectat}}*, com a (59b: 1331), en què el parlant demana als interlocutors que no cridin tant.

(59a)	179	JRR:	{{(AC) dos a la	(59b)	1328	MSA:	p(e)rò mama\
			banda_		1329		jo t'ho he dit
	180		noranta-nou_				[alguna
	181		noranta-vuit_				vegada/]
	182		noranta-set_		1330	SER:	[xi:~]
	183		noranta-sis_		1331		{{(DC) que a
	184		noranta-cinc_				l'Oriol no li
	185		noranta-quatre_				deixeu dormi:r\}
	186		noranta-tres_			(COC: 09, adaptat)	
	187		noranta-dos_				
	188		noranta-u_				
	189		noranta\}				
			(COR: EMP3)				

El volum de veu és un dels correlats prosòdics més versàtils, perquè pot servir, entre altres funcions, per marcar certes construccions sintàctiques, com ara dislocacions i focalitzacions; donar expressivitat al discurs o guanyar el torn de parla. Com en el cas del *tempo*, en la transcripció només reflectirem canvis d'intensitat que siguin rellevants en relació amb l'idiòlecte de cada informant. Farem servir el codi *{{(F) text afectat}}* quan la intensitat sigui forta (*forte*), com a (60a: 830-831), que reproduïx una discussió familiar, i el codi *{{(P) text afectat}}* quan sigui fluixa (*piano*), com a (60b: 734-737), en què s'expressa afecte.

(60a)	827	MEC:	no parlis així: home:\
	828	ONI:	oh és que_
	829		(. 0.43) saps lo que mengen\
	830		(. 0.19) {{(F) saps lo que [mengen aquests dos\
	831		estan--]}
	832	MAT:	[mo:lt\}



	833	ONI:	estan--		736	PPP:	(... 2.84) {(P) no
	834		estan bojos\				discutim tant tu i
		(COC:	04, adaptat)				jo\
(60b)	733	PPP:	(. 0.33) bonica\		737		oi que no com
	734	CME:	(. 0.24) {(P) hola\				abans discutíem\}
	735		(. 0.18) simpàtic\}		(COC:	07, adaptat)	

Les variacions tonals són un altre recurs expressiu freqüent. Per exemple, a vegades trobem fragments de parla pronunciats amb un to alt o baix mantingut per dotar-los d'un efecte comunicatiu concret. Si el to és alt, com en la veu de falset, emprarem el codi {(A) *text afectat*}, i si és baix, {(B) *text afectat*}. Vegeu com a mostra (61a: 1303), en què LAU es queixa d'un cop que li han donat, i (61b: 232), en què MIA abaixa la seva tessitura habitual per denotar seriositat, gravetat, perquè li tornin unes pintures.

(61a) 1303 LAU: {(A) no em pi:quis\}  
1304 MMA: si només t'he tocat\  
1305 LAU: {(F) xata\  
1306 em fa molt mal aquest braç\  
(COC: 01)

(61b) 232 MIA: i a mi: m'ha desaparegu:t {(B) la capsa de pintures meves\  
233 eh Arnau\  
234 (. 0.27) {(P) sí\  
235 ARN: (. 0.22) quines pintures\  
(COC: 03)

Sovint el que es produeix és una combinació dels recursos anteriors en el mateix segment discursiu per imitar la parla d'algú altre. Sempre que sigui pertinent, anotarem aquests fragments amb el codi d'extrafer la veu {(EV) *text afectat*}, que a part de ser molt propi de les citacions, veg. (62a: 671, 676), i les imitacions, veg. (62b: 264-266, 62c: 907), també sol aparèixer en modalitats discursives molt específiques, com ara la parla adulta adreçada a infants (to alt mantingut, *tempo* lent, etc.), exemplificada a (62d: 299).

(62a)	699	TON:	què fa la Rose:r\	(62c)	905	NAA:	(. 0.62) diu ai\
	670		què fa d'això:\		906		diu perquè_
	671		(. 0.59) {(EV) i en Berna:t\}		907		{(EV) es que habla de una manera así_}
	672	AMM:	[1 @ @ @ @]		908		(. 0.79) oi que sí/ (COC: 02, adaptat)
	673	???:	[1 xX(1.11)Xx]				
	674	AMM:	[2 @ @ @ @ @]				
	675	TON:	[2 m'ho pregunta aixís eh\ {(EV) i en Berna]:t\}	(62d)	296	MAT:	(. 0.75) i per què el vols tu el xupete\ 297 COC: ah\ 298 no ho sé\ 299 MAT: {(EV) que tu ja ets una nena gran\}
	676		(COC: 03, adaptat)				(COC: 04)
(62b)	259	NAA:	(. 0.24) el dia que van parlar amb el mossèn_ 260 (. 0.52) perquè_ 261 va entrar_ 262 i bueno\ 263 allò que els començava a parlar_ 264 {(EV) que si el casament_ 265 i que si Déu_ 266 (. 0.26) i que si no sé quantos_}				
			(COC: 02, adaptat)				

Com en el cas dels encavalcaments, és possible que dos o més codis coincideixin en un mateix fragment de parla. Si afecten exactament el mateix segment, els anotarem dins de les mateixes claus, sense deixar espai, i els ordenarem per ordre alfabètic, com a (63a), en què MOL, que explica un conte, pronuncia els grups tonals 65-68 amb una intensitat baixa i estafent la veu. En canvi, si els codis no afecten exactament el mateix fragment de parla, assenyallarem on acaba el codi més llarg repetint-ne la lletra abans de la clau de tancament per evitar ambigüitats; vegeu com a mostra (63b: 923), amb un mot pronunciat amb intensitat forta dins un segment accelerat, i (63c: 59), amb un grup tonal amb intensitat baixa enmig d'un codi de veu estafeta. Sempre es deixarà un espai entre el

símbol d'entonació i els parèntesis que contenen la lletra del codi de tancament. D'altra banda, si dos o més codis que afecten segments diferents acaben o comencen en el mateix punt —com a (63d: 1675), en què coincideixen (*P*) i (*L2*), codi de segona llengua (veg. § 4.4.15.5)—, no deixarem espai entremig.

<p>(63a) 63 MOL: (.. 0.78) obriu obriu\          64 que sóc la vostra          mare\          ((se senten dos          cops))          65 {(EV)(P) oh\          66 (.. 0.85) ai no\          67 la mare no pot ser\          68 (.. 0.51) que la mare          no té aquesta veu\          69 (... 1.64) deu ser el          llop\          (COR: EDUC2, adaptat)</p>	<p>(63c) 57 MOL: (.. 0.70) {(EV) ui\          58 (.. 0.50) {(P) qui          truca\          59 la mare ens ha dit que          no obrim\          (EV)}          (COR: EDUC2)</p>	<p>(63d) 1675 TUT: (.. 0.33) {(L2)}{(P)          pues_          1676 buscar_          1677 una persona en          Pineda_ (L2)}          1678 (.. 0.42) [que--]          1679 PRO: [que] sàpiga          llengua de signes\          (COR: EDUC1, adaptat)</p>
<p>(63b) 921 GIA: (. 0.26) {(AC) ahora          que recordem_          922 que ell_          923 {(F) comença} a          conèixer els artistes          surrealistes_          924 estant a París\          (AC)}          (COR: CULT2)</p>		

#### 4.4.9.2 *Riure*

Representarem el riure mitjançant el símbol @, que correspondrà a un sol cop de veu, és a dir, a un *ha!* de la transcripció ortogràfica de les rialles en còmics, obres de teatre, etc., veg. (64a: 642-643). Per defecte, el riure tindrà línia pròpia, com a (64a: 642-643), però si resulta difícil destriar-lo de la parla, l'inclourem dins el grup tonal corresponent, com a (64a: 644); si, a més, a causa del riure costa identificar la seqüència tonal terminal, excepcionalment podem deixar el grup entonatiu sense la marca prosòdica final, veg. (64b: 473).

(64a) 638 T??: @R(3.97)R@  
 639 MAM: {(F) a mi també\  
 640 BEB: (.. 0.99) molt [bé\  
 641 ANI: {(F) [la] Marina també vol polvos\  
 642 ???: (.. 0.37) @@  
 643 BEB: @@  
 644 (.. 0.69) {( @ ) ai ai ai\  
 645 MAM: @R(7.10)R@  
 (COC: 06)

(64b) 469 MER: i em diu\  
 470 AMM: no se n'havia posat ma[:i\  
 471 MER: [han] [sortit\  
 472 AMM: [xxxx]xxx de Mataró\  
 473 i mira es @@@@  
 (COC: 03, adaptat)

Per norma general, no hi haurà espai entre dues o més arroves, veg. (64b: 473), però si percebem que les rialles es reparteixen en grups, intentarem reflectir-ne la distribució deixant espais on correspongui, com a (65: 824, 827). Quan el riure és simultani amb la parla, farem servir el codi {( @ ) *text afectat*}, veg. (66: 583), i si convé, el combinarem amb altres codis —precedint els que continguin una lletra—, com a (66: 584).

(65) 822 ANI: (... 1.38) per què em cuida tant\  
 823 la Beli a mi\  
 824 BEB: @ @ @ @  
 825 ANI: {(F) què he fet jo\  
 826 perquè em cuidi [tant\  
 827 BEB: [ @ @ @ ]  
 (COC: 06, adaptat)

(66) 583 MER: {( @ ) i va continuar bevent\  
 584 AMM: {( @ )(F) i tota l'aiga per sobre així\  
 (COC: 03)

Si tots els parlants riuen alhora i no podem determinar a qui corresponen les rialles, emprarem el símbol @RR@ per atribuir-les a tot el conjunt dels interlocutors, representats per T?? (veg.

§ 4.4.5). Normalment assenyalarem la durada del fenomen entre parèntesis mitjançant *@R(durada)R@*, com a (67: 999).

- (67) 998 TON: en lloc d'anar així anàvem fotent així\ ((simulant l'avió amb les mans))  
999 T??: @R(0.91)R@  
1000 ????: (.. 0.80) (INH)  
1001 TON: i amb una de les baixades vam ensapegar l'illa\  
(COC: 03, adaptat)

Altres vegades, més que una rialla, el parlant fa un somriure, que pot afectar un fragment de parla. Si el fenomen és simultani amb el text, utilitzarem el codi *{(somrient) text afectat}*, com a (68a: 51-53); si no és simultani, disposem dels comentaris del transcriptor *(somriure)* i *(somriu)*, veg. (68b: 661) (veg. § 4.4.13 per als comentaris del transcriptor).

- (68a) 45 I05: (.. 0.38) per sort no va passar re:s\_  
46 vull dir no:\_  
47 (INH)  
48 no va ser re:\_  
49 l'únic que jo vaig estar\_  
50 doncs\_  
51 {(somrient) am(b) el collarí:n\_  
52 dura:nt\_  
53 un piló: de te::mps\_}  
(CAP: C1CN05CC)
- (68b) 658 ALK: (.. 0.60) les coses són barates\ ((als encants))  
659 regateja-ho fort\_  
660 [i::\_]   
661 FAA: [(somriure)]  
662 ALK: sí home\  
663 (... 1.47) vinga\  
(COR: CULT3)

#### 4.4.9.3 Inhalacions i exhalacions

En enregistraments amb bona qualitat de so és possible distingir les inhalacions i exhalacions que fan els informants durant l'elocució, i que aquí representarem amb els símbols

(*INH*) i (*EXH*) respectivament. De fet, en alguns enregistraments la proximitat del parlant al micròfon fa que aquests sons orals no verbals siguin tan perceptibles que haurem de triar-los i anotar només els més rellevants segons les circumstàncies comunicatives. Per exemple, una inhalació pot indicar que l'interlocutor es disposa a ocupar el torn de parla, com a (69a: 1442), i una exhalació pot adquirir diversos matisos expressius, com passa a (69b: 367), en què serveix per mostrar desacord. Tant (*INH*) com (*EXH*) tindran línia pròpia, i si són simultanis amb un fragment de parla, els anotarem mitjançant els codis *{{(INH) text afectat}}* i *{{(EXH) text afectat}}*, veg. (69c: 1661).

(69a) ((sobre les dificultats de mossegar llonganissa massa seca))

1441 MAG: *{{(AC) nosatres vam anar bastant bé}}*

1442 JON: *[(INH)]*

1443 MAG: [en Joan com] que no té pas problema\

1444 JON: jo gràcies a Déu\_

1445 amb les dents és amb lo que vaig més bé\

1446 més que amb les cames\

(COC: 10, adaptat)

(69b) 366 REP: explica-m'ho que a mi no m'expliques res\

367 MMA: (*EXH*)

368 sí que t'ho explico\

369 REP: tu no m'expliques re\

(COC: 01, adaptat)

(69c) 1661 ALL: *{{(EXH) bue:no}}* ((fressa de papers))

1662 (... 2.12)

1663 ALL: doncs què\

1664 com ha anat aquest- s dies\

(COC: 10)

Du Bois *et al.* (1993) proposen símbols específics per a les inhalacions i les exhalacions llargues, que aquí adaptem com a (*INH**HHH*) i (*EXH**HHH*). Nosaltres, però, deixarem aquests símbols per a les transcripcions estretes i preferirem els comentaris del transcriptor (*sospir*), (*sospira*), etc. (veg. § 4.4.13).

Finalment, cal tenir en compte que sovint tots dos sons van precedits d'una pausa i pot ser difícil saber on comencen exactament. Per això recomanem que en cas de dubte no s'anoti la pausa, sobretot si és molt curta, i es consideri que forma part de la inhalació o l'exhalació.

#### 4.4.9.4 Sons paralingüístics: retroalimentadors i omplidors de parla

En la llengua oral són abundants els sons vocals sense significat lingüístic que han adquirit valors discursius per gestionar la interacció verbal. Per la seva freqüència, cal destacar-ne dos: els retroalimentadors o *backchannels*, que comuniquen que el canal continua obert (veg. Tusón 1995: 46), i els omplidors de parla, que permeten conservar el torn mentre es rumia què es dirà a continuació (veg. Payrató 1988: 206).

Els retroalimentadors de so nasal els representarem mitjançant la forma *mhm*, veg. (70a: 939), que correspondrà a dues síl·labes —per tant, només anotarem *mhm mhm* si el parlant fa servir dos *backchannels* seguits, cosa poc freqüent. En cas que s'encavalquin parcialment amb un altre fragment de parla, només marcarem amb claudàtors la primera ema (*[m]hm*) o la segona (*mh[m]*), com a (70b: 346). Pot tenir el mateix ús la forma *aha* (70c: 96), que encavalcada serà *[a]ha* o *ah[a]*. Veg. § 4.4.3.3 per a l'anotació prosòdica d'aquests elements.

- (70a) 936 AFT: (.. 0.38) hi volem anar bé\_ ((de vacances))  
 937 volem ana:r\_  
 938 (. 0.19) còmodament\_  
 939 DRO: (. 0.16) [mhm\  
 940 AFT: [però] també volem estar una mica en contacte amb  
 la gent d'allà\  
 941 DRO: (. 0.20) bueno\_  
 (COR: EMP4, adaptat)
- (70b) 343 NAR: i allavores\_  
 344 posaria\_  
 345 els diners a plaço fijo\  
 \

346 NNA: (.. 0.43) {(P) mh[m\]}

347 NAR: {(F) [i pe]r lo que em dongue:ssin\_  
 348 ja no treballaria més\}

(COS: NAR)

(70c) 93 I05: (. 0.21) sempre estic fora\  
 94 (.. 0.99) però no faig res allò:\_  
 95 {(P) especial ni\_  
 96 ECC: (. 0.27) {(P) aha\}

(CAP: C1CE05CC, adaptat)

Pel que fa als omplidors de parla, podem optar per dues formes vocàliques que utilitzarem segons l'idiolècte del parlant: a, si el so correspon a una vocal neutra, com a (71a: 20, 22), i e si té un grau de tancament pronunciat, com a (71b: 176). També és corrent el so bilabial nasal *m*, que a vegades es combina amb els omplidors de parla vocàlics, com a (71c: 44), i que separarem amb un espai. Generalment, totes tres formes presenten allargaments i una entonació terminal mantinguda (veg. § 4.4.3.3).

(71a) 19 I06: (.. 0.55) llavors\_  
 20 a::\_  
 21 quan hakis recollit el teu ami:c\_  
 22 a::\_  
 23 (.. 0.46) continues\_  
 24 (. 0.20) fent que el\_  
 25 poliesportiu et quedi a mà dreta:\_  
 (CAP: C1TI06CC)

(71b) 175 GIA: (.. 0.73) i després\_  
 176 (.. 0.30) e::\_  
 177 (.. 0.34) per desig\_  
 178 dels seus pares\_  
 179 va estudiar arquitectura\_  
 (COR: CULT2)

(71c) 41 ECC: i fas algunes activita:ts\_  
 42 en concre:t\_  
 43 o:\_  
 44 I05: a:: m::\_  
 45 (.. 0.36) no\  
 46 (.. 0.42) quedem amb\_



47 tot el::\_  
 48 {(AC) tota la colla en un bar que anem sempre\_  
 (CAP: C1CE05CC)

Un altre so oral no verbal que sovint funciona com a omplidor de parla és l'espetic de llengua, que representarem amb la forma *ntx* i una terminació tonal descendent, veg. (72: 97).

(72) 93 l05: a ve(u)re\  
 94 de\_  
 95 llegi:r\_  
 96 (. 0.22) llegeixo si::\_  
 97 (.. 0.44) ntx\  
 98 m'ha d'enganxar el llibre\  
 99 i aquest\_  
 100 realment em va enganxar\  
 (CAP: C1CA05CC)

#### 4.4.10 Interjeccions i onomatopeies

Un dels recursos expressius més característics de la llengua oral, sobretot en registres informals, són les interjeccions, enteses com a formes lingüístiques que generalment no tenen un significat lèxic referencial però que han adquirit valors pragmàtics molt diversos (veg. Cuenca 2002: 3199). Aquí resulta impossible oferir una llista exhaustiva de les interjeccions pròpies i impròpies més habituals del català i proposar-ne una classificació segons els usos comunicatius més freqüents. Per a aquesta qüestió, remetem el lector a la bibliografia existent (veg. Cuenca 2002) i sobretot al *Diccionari d'onomatopeies i mots de creació expressiva (DOMCE)*, fins ara el recull més complet d'interjeccions i onomatopeies en català. En tot cas, a l'hora de transcriure una interjecció pròpia procurarem triar una forma gràfica que ja tingui una certa tradició escrita (incloent-hi diccionaris generals, glossaris lèxics, etc.) per facilitar-ne la identificació i evitar confusions.

A continuació presentem les grafies que utilitza el sistema de transcripció discursiva del CCCUB per representar algunes

de les interjeccions pròpies més freqüents de la llengua oral, juntament amb alguns dels valors pragmàtics amb què se solen associar. Entre parèntesis, indiquem les obres de referència i corpus en què apareixen documentades. Pel que fa a l'anotació prosòdica de les interjeccions, veg. § 4.4.3.2.

Valor pragmàtic	Forma gràfica
Assentiment	<i>hi hi</i> (DOMCE, CTILC) <i>hu hu</i> (DOMCE, CTILC)
Crida	<i>psist</i> <i>txist</i> <i>xit, psit, pst</i> (DOMCE)
Crida, salutació	<i>eh, ehem, ei, oh</i> (DIEC2, GDLC, DOMCE, CTILC)
Demanda de conformitat	<i>eh, oi</i> (DIEC2, GDLC, DOMCE, CTILC)
Demanda de silenci	<i>xi</i> <i>ss, pst, psit, xit, xist</i> (DOMCE)  Per a la demanda de silenci, el DOMCE també recull <i>psxt, txsit, txit</i> i <i>sst</i> .
Desacord	<i>boah</i> <i>ps</i> (DOMCE)
Desconfiança, dubte, incredulitat	<i>hum</i> (DIEC2, GDLC, DOMCE, CTILC)
Dolor	<i>ai, oi, ui</i> (DIEC2, GDLC, DOMCE, CTILC)
Èmfasi	<i>buf</i> (DOMCE, CTILC) <i>fu</i> (CTILC)
Estranyesa	<i>ai</i> (DIEC2, GDLC, DOMCE, CTILC)
Fàstic	<i>ecs, uf, uix</i> (DIEC2, GDLC, DOMCE, CTILC)
Fruïció, desig	<i>hum</i> (DIEC2, GDLC, DOMCE, CTILC)
Menyspreu	<i>eh</i> (DIEC2, GDLC, DOMCE, CTILC)
Opressió, aclapament	<i>uf</i> (DIEC2, GDLC, DOMCE, CTILC)
Rectificació, objecció	<i>ei</i> (DIEC2, GDLC, DOMCE, CTILC)
Satisfacció, triomf	<i>ha ha</i> (DIEC2, GDLC, DOMCE, CTILC)
Sorpresa, incredibilitat	<i>ai, ha, oi</i> (DIEC2, GDLC, DOMCE, CTILC) <i>hu</i> (DOMCE, CTILC)
Valoració	<i>psepsè</i> <i>pse</i> (DIEC2, GDLC, DOMCE, CTILC)

Altres	<i>ah, oh, xe</i> ( <i>DIEC2, GDLC, DOMCE, CTILC</i> )
--------	--

Taula 4.3 Grafies utilitzades pel CCCUB per representar interjeccions pròpies

Les onomatopeies, que podem definir com a interjeccions amb funció representativa (veg. Cuenca 2002: 3228), també les transcriurem seguint formes gràfiques amb certa tradició escrita perquè siguin fàcils de reconèixer (veg. també el *DOMCE*). Cal recordar que assenyalarem els allargaments vocàlics tan característics de les interjeccions i les onomatopeies fent servir els dos punts, veg. (73: 581, 582), mai repetint la vocal allargada (cf. *bruum, xuuc*).

- (73) 580 PPP: m'ha despertat quan baixava la la la a: (. 0.12) el llit\  
581 (.. 0.76) br[u::m\  
582 MMM: [xu:c\  
583 sí\  
584 jo també ho he sentit\  
(COC: 07)

#### 4.4.11 Altres fenòmens vocals: xiuxiueig, cant, etc.

Hi ha tot un seguit de fenòmens vocals, com ara el xiuxiueig, la tos, els aclariments de gola i el cant, entre d'altres, que no tenen assignat cap símbol específic i que haurem de transcriure com a comentaris del transcriptor o mitjançant un codi. Si aquests sons vocals no són simultanis amb la parla, els anotarem com a comentari del transcriptor, entre parèntesis simples (veg. § 4.4.13) i línia pròpia. En donem alguns exemples a (74): xiuxiueig (74a: 1659), soroll de mocar-se (74b: 961), esternut (74c: 1092), xiulet (74d: 869), imitació d'un soroll (74e: 172, 74f: 1377), esbufec (74g: 480), aclariment de gola (74h: 222) i tos (74i: 992).

- (74a) 1659 ???: (.. 0.48) (xiuxiueig) 1663 ALL: doncs què\  
1660 MAG: (.. 0.34) m bue:no 1664 com ha anat  
bueno\  
aquest- s dies\  
1661 ALL: {(EXH) bue:no} (COC: 10, adaptat)  
1662 (... 2.12) (74b) 959 MMA: (.. 0.33) @@@@

960	(... 2.06)				
961	(soroll de MMA		1375	estrident brutal\	
	mocant-se)			perquè no el toco	
962	DIT: {( @) Mar}ta\		1376	mai\	
963	que et sortirà sang		1377	no\	
	del nas\			fa (soroll imitant el	
	(COC: 01, adaptat)			soroll del fre)\	
				(COC: 05, adaptat)	

(74c) 1088 REP: (.. 0.30) Marta\  
1089 (... 1.32) ara la  
Marta--  
1090 (.. 0.87) bueno\  
1091 primer esternuda\  
1092 (MMA esternuda)  
1093 REP: (. 0.24) {(F) vale\  
1094 Jesús\  
(COC: 01)

(74d) 867 MER: (.. 0.99) i escolta\  
868 (... 0.56) i vam anar  
primer a l'exposició  
de la Vir[reina]  
869 ???: [(xiulet)]  
870 MER: {(AC) vam dir  
després anirem  
a veure l'ET\  
(COC: 03)

(74e) 168 MJJ: el que és obrir\_  
169 una cervesa\_  
170 no se'm dóna bé\  
171 ara\_  
172 (soroll vocal  
imitant el de  
l'obridor)  
173 (.. 0.36) {(P)  
bueno\  
174 és igual\  
(COC: 05)

(74f) 1373 MJJ: i el freno de  
davant\_ ((de la  
bicicleta))  
1374 fa un soroll

(74g)	476	MJJ:	però no\	222	(acla- riment de gola)
	477		no_		
	478		hi ha un munt de morts eh\	223	(. 0.13) de quartels de:_
	479	PUY:	allà_	224	(. 0.12) estratègics de guerra\
	480		(esbufec)		
	481	TRS:	és_	(COC: 10, adaptat)	
	482	MJJ:	un munt de morts\		
	483	TRS:	és horrorós\	(74i) 991	NAA: jo que em volia casar la setmana que ve\ (algú tus)
		(COC: 05, adaptat)			
(74h)	219	VIE:	molts d'aquests_ ((castells))	992	
	220		(. 0.11) després d'haver matat els càtars_	993	FAN: (... 1.10) aquests els america- els americans de les pel·lícules que diu_
	221		(. 0.46) després e:n Napoleón n'havia fet servir de_	994	{{(EV) sí\
				995	ens casem\}
				(COC: 02, adaptat)	

Si el fenomen vocal és simultani amb la parla, l'anotarem mitjançant el codi *{{(descripció del fenomen) text afectat}}*, que pot incloure des de fragments de cançons, com a (75a: 816-819) —en què el codi (L2) es refereix a una segona llengua, veg. § 4.4.15.5—, fins a fragments de parla pronunciats enmig d'un badall (75b: 1182), somicant (75c: 1021), mastegant (75d: 66), sense vocalitzar (75e: 1694) o imitant el cant d'un ocell (75f: 706), entre d'altres. Com mostren (75a) i (75c), es pot combinar amb altres codis.

(75a)	814	MJJ:	a mi l'Alaska m'encanta eh\		importa lo que yo diga\}
	815	PUY:	sí:\	819	MJ?: {{(cantant)(L2)
	816		(. 0.31) {{(cantant) (L2) a quién le importa lo que yo haga\}		porta lo que yo diga\}
				(COC: 05, adaptat)	
	817	MJ?:	{{(cantant)(L2) porta lo que yo haga\}	(75b) 1177	LAU: o:stres\ 1178 mira\ ((una carrera a les mitges))
	818	PUY:	(. 0.33) {{(cantant) (L2) a quién le	1179	MMA: ha:la:\

1180 DIT: què/  
1181 REP: [{{(??) què--}}]  
1182 MMA: {(badallant) [i  
són] les noves\  
(COC: 01, adaptat)

(75c) 1018 MMA: (... 1.60) {{(??)  
txus\  
l'esternut)}  
1019 REP: (... 1.11) mira\  
1020 mira la làmpara\  
1021 MMA: {{(F)(somicant)  
que me'l  
trenque::u\  
1022 REP: Marta\  
1023 Marta\  
1024 (. 0.37) mira la  
làmpara\  
(COC: 01, adaptat)

(75d) 64 COC: tu te'n recordes  
Ramon per v on  
s'hi va a la  
carretera de xxxx/  
65 MAT: ja t'ho diré jo\  
66 (. 0.61)  
{(mastegant)  
mira\  
67 COC: és que hi vem  
anar per l'altre  
cantó:\  
(COC: 04)

(75e) 1694	ANA: {(sense vocalitzar vull magdalenes\}}	(75f) 704	TON: m'ho pregunta aixís eh\
1695	MEC: demana-ho bé\ (COC: 04, adaptat)	705	{(EV) i en Berna:t\}
		706	NAT: {(imitant el so d'una cucut) cucu:t\}
			(COC: 03, adaptat)

Recordem que no cal transcriure tots els fenòmens orals no verbals que es produeixen al llarg del fet de parla. Només anotarem els que ajudin a contextualitzar el discurs i els que siguin rellevants per al desenvolupament de la interacció —a (75e), per exemple, el codi és molt pertinent perquè MEC ordena a ANA que demani les magdalenes amb una fórmula de cortesia i pronunciant correctament l'enunciat—, d'aquesta manera evitarem donar informació innecessària i complicar massa la lectura del text. Per tant, no anotarem fenòmens de tos, esternuts, aclariments de gola, etc. que no afectin el parlant que ocupa el torn de parla o no tinguin una incidència directa en el transcurs de la interacció.

#### 4.4.12 Gestualitat

Per transcriure adequadament els recursos gestuals que s'utilitzen en l'esdeveniment comunicatiu (il·lustradors, emblemes, aspectes proxèmics, etc.) (veg. Payrató 1988: 130, McNeill 1992, 2005), ens caldrà un sistema de notació específic que ens permeti assenyalar, entre altres informacions, on comença i on acaba un gest determinat, quines parts del cos hi intervenen i com ho fan, quantes fases té el gest, quant dura cada fase, etc. El model de transcripció discursiva del CCCUB és prou flexible per adaptar-hi els sistemes de notació gestual més corrents, com mostra l'exemple (76), en què la transcripció discursiva es completa amb la notació de gestos relacionats amb la dixi d'espai a partir de la proposta de McNeill (1992), adaptada al català per Fitó Pardo (2009a) (veg. també Fitó Pardo 2009b). Les convencions que es fan servir són les següents: inici del

gest (<), acabament (>), punt àlgid (negreta), manteniment (subratllat) i descripció del moviment gestual (cursiva).

(76) 110 114: carr<er\_  
111 (**.. 0.35**) gires cap a la:  
112 dreta>  
113 (INH)

*Configuració de les mans: dits estirats cap a la dreta*

(adaptat de Fitó Pardo 2009a: 185)

Ara bé, afegir més convencions a una transcripció de nivell intermedi-estret o estret sempre té l'inconvenient d'enfardar el text i dificultar-ne la lectura. Per això aconsellem que la transcripció discursiva només incorpori notació gestual detallada quan ho exigeixi l'estudi a què la destinem i després d'haver comprovat, mitjançant una prova pilot, que la lectura no se'n ressent massa. Per tant, fora dels casos que demanen un tractament més acurat dels aspectes gestuals, el sistema de transcripció discursiva del CCCUB optarà per descriure, més que no pas transcriure, els gestos que es combinen amb el discurs oral —i sempre, és clar, que siguin rellevants per al transcurs de la interacció. Ho farem mitjançant dos procediments: d'una banda, els comentaris del transcriptor, que aniran entre parèntesis senzills o dobles segons el tipus de gest; i, d'una altra, el codi *{(descripció del fenomen) text afectat}*, emprat també per a sons orals no verbals (veg. § 4.4.11).

Inserirem un comentari entre parèntesis simples i preferentment en línia pròpia quan el gest sigui puntual i fàcil de delimitar en el temps, com ara un cop de puny en una taula o el gest de fer petar els dits. Ho il·lustren els exemples de (77): en el primer MOL, que explica a uns infants el conte del llop i les cabretes, dóna uns cops a la taula per fer veure que el llop pica a la porta, veg. (77a: 118); i en el segon, procedent d'una subhasta de peix, un comprador alça el braç per aturar el compte, veg. (77b: 1607). És possible situar aquest tipus de



comentaris darrere de grup tonal, com a (77c: 931-932), extret d'una conversa tècnica entre informàtics, però, en principi, no ha de ser la posició preferent.

(77a) 114 MOL: (.. 0.81) {(EV) i me les menjaré\_  
115 que tinc molta gana\  
116 (... 1.24) se'n va a casa de les cabretes\_  
117 i torna a picar\_  
118 (.. 0.36) (quatre cops de MOL i després uns quants dels  
nens)  
119 MOL: {(EV) qui hi ha:\  
(COR: EDUC2, adaptat)

(77b) 1602 AAC: qui més va ara\  
1603 T??: xX(3.33)Xx  
1604 JRR: en queden quatre\ ((caixes de peix))  
1605 T??: xX(3.03)Xx  
1606 JRR: quatre a la una\_  
1607 (algú aixeca el braç)  
1608 JRR: {(L2) qué nombre\  
1609 T??: xX(0.88)Xx  
1610 AA3: Pedro\  
(COR: EMP3, adaptat)

(77c) 926 XFS: (.. 0.50) amb l'únic que ho podrien muntar així més o  
menys\_  
927 (.. 0.38) ara\_  
928 perquè hi ha equips\_  
929 és amb Cisco\_  
930 però:\_  
931 (.. 0.52) tothom té la promesa\_ (pica a la taula)  
932 Bay té la promesa\_ (pica a la taula)  
933 Trescom té la promesa de veu\_  
(COR: EMP2, adaptat)

Normalment costa més delimitar els gestos, perquè molt sovint es repeteixen durant un cert període de temps i són simultanis amb la parla. En aquests casos introduïrem el comentari entre parèntesis dobles i el col·locarem preferentment al final del primer grup entonatiu amb què coincideix el moviment gestual, que pot incloure des d'aspectes proxèmics fins a configuracions de les mans, moviments del cap o de tot el cos, expressions

facials, etc. En donem una mostra a (78) amb quatre tipus de gestos diferents: de reprensió (78a: 800); per descriure el vol d'un avió (78b: 998) i la forma d'un vestit de núvia (78c: 429), i les accions d'aixecar-se i senyar-se pròpies d'una missa (78d: 130, 134).

- (78a) 797 REP: ((a ROM)) tu com és que no haes\  
 798 ROM: (.. 0.32) jo he menjat l'entrepà a les nou\  
 799 REP: (.. 0.88) {(L2) hombre\  
 800 tu també: \_ ((movent el cap com a senyal de crítica))  
 801 ROM: (.. 0.36) o me'l menjava o el llençava\  
 (COC: 01, adaptat)
- (78b) 987 TON: en lloc d'avió\_  
 988 ens van fotre una tartana\_  
 [...]  
 998 en lloc d'anar així anàvem fotent així ((simulant l'avió amb les mans))  
 999 T??: @R(0.91)R@  
 (COC: 03, adaptat)
- (78c) 423 NAA: a mi els vestits de núvia m'agraden aquests vestits així  
 tipo:\_  
 424 de hippie\_  
 425 així\_  
 426 bueno de hippie\  
 427 tipo eivissenc\_  
 428 una cosa així\  
 429 fiu fiu ((movent el braç amunt i avall))  
 (COC: 02)
- (78d) 130 CPP: posem-nos drets ((els assistents s'aixequen))  
 131 {(??) amb} l'Evangelí\_  
 132 (.. 0.53) el Senyo:r\_  
 133 (. 0.27) sigui amb vosaltres\  
 134 T??: (. 0.29) i amb el vostre esperit ((tothom se senya))  
 135 CPP: lectura de l'Evangelí:\_  
 136 segons sant Marc\  
 (COR: REL2)

Si el gest no és simultani amb la parla però té una durada indeterminada, assignarem línia pròpia al comentari. Vegeu com a exemples (79a: 1002), en què el capellà que oficia un

casament s'apropa el micròfon, i (79b: 574), en què els assistents a un casament civil aplaudeixen després del parlament de l'alcalde.

(79a) 1001 CPP: que el Senyor confirmi\_  
1002 (... 1.89) ((s'apropa el micròfon))  
1003 que el Senyor confirmi\_  
1004 el consentiment\_  
1005 que us heu manifestat\_  
(COR: REL2)

(79b) 567 ALC: (.. 0.40) avui el camí és dolç\  
568 (.. 0.30) {(P) gaudiu-lo\  
569 (.. 0.74) a tots\_  
570 de ben segur\_  
571 (.. 0.44) ens agradaria poder dir\_  
572 un sí\_  
573 (. 0.28) com el vostre\  
574 (... 11.43) ((aplaudiments))  
(COR: CIV2)

Finalment, utilitzarem el codi *{{(descripció del fenomen) text afectat}}* per indicar que un gest que podem delimitar amb precisió en el temps és simultani amb un fragment de parla. Així ho representen (80a: 1675), que recull les instruccions que dona un director de teatre a un actor; (80b: 177), en què un subhastador de peix assenyala què es vendrà a continuació, i (80c: 352), en què una informant del COS explica com jugaven de petits amb l'armament de la Guerra Civil.

(80a) 1672 ALK: (.. 0.76) fotrà un efecte bastant impressionant\_  
1673 aquest llibre aquí\_  
1674 {(AC) tot just arribis aquí davant\_  
1675 (.. 0.45) {(fent veure que estripa el llibre) rec\  
1676 (.. 0.30) r:a:st\  
1677 (. 0.29) plaf\  
(COR: CULT3)

(80b) 177 RRR: (.. 0.67) {(assenyalant amb el dit) el lluç també va\  
178 T??: xX(3.95)Xx  
179 JRR: {(AC) dos a la banda\_  
180 noranta-nou\_  
181 noranta-vuit\_

182            noranta-set\_}  
(COR: EMP3, adaptat)

(80c) 346    MAR: (.. 0.38) después jugàvem amb a: le:s bales\_  
347            (. 0.15) els hi trèiem la pólvora\_  
348            (.. 0.54) i les ficàvem dintre una olla\_  
349            li fumíem foc\_  
350            (. 0.27) i ens amagàvem\_  
351            i veies l'olla\_  
352            {(F)(gesticulant com si disparés) pam\_  
353            pam\_  
354            pam\_}  
(COS: MAR, adaptat)

Aquest tipus de codi serà especialment útil per a transcripcions a partir d'enregistraments audiovisuals, en què és relativament fàcil determinar on comença i on acaba un gest concret. En transcripcions a partir de so només, preferirem els comentaris entre parèntesis dobles darrere de grup tonal.

#### 4.4.13 Comentaris del transcriptor

En la majoria de transcripcions ens veurem obligats a incloure-hi informació relativa al context per ajudar el lector a seguir el desenvolupament de la interacció verbal, sobretot si no n'hi ha imatges enregistrades o conté molta informació pressuposada per part dels interlocutors. Aquestes informacions les introduïrem com a comentaris del transcriptor, que, com hem vist a propòsit de la notació de la gestualitat (veg. § 4.4.12), poden presentar dos formats diferents segons el seu contingut: entre parèntesis simples i línia pròpia, o entre parèntesis dobles i preferentment darrere de grup tonal.

Els comentaris referits a fenòmens contextuais que es puguin delimitar amb precisió en el temps aniran entre parèntesis simples i tindran per defecte línia pròpia, per exemple, (*cop de porta*), (*agafa el setrill*), (*sona el telèfon*). Tenen una funció principalment descriptiva i, per tant, tendiran

a recollir fenòmens que es produeixen en l'entorn en què s'inscriu la interacció verbal i que la poden condicionar. En donem una mostra a (81): a (81a: 162) el comentari serveix per identificar el soroll d'un obridor; a (81b: 866), per indicar que ANI penja el telèfon després de fer una trucada; a (81c: 1525), extret d'un assaig de teatre, per explicar què passa a l'escenari, i a (81d: 1377), per assenyalar que COT obre la porta abans de sortir de la sala on ha tingut una reunió.

(81a) 159 EUU: no saps ni obrir [una cerve]sa\  
 160 PUY: [ah sí\  
 161 (.. 0.51)  
 162 (soroll de l'obridor)  
 163 (.. 0.34)  
 164 MJJ: home\  
 165 mira\  
 (COC: 05)

(81b) 862 ANI: (... 2.39) vale\  
 863 vale\  
 864 doncs ja la trucaré\  
 865 (.. 0.36) (a)dé:u\  
 866 (penja el telèfon)  
 867 (... 1.72) {(F) ha marxat a Llore:t\  
 (COC: 06)

(81c) 1522 CJJ: bona nit\_  
 1523 pare\  
 1524 (... 1.17) em sembla que tinc son\  
 1525 (.. 0.33) (truquen a la porta i se sent algú que puja l'escala)  
 1526 IHB: {(F) Gestapo\  
 (COR: CULT3)

(81d) 1372 PSI: (.. 0.49) {(??) va\  
 1373 (.. 0.54) me'n vaig eh\  
 1374 TUT: grà[cies\  
 1375 COT: [adéu\  
 1376 PSI: (a)déu\  
 1377 (soroll de porta que s'obre)  
 1378 PSI: bueno [escolta'm\  
 1379 TUT: [bé\  
 1380 PSI: (. 0.15) ara\_  
 1381 fora de la reunió\_

(COR: EDUC1)

D'altra banda, els comentaris entre parèntesis dobles serviran per anotar fenòmens que també tenen lloc en l'escena comunicativa però són difícils de delimitar en el temps, com ara accions repetitives dels participants o sorolls de fons d'una certa durada. Aquest tipus de comentaris, de caràcter més explicatiu que els anteriors, se situaran al final de grup tonal si són simultanis amb la parla, com a (82a: 1188, 82b: 1356, 82c: 725); si no, podran ocupar línia pròpia, com ara els que recullen desplaçaments dels parlants, veg. (82d: 457, 82e: 48, 82f: 1478), accions en què no intervé la parla, veg. (83a: 23, 83b: 106) —corresponents a una missa i un míting polític respectivament—, o fragments que no es transcriuen, veg. (83c: 1), extret d'un judici.

(82a) 1188 MER: doncs a veure\ ((tallant el rostit))  
1189 [a veure què aconseguirem eh\  
1190 RAS: {(??) [treus la cuixa\_  
1191 i deixes la] paleta\  
(COC: 03)

(82b) 1353 MJJ: ja ha marxat el\_  
1354 el camió\  
1355 {(P) si us plau\  
1356 {(??) fa molta calor\} ((soroll que fa el balcó en obrir-lo))  
1357 (.. 0.62) molta calor\ (P)}  
(COC: 05)

(82c) 725 ANI: (... 1.77) qui són aquests\ ((mira una fotografia que hi ha  
damunt la taula))  
726 MAM: (... 1.40) els iaïos\  
727 ANI: (... 2.80) que guapo:s\  
(COC: 06)

(82d) 456 VIE: compris el que compris ho notes\  
457 (... 3.06) ((JON parla de lluny))  
458 JON: sí sí\  
459 no no\  
460 esclar\  
461 oh i tant\

- 462 (... 3.29) ((JON i MAG entren al menjador, JON amb la jaqueta))
- 463 JON: ja us en (a)neu/
- 464 RIA: {(P) sí\}
- (COC: 10)
- (82e) 45 BEB: (... 1.67) Anna vols cafè\  
 46 ANI: [sí\_]  
 47 BEB: [o:] desca- --  
 48 ((xoquen BEB i ANI))  
 49 ANI: ai\  
 50 perdó\  
 (COC: 60)
- (82f) 1477 PPP: (... 1.01) amb dos paquets n'hi haurà prou\  
 1478 ((MMM torna de l'habitació de CME))  
 1479 MMM: he anat a preguntar al Jaume\_  
 1480 si volia que li escalivés una poma\_  
 1481 perquè té mal de panxa\  
 (COC: 07, adaptat)
- (83a) 21 CLL: (.. 0.42) per tots els segles dels segles\  
 22 T??: amén\  
 23 (... 12.02) ((CLL seu en una butaca que hi ha al costat de l'altar; un dels assistents, OR1, s'aixeca, s'acosta a un faristol i comença a llegir))  
 24 OR1: comença\_  
 25 la segona carta de sant Pere\  
 (COR: REL3, adaptat)
- (83b) 102 JHH: (.. 0.58) des de la fraternal Euskadi\_  
 103 (.. 0.43) Begoña Lasagabaster\  
 104 (.. 0.33) {(AC)(F) candidata número u d'Eusko Alkartasuna per Guipúscoa\_  
 105 xX(5.00)Xx\  
 106 ((aplaudiments durant 19 segons, mentrestant BLL s'aixeca i puja a l'escenari))  
 107 BLL: {(L2) egun} Catalunya\  
 (COR: POL1)
- (83c) 1 (... 19.42) ((el jutge i la fiscal conversen una estona mentre esperen que entrin els testimonis i els agents judicials; hi ha soroll i no s'entén gaire el que diuen))  
 2 JUG: sí sí\  
 3 passin passin\

4           (.. 0.87) {(??) bé\  
5           se'ls ha citat a vostès per tal\_  
6           que vinguessin a aquest jutjat\_  
(COR: CIV1, adaptat)

Els comentaris entre parèntesis dobles també tenen la funció d'aportar dades que els informants pressuposen pel seu coneixement compartit del món però que el lector ignora i que, per tant, nosaltres li haurem de proporcionar perquè pugui interpretar el text. En aquest cas els comentaris aniran per defecte darrere de grup tonal. Vegeu, com a mostra, la conversa 08 del COC, en què els comentaris del transcriptor es fan imprescindibles per saber a què es refereixen els participants, que parlen de temes personals, veg. (84a: 7, 84b: 264), mentre comenten notícies que surten per televisió, veg. (84c: 758).

(84a) 6   RRR: el del:\_  
7           aparcament ens ha deixat\ ((RRR tenia un aparcament llogat))  
8   PEI: (.. 0.66) a:i\  
9           quin greu\  
(COC: 08)

(84b) 264   RRR: en Joan ja està més tranquil\ ((és germà de RRR i PEI))  
265   PEI: eh/  
266   RRR: (.. 0.31) ja està més tranquil\  
267   PEI: (.. 0.40) trobo que n:o\  
(COC: 08)

(84c) 757   RRR: ostres\  
758           tothom {( @) reps trets\} ((es refereix a la notícia que acaben de donar sobre un assassinat d'ETA))  
759           (... 1.84)  
760   RRR: on això:/  
(COC: 08)

També situarem darrere de grup tonal els comentaris que s'afegeixen per garantir l'anonimat dels informants: ((*nom fictici*)), ((*número fictici*)), ((*adreça fictícia*)), etc. (veg. § 2.2). Vegeu com a exemple el missatge de contestador automàtic de (85), en què s'han substituït les dades originals per un nom inventat —però que sempre procurarem que tingui el mateix



nombre de síl·labes que el nom real— i un número de telèfon també fictici, acompanyats del comentari corresponent.

- (85) 714 EMM: Maria Àngels\_  
715       truca sisplau a l'Ernest Cugat\_ ((nom i cognom ficticis))  
716       al quatre cinquanta-u\_ ((número fictici))  
717       trenta-cinc\_  
718       dos cinc\  
719       (. 0.22) és per la renta\  
720       (. 0.20) gràcies\  
(COR: PRIV2)

Excepcionalment, els comentaris del transcriptor entre parèntesis dobles aniran davant de grup tonal quan calgui assenyalar a qui va dirigit el fragment de parla següent —en aquest cas el contingut del comentari sempre serà el mateix, això és, ((a *INICIALS*)), veg. (86a: 1220)— o donar una informació contextual de manera immediata per evitar confusions, veg. (86b: 848).

- (86a) 1220 LAU: ((a ROM)) què has de menjar tu de fruita\  
1221 ROM: què hi ha de fruita\  
1222 LAU: (.. 0.68) {(F) pe:ra\  
1223       po:ma\  
(COC: 01, adaptat)

- (86b) 848 ANI: (... 1.15) ((al telèfon)) que hi ha la Marta/  
849       (... 1.49) [de l'Anna\_  
850 MAM: [vosaltres] a quina hora voleu anar-hi\  
(COC: 06)

En resum, utilitzarem els comentaris del transcriptor entre parèntesis simples i en línia pròpia per descriure fenòmens contextuais ben delimitats en el temps que tinguin lloc durant l'esdeveniment comunicatiu, mentre que reservarem els comentaris entre parèntesis dobles per explicar fenòmens de durada indeterminada i donar informacions complementàries que ajudin a interpretar la transcripció. Cal recordar, però, que els comentaris del transcriptor s'han d'administrar amb certa

moderació, perquè posar-ne massa alenteix la lectura del text i, per tant, només n'introduïrem si ajuden el lector a seguir el transcurs del fet de parla o aporten dades rellevants per a l'anàlisi de la transcripció.

#### 4.4.14 Regularitzacions

Sovint trobarem solucions lingüístiques de tipus fonètic, morfològic, lèxic, etc. que no es corresponen amb les formes ortogràfiques estàndards de la llengua (per exemple, les pronúncies *dicidir* per *decidir* i *allavonses* per *llavors*). En aquest manual anomenarem *regularització* el procés pel qual adaptem aquestes formes no estàndards a les convencions ortogràfiques generals (escrivint, per exemple, *dicidir* i *allavonses*) en funció dels objectius de cada transcripció. Si, per exemple, el propòsit de l'estudi és analitzar sintàcticament les oracions de relatiu en la llengua col·loquial, segurament no caldrà que la transcripció reculli fenòmens d'assimilació fonètica com el que trobem a *dicidir*. En canvi, si l'objectiu és detectar marques lingüístiques relacionades amb les diferents varietats socials de la llengua, sí que serà pertinent distingir, posem per cas, les diferents pronúncies del connector *però* (cf. *però*, *pro*, *pero*). Per tant, abans de començar a transcriure caldrà preguntar-se fins a quin punt haurem de regularitzar les formes no estàndards que apareguin en la transcripció, sempre tenint en compte que com menys s'alterin les convencions ortogràfiques habituals més fàcils seran les cerques posteriors (veg. Duranti 1997: 193). En general, tenim quatre opcions a l'hora de regularitzar una transcripció, tal com exemplifiquem tot seguit a partir del mateix fragment de parla, extret del COS:

1. No regularitzar les formes no estàndards. Això implica transcriure respectant al màxim les convencions ortogràfiques i, per exemple, escriure *decidir* tant si el parlant ha assimilat la primera vocal com si no. Podem considerar que aquesta opció falseja la realitat, perquè només reflecteix parcialment el que ha tingut lloc en la interacció verbal. Tot i així, hem de recordar que

tota transcripció sempre serà, per força, parcial i incompleta i que caldrà triar la informació d'acord amb la finalitat de l'estudi coneguent. A més, una transcripció que segueix fidelment les convencions ortogràfiques sempre és més fàcil de difondre i permet una cerca d'ítems, sigui automàtica o manual, molt més ràpida. Per això és el procediment que solen adoptar els sistemes de notació estàndards internacionals, com ara el TEI (veg. § 3.5). Vegeu que a (87) no s'han regularitzat les formes lingüístiques originals i s'han transcrit amb la grafia estàndard corresponent, encara que s'allunyi de la pronúncia real:

- (87) 430 VDC: (... 1.12) però això no vol dir que siguis feliç\_  
 431 {(AC) perquè de problemes en tindràs sempre\_  
 432 tant per la dreta\_  
 433 com per l'esquerra\_  
 434 (... 1.20) !:a felicitat\_  
 435 no existeix\  
 (COS: VDC, adaptat)

2. No regularitzar les formes no estàndards, però incloure en el text una transcripció fonètica o regularitzada dels mots més significatius. Aquesta segona opció consisteix a acompanyar certes formes en grafia estàndard d'una transcripció fonètica o una transcripció ortogràfica adaptada que s'aproximi a la pronúncia original, de manera que reflectim més fidelment la realitat i alhora garantim una cerca ràpida. El gran desavantatge d'aquest procediment és que fa molt més feixuga la lectura de la transcripció. Per aquesta raó els sistemes que opten per aquesta fórmula solen situar les formes regularitzades al marge de la transcripció, com ara el model desenvolupat per Blanche-Benveniste i Jeanjean (1987), que les anota a peu de pàgina. Vegeu que a (88) s'ha completat el text de (87) afegint transcripcions regularitzades (entre parèntesis triangulars) allí on la forma original s'aparta de l'estàndard. Ara tenim totes dues solucions, les estàndards i les no estàndards, però, sens dubte, la lectura del text se'n ressent.

- (88) 430 VDC: (... 1.12) però <pero> això no vol dir que siguis feliç\_  
 431 {(AC) perquè de problemes <problemes> en tindràs

sempre\_}  
 432 tant per la dreta <drea>\_  
 433 com per <por> l'esquerra\_  
 434 (... 1.20) l:a felicitat <felicità>\_  
 435 no existeix <existeixa>\

(COS: VDC, adaptat)

Una altra possibilitat és presentar en la mateixa pàgina una transcripció en grafia estàndard combinada amb una transcripció fonètica completa o parcial (disposant-les en columnes paral·leles o anotant sota cada grup tonal la transcripció fonètica respectiva), però, evidentment, això exigeix un esforç en temps i mitjans considerable, i no sempre necessari.

3. Regularitzar parcialment les formes no estàndards. En aquest cas l'opció per defecte és emprar l'ortografia convencional excepte en aquelles formes en què sigui pertinent aproximar-se a la pronúncia original. Aquesta solució, adoptada sovint per models dissenyats per transcriure llengua col·loquial, com ara el sistema Valesco (veg. Briz 2002), permet reproduir amb més fidelitat algunes formes lingüístiques no estàndards, però demana dues feines complementàries. La primera, establir uns criteris clars de regularització i determinar d'entrada què adaptarem i què no. Hem de tenir present que, un cop iniciada la transcripció, resultarà molt costós canviar un criteri i tornar a revisar el text des del principi, sobretot si treballem amb corpus extensos. A més a més, ens haurem de plantejar com alterem les convencions ortogràfiques generals per no dificultar gaire la lectura del text. Per exemple, la pronúncia *dos* per *doncs*, freqüent en llengua col·loquial, pot esdevenir ambigua en certs contextos, perquè el lector pot interpretar que es refereix al numeral i no pas al connector. Com veurem més endavant, una possible solució és modificar la forma estàndard del mot però garantint que es pugui reconèixer, per exemple, fent servir parèntesis: *do(nc)s*. La segona tasca serà elaborar un document o una base de dades que reculli totes les formes regularitzades de la transcripció per garantir que les cerques posteriors també recuperaran les solucions no estàndards.

A (89) regularitzem, en part, el fragment de (87). Col·loquem entre parèntesis els sons elidits i adaptem ortogràficament les formes que mostren canvis morfològics, però mantenim en grafia estàndard les que només presenten canvis d'accentuació (*però*, i no *pero*).

(89) 430 VDC: (... 1.12) però això no vol dir que siguis feliç\_  
 431 {(AC) perquè de p(r)oblemes en tindràs sempre\_  
 432 tant per la dre(t)a\_  
 433 com por l'esquerra\_  
 434 (... 1.20) l:a felicità(t)\_  
 435 no existeixa\  
 (COS: VDC)

4. Regularitzar totes les formes no estàndards. Lògicament, aquesta quarta opció permet aproximar-se molt més a les solucions lingüístiques reals, però com a contrapartida altera de manera important les convencions ortogràfiques i pot comprometre seriosament tant la lectura del text com les cerques posteriors. Vegeu (90), en què s'han adaptat ortogràficament totes les pronúncies no estàndards que utilitza l'informant.

(90) 430 VDC: (... 1.12) pero això no vol dir que siguis feliç\_  
 431 {(AC) perquè de p(ro)blemes en tindràs sempre\_  
 432 tant per la drea\_  
 433 com por l'esquerra\_  
 434 (... 1.20) l:a felicità\_  
 435 no existeixa\  
 (COS: VDC, adaptat)

Per norma general, el sistema de transcripció discursiva del CCCUB segueix la tercera opció. Malgrat les dificultats que, com hem vist, implica aquest procediment, també sembla que és una bona manera d'aconseguir un equilibri raonable entre, d'una banda, la fidelitat a la parla natural i, d'una altra, la voluntat d'oferir als investigadors transcripcions àgils de llegir i aptes per a les cerques. Cal recordar que la principal finalitat de les transcripcions discursives del CCCUB és proporcionar a

investigadors, docents i altres professionals de la llengua textos neutres que es puguin fer més amples o més estrets en funció del propòsit de l'estudi o la tasca que es vulgui desenvolupar. Per això des del principi es va optar per un sistema de regularització que, sense ser massa exhaustiu, recollís els trets més característics de la llengua oral.

En l'apartat següent detallarem els criteris de regularització que aplica el CCCUB en la transcripció discursiva de nivell intermedi-estret. La majoria de convencions que presentem es basen en les solucions pròpies de la varietat oriental central, que és la que fins ara s'ha utilitzat més en aquest tipus de transcripcions, però es poden adaptar a qualsevol altra modalitat geogràfica, funcional i/o social, tal com ja s'ha fet amb la varietat nord-occidental en el Corpus Oral de Registres de la Universitat de Lleida (veg. Universitat de Lleida).

#### 4.4.15 Regularitzacions del CCCUB

La regularització de formes no estàndards en les transcripcions discursives del CCCUB segueix el que podríem denominar el *principi de predictibilitat*: com més predictable és una solució lingüística, menys tendirà a ser regularitzada. D'acord amb això, la transcripció només adaptarà ortogràficament aquelles formes més marcades, menys previsible, per aproximar-se al discurs oral sense alterar gaire les convencions ortogràfiques ni entorpir la recuperació d'informació. Així, per exemple, davant les diferents possibilitats de transcriure el mot *pressió* pronunciat amb [-z-] —*pressió*, *presió* o *pres(s)ió*—, triarem per defecte *pressió*, perquè és previsible que en molts contextos la pronúncia habitual sigui amb essa sonora. Per contra, la dificultat de predir l'elisió de [-n-] en un verb com *transformar* afavorirà que el mot sigui regularitzat —mitjançant parèntesis, *tra(n)sformar*, veg. § 4.4.15.1— en cas d'omissió de la consonant nasal. Per tant, en un nivell de transcripció intermedi-estret tindrem com a criteri prioritari evitar la regularització de les formes no

estàndards més predictibles i, en cas de dubte, sempre triarem com a solució preferent la que no impliqui cap alteració de les convencions ortogràfiques. Excepcionalment, però, hi ha certs casos que fan recomanable la regularització. Ho resumim amb els dos supòsits següents:

1. Quan considerem que la grafia convencional no reflecteix de manera prou satisfactòria una forma no estàndard representativa de la varietat geogràfica, funcional o social en què apareix. Per exemple, la grafia de *text*, *complex* i *reflex* no recull l'epèntesi vocàlica final que sovint adjunten els mots acabats en so africàtic, però tenint en compte que es tracta d'un fenomen fonètic molt característic de la llengua oral, tendirem a regularitzar-lo afegint-hi, quan escaigui, una -e (*texte*, *complexe*, *reflexe*). Farem el mateix en la transcripció de formes degudes a errors o lapsus de l'informant (per exemple, *estigmantització* per *estigmatització*).

2. Quan la grafia convencional s'aparta sistemàticament d'una realització no estàndard. Com a exemples podem esmentar els numerals *quaranta* i *seixanta*, que en català oriental rarament mantenen els diftongs [-••-] i [-•••-], cosa que afavorirà que regularitzem tots dos mots amb les solucions *coranta* i *sixanta*, que no malmeten gaire la grafia convencional.

Sovint, però, haurem de vigilar que la regularització no deformi massa l'ortografia per garantir que el lector pugui reconèixer fàcilment el mot en qüestió; en cas contrari, mantindrem la forma estàndard. Per exemple, regularitzar la dissimilació gairebé sistemàtica de [-n-] a *bona nit* dona com a resultat una alteració ortogràfica que considerem massa brusca (cf. *bora nit*) i, per tant, tendirem a evitar-la no adaptant l'expressió.

En resum, tot i que el principi general serà respectar al màxim les convencions ortogràfiques, hi haurà mots que sí que regularitzarem, bé perquè són marques identificatives d'una varietat lingüística, o bé perquè la grafia estàndard ja no en

reflecteix la forma més habitual i previsible. Com hem apuntat més amunt, això ens obligarà a acompanyar la transcripció o tot el corpus d'un inventari actualitzat de regularitzacions per preveure els ítems que caldrà buscar en fer les cerques.

#### 4.4.15.1 Ús de parèntesis

Els parèntesis seran un dels recursos gràfics que més utilitzarem per regularitzar solucions no estàndards. Bàsicament, els farem servir per indicar que s'ha elidit material fonològic, ja sigui per raons fonètiques, com ara la caiguda de vocal àtona entre consonant oclusiva i líquida pròpia del català oriental (*t(a)ronja*, *g(a)ratge*, *T(e)resa*), o per errors i lapsus de pronúncia (*dete(r)gent*, *aju(n)tament*). Aquest signe de puntuació té força avantatges. D'entrada, un dels significats que té fora de la transcripció discursiva ja és, precisament, el de l'elisió *i*, per tant, utilitzar-lo amb aquesta finalitat no suposa atribuir-li un ús completament nou que pugui confondre el lector. A més a més, no deforma gaire la grafia convencional i permet reconèixer el mot; compareu, per exemple, les solucions *t(a)ronja*, *g(a)ratge* i *T(e)resa* amb les formes corresponents amb la vocal suprimida *tronja*, *graig* i *Tresa*, molt més difícils d'identificar.

Malgrat tot, procurarem que la transcripció no estigui massa atapeïda de parèntesis per fer més àgil la lectura del text, sobretot en transcripcions de conversa col·loquial, que ja solen incorporar un gran nombre de claudàtors, claus i altres símbols. Per això evitarem els parèntesis en pronúncies molt previsibles, i especialment en casos de fonètica sintàctica com els que recull la taula 4.4.

Formes recomanables	Formes no recomanables
<i>i ja està!</i>	<i>i ja (e)stà!</i>
<i>vaig (a)nar a esquiar</i>	<i>vaig (a)nar (a) (e)squiar</i>
<i>m'he aixecat</i>	<i>m'he (ai)xecat</i>



<i>això li va passar a aquell noi</i>
---------------------------------------

<i>això li va passar (a) aquell noi</i>
---

Taula 4.4 Ús de parèntesis en contextos de fonètica sintàctica

Com es pot veure, gairebé totes les formes incloses a la taula 4.4 poden prescindir dels parèntesis perquè és fàcil preveure-hi elisions vocàliques per fonètica sintàctica, sobretot si l'elocució és ràpida. L'única excepció és *Vaig (a)nar a esquiar*, en què no es produeix un fenomen de fonètica sintàctica pròpiament dit, sinó una elisió molt freqüent —però no sistemàtica— que afecta la vocal inicial del verb *anar*. En aquest cas l'ús de parèntesis ens estalvia haver de recórrer a una forma no estàndard que seria difícil de reconèixer (cf. *vaig nar a esquiar*).

En una transcripció discursiva de nivell intermedi-estret també utilitzarem parèntesis, si escau, en els contextos que esmentem a comunicació:

1. Reducció de grups consonàntics: *pro(b)lema*, *p(r)oblema*, *p(r)o(b)lema*; *p(r)ograma*; *co(n)stipat*, *tra(n)satlàntic*; *i(n)stitut*; *tra(n)sparent*; *escri(p)tor*, *por(t)*; *mor(t)*; *San(t) Andreu*. En alguns casos, però, preferirem una altra grafia per reproduir més fidelment la pronúncia real (*organisar* per *organitzar*, *tranquil·lisar* per *tranquil·litzar*, *hospitalisar* per *hospitalitzar*, *contar* per *comptar*) o evitar l'excés de parèntesis (*atre*, *-a*, *-es* i *antre*, *-a*, *-es* per *altre*, *-a*, *-es*).

2. Elisió d'oclusives en posició intervocàlica (*mo(v)iment*), incloent-hi la caiguda de [-•-] en valencià (*vega(d)a*).

3. Elisió de vocal neutra entre oclusiva i líquida: *t(a)ronja*, *b(e)renar*, *T(e)resa*, *v(e)ritat*.

4. Elisió de [-b] en la preposició *amb* en contacte amb vocal inicial de mot: *am(b) ell*. Però transcriurem *an*, *am* o *amb* quan aquestes formes corresponguin a les preposicions *a* i *en* regides per un verb: *l'hi va dir a/an/am/amb ell*; *es converteix en/am/amb algú desconegut*.

5. Elisió de [-s-] en el demostratiu masculí singular *aquest* en contacte amb vocal inicial de mot (*aque(s)t home*) i en les formes femenines *aque(s)ta* i *aque(s)tes*. Si en el dialecte de l'informant l'elisió és sistemàtica, tenim l'opció de prescindir dels parèntesis (*aqueta*, *aquetes*) per evitar la profusió d'aquesta grafia.

6. Elisions vocàliques en els verbs *anar* i *haver* no degudes a fonètica sintàctica; és a dir, no regularitzarem els enunciats *Hi havia molta gent* i *Hi anava molta gent* —pronunciats *Hi via molta gent* i *Hi nava molta gent*—, però sí que marcarem entre parèntesis les elisions en enunciats com *Ell (ha)via de venir ahir*, *(Ha)guessis vingut més aviat* i *(A)nar-hi ara no val la pena*, perquè la caiguda de la vocal no és causada per fonètica sintàctica.

7. Elisions en els connectors *per (ai)xò*, *do(nc)s*, *p(e)rò* i *per (e)xemple*; en el marcador discursiu *b(u)eno* i en les formes *(a)déu*, *per (a)quí*, *per (a)llà* i *mi(ra)-te'l*.

No distingirem amb parèntesis els pronoms *res* i *re*, tots dos recollits en obres de referència.

#### 4.4.15.2 Fonètica

A part de l'elisió, hi ha altres fenòmens fonètics que també regularitzarem. A sota esmentem els més freqüents:

1. *Epèntesi*. Representarem l'addició d'una vocal epentètica mitjançant la grafia *e*, que s'adjuntarà al mot que reforça: *texte*, *complexe*, *reflexe*, *fixe*. En contextos pronominals, però, anirà entre espais (*ens e l'ha dut ell*) o lligada al pronom mitjançant guionet (*di'ns-e*) (veg. § 4.4.15.3). Les epèntesis consonàntiques (*v*, *g*, *t*, *s*) també s'uniran al mot (*cort* per *cor*, *cart* per *car*, *mart* per *mar*, *aixís* per *així*), i si les normes

ortogràfiques ho requereixen, modificarem l'accentuació del mot després d'afegir-hi l'epèntesi: *col·lègit, prèmit*.

Un cas particular és el pronom *on*, que incorpora sovint epèntesis consonàntiques i vocàliques. Si l'epèntesi és consonàntica i posterior, l'ajuntarem al pronom (*ont és?*), i si, a més, s'hi afegeix una vocal, hi haurà un espai entre aquesta i la resta del mot (*ont e vas?*). En canvi, quan l'epèntesi és anterior, la separarem del pronom amb un espai (*a v on vas?*, *a g on vas?*, *a u on vas?*).

Igualment, considerem epèntesis les formes que especifiquem tot seguit:

a) La forma *en* davant dels adverbis *aquí* i *allà*: *en aquí no ho trobaràs, en allà és més barat*.

b) Les formes *an*, *am*, *amb* i *ambe* com a reforç de preposicions rígides (*li he dit a/an/am/amb/ambe un amic; viu en/am/amb un pis de Mataró; no confia en/am/amb ell; està en/am/amb estat*) i de tematitzacions (*an/am/amb el Pere l'hi vaig dir jo*). Una altra variant és *ane*, que apareix sovint davant pronoms: *ane qui ho ha dit?, m'ho ha dit ane mi*.

c) Les formes *am* i *amb* en lloc de la preposició *en* en contextos com *en/am/amb aquestes que...* i *en/am/amb això que....*

2. *Accentuació*. No reflectirem el canvi de timbre dels accents respecte a les pronúncies estàndards i, per exemple, transcriurem *ciència* encara que s'hagi pronunciat *ciéncia*. Seguirem el mateix criteri en relació amb la conjunció àtona *que* —a vegades articulada amb *e* oberta o tancada en català oriental— i el pronom *què*. Així doncs, mantindrem *que* en casos de no caiguda de preposició (*es queixa de que fa molta calor*) i en usos exclamatius (*que fort!*) i *què* en oracions interrogatives i de relatiu. Les úniques excepcions seran el *que* substantivat en enunciats com ara *El què?* i *Ja et diré el què* i la interjecció d'origen castellà *què va!*

Sí que indicarem els canvis de posició de l'accent respecte a les formes estàndards (*atmòsfera, fútbol, canvia, hòquei, estúdia, denúncia*) i opcionalment també reflectirem el canvi de

posició accentual en els connectors *perquè* (*perque*) i *però* (*pero*).

3. *Diftongs i hiats*. En general, tendirem a reproduir les reduccions de diftongs: *mitat*, *sixanta*, *coranta*, *aiga*, *aigo*, *llenga*, *con* per *quan*, *custió* per *qüestió*, *casi*, *cantitat*, *calitat*, *sisplau*. Tot i així, si considerem que la forma resultant altera massa la grafia convencional del mot, farem servir parèntesis, cf. *ve(u)ràs*, *v(e)uràs*, *a ve(u)re*, *ja v(e)urem*; (val.) *vore*, *voràs*, *vorem*; *b(u)eno*; *ll(e)uger*. També adaptarem la diftongació de vocal inicial de mot (*aufegar*) i els sons antihiatícs (*ideia*, *paiella*, *agraiexo*, *dugues*).

4. *Assimilacions i dissimilacions*. Regularitzarem tant les assimilacions (*istiu*, *dicidir*, *salxitxa*, *xíndria*, *xixanta*, *jonoll*, *istil*) com les dissimilacions (*vegilar*, *dingú*, *altre/antre*, *comanió*, *fandilla*), excepte en aquells casos sistemàtics en què considerem que s'altera massa la grafia convencional del mot (*bona nit*, no *bora nit*; *juliol*, no *juriol*).

5. *Altres fenòmens fonètics*. No reflectirem ortogràficament els fenòmens de ieisme etimològic (*ull*, no *ui*; *palla*, no *paia*) o no etimològic (*projecte*, no *proiecte*) ni l'africació a l'inici de mot (*xocolata*, no *txocolata*; *xiclet*, no *txiclet*), però regularitzarem la palatalització (*llògica*, *lliteratura*, *fenya*, *cunya*) i la metàtesi (*radere*).

#### 4.4.15.3 *Morfologia*

A continuació descrivim alguns dels contextos morfològics que solen comportar processos de regularització:

1. *Article*. No apostrofarem els articles singulars *el* i *la* si no s'elideix la vocal (*el home*, *la època*) i en contacte amb la preposició *per a* transcriurem per defecte *pel*, *per la*, *pels* i *per les*, excepte en aquells casos en què el parlant opti per les

formes *per al*, *per a la*, *per als* i *per a les*. I, si cal, apostrofarem l'article femení singular davant de *i/u* àtones: *l'història*, *l'universitat*.

2. *Substantius i adjectius*. Regularitzarem les formes analògiques de gènere i nombre de substantius i adjectius (*eleganta*, *igualta*, *capaça*, *llàpissos*, *dillunsos*), així com totes les variants geogràfiques que puguin presentar (*hòmens*). Els substantius acabats amb la grafia *-us* pronunciats amb [-u] els adaptarem amb *-o*: *globo*, *tipo*, *foco*.

3. *Demonstratius, possessius i indefinits*. Si escau, transcriurem *aquestos*; *nostro*, -s, *vostro*, -s i les formes analògiques *bastanta*, -es; *masses*; *prous*, *prouta* i *proutes*, i *forces*.

4. *Morfologia verbal*. Adaptarem ortogràficament totes les variants verbals no estàndards, i en cas d'elisió farem servir parèntesis com a opció preferent. A la taula 4.5 recollim algunes de les regularitzacions més habituals en català oriental (central):

<b>Formes no personals</b>
Infinitiu: <i>haguer</i> , ( <i>ha</i> ) <i>ver</i> , <i>poguer</i> , <i>quebre</i> , <i>sapiguer</i> , <i>tindre</i> , <i>ve(u)re</i> , <i>vindre</i>
Gerundi: <i>creguent</i> , <i>diguent</i> , <i>duguent</i> , <i>estiguent</i> , <i>haguent</i> , <i>nasquent</i> , <i>sapiguent</i> , <i>siguent</i> , <i>tenguent</i> , <i>tinguent</i> , <i>traquent</i> , <i>vegent</i> , <i>veguent</i> , <i>vinguent</i>
Participi: <i>aparescut</i> , <i>capigut</i> , <i>imprimit</i> , <i>molgut</i> , <i>sapigut</i> , <i>set</i>
<b>Formes personals</b>
<i>Anar</i> : pres. ind. <i>vai(g)</i> , <i>vaic</i> ; perf. ind. <i>vai(g)</i> (p. ex. <i>vai(g) anar-hi tard</i> ), <i>vàreig</i> , <i>vem</i> , <i>veu</i>
<i>Cabre</i> : pres. ind. <i>quepo</i> , <i>quebo</i> ; pres. subj. <i>cabi</i> , <i>cabis</i> ...
<i>Deure</i> : fut. <i>d(e)uré</i> , <i>d(e)uràs</i> , <i>d(e)urà</i> ...
<i>Donar</i> : pres. ind. <i>dón(a)</i> (p. ex. <i>no em dón(a) la gana de fer-ho</i> ), <i>dón(e)s</i> (p. ex. <i>em dón(e)s un caramel?</i> ); pres. subj. <i>dongui</i> , <i>donguis</i> ...; imper. <i>dón(a)</i> (p. ex. <i>dón(a) un caramel al teu germà</i> )
<i>Esperar</i> : imper. ( <i>es</i> ) <i>pera</i> (p. ex. ( <i>es</i> ) <i>pera</i> , <i>que vinc!</i> )
<i>Ésser</i> : imp. ind. <i>érets</i>
<i>Estar</i> : imp. subj. <i>estés</i> , <i>estessis</i> ...
<i>Fer</i> : pres. ind. <i>fai(g)</i> , <i>faic</i> ; pres. subj. <i>fagi</i> , <i>fagis</i> ...
<i>Haver</i> : indef. <i>hai</i> , <i>hi</i> (p. ex. <i>ja hi dinat</i> ); plusq. ind. ( <i>ha</i> ) <i>via</i> , ( <i>ha</i> ) <i>vies</i> ...; imp. subj. ( <i>ha</i> ) <i>gués</i> , ( <i>ha</i> ) <i>guessis</i> ..., ( <i>ha</i> ) <i>vés</i> , ( <i>ha</i> ) <i>vessis</i> ...; perífrasi d'obligació

<i>hai(g)</i> (p. ex. <i>hai(g) d'anar-hi</i> ).
<i>Morir</i> : pres. ind. <i>et mores, es mora</i>
<i>Saber</i> : pres. ind. <i>sa(p)s, sats</i> ; imp. subj. <i>sapigués, sapiguessis...</i>
<i>Veure</i> : pres. ind. <i>vei(g), vec, veic</i> ; fut. ind. <i>v(e)uré, ve(u)ràs...</i>

Taula 4.5 Regularització de formes verbals

El Corpus Oral de Registres de la Universitat de Lleida regularitza les formes verbals de la varietat nord-occidental adaptant-les ortogràficament (*ell cante, jo podie, conduíam, portéssam, vingos*) o fent servir parèntesis (*porta(v)en, calenta(v)e*) (veg. Universitat de Lleida).

5. *Pronoms personals*. Regularitzarem totes les variants dels pronoms personals tòpics *nosaltres* i *vosaltres*: *nosaltros/vosaltros, nosatres/vosatres, nosatros/vosatros, naltres/ valtres, naltros/valtros*, etc. Pel que fa als pronoms febles, seguirem les grafies i les combinacions estàndards, incloent-hi les convencions que se solen emprar en registres informals, com ara *l'hi* per *la hi* i *li ho* (vegeu-ne una mostra als portals lingüístics Optimot i Ésadir). Si la forma estàndard no representa adequadament la pronúncia, transcriurem els pronoms febles evitant els parèntesis i potenciant l'ús del guionet, tal com proposem en els subapartats següents:

a) *Formes reforçades i elidides*. En els pronoms de 1a, 2a i 3a pers. pl. en forma reforçada seguits de vocal epentètica deixarem un espai entre el clíctic i el suport vocàlic.

Pronoms febles en forma reforçada (grafia estàndard)	Pronoms febles en forma reforçada (grafia regularitzada)
<i>ens sentia</i>	<i>ens e sentia</i>
<i>us sentia</i>	<i>us e sentia</i>
<i>els sentia</i>	<i>els e sentia</i>
<i>les sentia</i>	<i>les e sentia</i>

Taula 4.6 Regularització de pronoms febles en forma reforçada

Considerarem epentètica la vocal neutra que trobem entre els pronoms de 1a, 2a i 3a pers. pl. i els pronoms en forma elidida davant de verb iniciat en vocal.

Pronoms febles en forma reforçada davant de pronom en forma elidida (grafia estàndard)	Pronoms febles en forma reforçada davant de pronom en forma elidida (grafia regularitzada)
<i>ens n'anem</i>	<i>ens e n'anem</i>
<i>us n'aneu</i>	<i>us e n'aneu</i>
<i>ens n'ha parlat</i>	<i>ens e n'ha parlat</i>
<i>us n'ha parlat</i>	<i>us e n'ha parlat</i>
<i>ens l'ha comprat</i>	<i>ens e l'ha comprat</i>
<i>us l'ha comprat</i>	<i>us e l'ha comprat</i>
<i>els l'ha comprat</i>	<i>els e l'ha comprat</i>

Taula 4.7 Regularització de pronoms febles en forma reforçada i forma elidida

Tractarem com a formes reforçades els clítics que trobem darrere la 2a pers. sing. de l'imperatiu acabat en consonant —així no alterem gaire la forma verbal, cf. *ajupe't*— i també en gerundis com ara *mirant-el*.

Pronoms febles darrere de 2a pers. sing. d'imperatiu acabat en consonant (forma estàndard)	Pronoms febles darrere de 2a pers. sing. d'imperatiu acabat en consonant (forma regularitzada)
<i>segueix-me</i>	<i>segueix-em</i>
<i>ajup-te</i>	<i>ajup-et</i>
<i>fot-te</i>	<i>fot-et</i>
<i>segueix-lo</i>	<i>segueix-el</i>
<i>segueix-los</i>	<i>segueix-els</i>
<i>segueix-nos</i>	<i>segueix-ens</i>
<i>serveix-ne</i>	<i>serveix-en</i>

Taula 4.8 Regularització de pronoms febles darrere imperatiu

Només alterarem l'ortografia del verb en casos en què la forma resultant no reflecteixi adequadament la pronúncia (*rebe'ls* i no pas *rep-els* o *reb-els*).

b) *Formes plenes*. Entre la 2a pers. sing. de l'imperatiu i les formes pronominals plenes afegirem, si escau, una vocal epentètica entre guionets.

<b>Pronoms febles en forma plena darrere de 2a pers. sing. d'imperatiu acabat en consonant (forma estàndard)</b>	<b>Pronoms febles en forma plena darrere de 2a pers. sing. d'imperatiu acabat en consonant (forma regularitzada)</b>
<i>coneix-la</i>	<i>coneix-e-la</i>
<i>cull-la</i>	<i>cull-e-la</i>
<i>coneix-les</i>	<i>coneix-e-les</i>
<i>cull-les</i>	<i>cull-e-les</i>
<i>cull-ho</i>	<i>cull-e-ho</i>
<i>promet-li</i>	<i>promet-e-li</i>

Taula 4.9 Regularització de pronoms febles darrere imperatiu amb epètesi

Si el verb en imperatiu elideix la [-s] final, evitarem els parèntesis (cf. *digue(s)-li*) i optarem per grafies com les de la taula 4.10, amb tradició escrita (veg. Optimot, Ésadir) (veg. també la taula 4.12).

<b>Pronoms febles en forma plena darrere de 2a pers. sing. d'imperatiu acabat en -s (forma estàndard)</b>	<b>Pronoms febles en forma plena darrere de 2a pers. sing. d'imperatiu amb -s elidida (forma regularitzada)</b>
<i>digues-la</i>	<i>digue-la</i>
<i>digues-les</i>	<i>digue-les</i>
<i>digues-ho</i>	<i>digue-ho</i>
<i>digues-li</i>	<i>digue-li</i>

Taula 4.10 Regularització de pronoms febles darrere imperatiu amb -s elidida

c) *Formes reduïdes*. Considerarem reduïdes a 's les formes -nos i -vos de 1a i 2a pers. pl. darrere imperatiu i les separarem mitjançant guionet de la vocal epentètica que incorporen. També farem servir el guionet quan es combinin amb altres clítics.

<b>Els pronoms febles -nos i -vos darrere imperatiu (grafia estàndard)</b>	<b>Els pronoms febles -nos i -vos darrere imperatiu (grafia regularitzada)</b>
<i>aixequem-nos</i>	<i>aixequem's-e</i>
<i>aixequen-vos</i>	<i>aixequen's-e</i>
<i>anem-nos-en</i>	<i>anem's-en</i>
<i>aneu-vos-en</i>	<i>aneu's-en</i>



<i>comprem-nos-el</i>	<i>comprem's-el</i>
<i>comprem-nos-la</i>	<i>comprem's-e-la</i>
<i>comprem-nos-ho</i>	<i>comprem's-ho</i>
<i>comprem-nos-en</i>	<i>comprem's-en</i>
<i>compreu-vos-el</i>	<i>compreu's-el</i>
<i>compreu-vos-la</i>	<i>compreu's-e-la</i>
<i>compreu-vos-ho</i>	<i>compreu's-ho</i>
<i>compreu-vos-en</i>	<i>compreu's-en</i>
<i>arribem-nos-hi</i>	<i>arribem's-hi</i>
<i>arribeu-vos-hi</i>	<i>arribeu's-hi</i>

Taula 4.11 Regularització dels pronoms febles -nos i -vos darrere imperatiu

L'elisió de -s en la 2a pers. sing. d'imperatiu pot donar lloc a formes reduïdes (veg. també taula 4.10). Si l'imperatiu no presenta cap tipus d'elisió, obtindrem formes com *dóna'ls-hi*, *dóna'ns-e-la*, etc.

<b>Pronoms febles en forma plena darrere de 2a pers. sing. d'imperatiu acabat en -s (forma estàndard)</b>	<b>Pronoms febles en forma reduïda darrere de 2a pers. sing. d'imperatiu amb -s elidida (forma regularitzada)</b>
<i>digues-me</i>	<i>digue'm</i>
<i>estigues-te</i>	<i>estigue't</i>
<i>digues-lo</i>	<i>digue'l</i>
<i>digues-nos</i>	<i>digue'ns</i>
<i>digues-los</i>	<i>digue'ls</i>
<i>digues-ne</i>	<i>digue'n</i>

Taula 4.12 Regularització de formes reduïdes darrere imperatiu

Transcriurem els infinitius que elideixen la -r final davant de pronom feble sense aquesta consonant i el clíctic en forma reduïda o plena (veg. taula 4.13), i, si hi ha epèntesi, la separarem dels pronoms mitjançant guionet (veg. taula 4.14). Quan calgui, modificarem l'accentuació del verb d'acord amb les normes ortogràfiques.

<b>Infinitiu amb pronom feble (forma estàndard)</b>	<b>Infinitiu amb -r elidida i pronom feble (forma regularitzada)</b>
<i>conèixer-me</i>	<i>coneixe'm</i>
<i>conèixer-te</i>	<i>coneixe't</i>

<i>conèixer-nos</i>	<i>coneixe'ns</i>
<i>conèixer-vos</i>	<i>coneixe-us</i>
<i>conèixer-lo</i>	<i>coneixe'l</i>
<i>conèixer-los</i>	<i>coneixe'ls</i>
<i>conèixer-la</i>	<i>coneixe-la</i>
<i>conèixer-les</i>	<i>coneixe-les</i>
<i>conèixer-li</i>	<i>coneixe-li</i>
<i>conèixer-se</i>	<i>coneixe's</i>
<i>conèixer-ho</i>	<i>coneixe-ho</i>
<i>conèixer-hi</i>	<i>coneixe-hi</i>
<i>conèixer-ne</i>	<i>coneixe'n</i>

Taula 4.13 Regularització de pronoms febles darrere infinitiu amb *-r* elidida

<b>Infinitiu amb pronom feble (forma estàndard)</b>	<b>Infinitiu amb <i>-r</i> elidida i pronom feble amb epèntesi o altres clítics (forma regularitzada)</b>
<i>donar-nos</i>	<i>donà'ns-e</i>
<i>donar-vos</i>	<i>donà-us-e</i>
<i>dir-vos</i>	<i>dí-us-e</i>
<i>donar-los</i>	<i>donà'ls-e</i>
<i>donar-les</i>	<i>donà'ls-e</i>
<i>donar-nos-el (-els, -en, -ho)</i>	<i>donà'ns-el (-els, -en, -ho)</i>
<i>donar-nos-la (-les)</i>	<i>donà'ns-e-la (-les)</i>
<i>donar-vos-el (-els, -en, -ho)</i>	<i>donà-us-el (-els, -en, -ho)</i>
<i>donar-vos-la (-les)</i>	<i>donà-us-e-la (-les)</i>
<i>donar-los-el (-la, -ho, -en...)</i>	<i>donà'ls-el (-la, -ho, -en, -hi)</i>
<i>anar-nos-en</i>	<i>anà'ns-en</i>
<i>poder-nos</i>	<i>podé'ns-e, pogué'ns-e</i>

Taula 4.14 Regularització de pronoms febles darrere infinitiu amb *-r* elidida i epèntesi

Finalment, a la taula 4.15 anotem alguns casos particulars.

<b>Formes particulars (grafia estàndard)</b>	<b>Formes particulars (grafia regularitzada)</b>
<i>dóna'm</i>	<i>do'm, de'm</i>
<i>dóna-me'n</i>	<i>dóna-me-ne'n</i>
<i>mira-te'l, mira-te-la</i>	<i>mi(ra)-te'l, mi(ra)-te-la</i>

<i>vés-li a dir que...</i>	<i>vés-e-li a dir que...</i>
----------------------------	------------------------------

Taula 4.15 Regularització de formes pronominals particulars

#### 4.4.15.4 Lèxic

Hi ha dos processos de naturalesa morfològica que solen tenir una incidència directa en la regularització de solucions lèxiques no estàndards. En primer lloc, el polimorfisme, el fenomen pel qual un mateix mot presenta més d'una variant morfològica, com ara l'adverbi i connector *llavors*, que pot aparèixer en llengua oral amb les variants *llavors*, *allavors*, *llavores*, *allavores*, *llavòrens*, *allavòrens*, *llavontes* i *llons*, entre d'altres. I, en segon lloc, la gramaticalització, entesa com el canvi gradual que experimenta un mot lèxic en transformar-se en mot funcional, normalment amb alteracions morfològiques i pèrdua de contingut semàntic (veg. Cuenca 2002, Lorente Casafont 2002); n'és un exemple ben conegut l'oració atributiva *és clar* (*És clar que això no pot continuar així*), que ha esdevingut una interjecció, sovint escrita *esclar* (*A mi la xocolata m'agrada, esclar!*).

Respecte als casos de polimorfisme, el criteri general de transcripció serà reproduir ortogràficament totes les possibles variants del mateix mot sense alterar-ne gaire la forma estàndard i, sempre que sigui possible, emprant grafies amb certa tradició escrita. Així, per exemple, en el cas de *llavors*, adaptarem la majoria de les seves variants formals tal com apareixen documentades al *DCVB*. La variant que pot resultar més complicada de regularitzar és *llons*, ja que podríem dubtar entre aquesta mateixa grafia i la forma *ll(av)ons*. Aquí, però, triarem *llons*, perquè l'ús de parèntesis en un mot tan freqüent en llengua oral pot carregar encara més una transcripció sovint ja prou atapeïda.

També serien un exemple de polimorfisme els mots truncats *col·le*, *pel·li*, *profe*, etc. Malgrat que podríem regularitzar fàcilment algunes d'aquestes formes mitjançant parèntesis (cf.

*col·le(gi)*, *profe(ssor)*), preferirem les solucions escurçades (*col·le*, *pel·li*, *profe*), utilitzades sovint en registres escrits informals.

Pel que fa als mots en procés de gramaticalització, intentarem reflectir l'estadi morfològic que hagi assolit cada forma, sobretot si ja disposem de grafies més o menys consolidades. La taula 4.16 compila algunes solucions amb tradició escrita, juntament amb formes regularitzades amb parèntesis o fins i tot amb un canvi de grafia. Hi afegim, en cada cas, un comentari addicional.

Forma estàndard	Forma regularitzada	Comentari
<i>a veure</i>	<i>a veure</i> , <i>a ve(u)re</i>	Malgrat que <i>a vere</i> seria una regularització possible, considerem que s'aparta massa de la forma estàndard i de l'adaptació de formes del mateix verb no gramaticalitzades (cf. <i>ja ve(u)ràs</i> , <i>ja ve(u)rem què passa</i> ).
<i>adéu</i>	<i>adéu</i> , ( <i>a</i> ) <i>déu</i>	Malgrat que l'elisió de la <i>a</i> inicial és gairebé sistemàtica, utilitzarem parèntesis i evitarem la forma <i>déu</i> , ambigua.
<i>Déu n'hi do</i>	<i>Déu n'hi do</i>	Hi hauria la possibilitat d'adaptar aquesta interjecció com a <i>déu-n'hi-do</i> , però mantindrem la forma estàndard, més habitual. Només emprarem guionets quan s'hi afegeix el diminutiu, cf. <i>Déu-n'hi-doret</i> .
<i>doncs</i>	<i>doncs</i> , <i>do(nc)s</i> , <i>donses</i>	L'adaptació <i>dos</i> , en lloc de <i>do(nc)s</i> , podria resultar ambigua. <i>Donses</i> , documentat al <i>DCVB</i> .
<i>és a dir</i>	<i>és a dir</i> , <i>és dir</i>	Preferirem l'elisió de la <i>a</i> més que no pas l'ús de parèntesis, cf. <i>és (a) dir</i> .

<i>és clar</i> (interjecció)	<i>esclar</i>	Documentada a Solà <i>et al.</i> (2002)
<i>me cago en l'ou, l'olla, Déu, etc.</i>	<i>me cago en l'ou, l'olla; mecagondena, cagondena; mecagoncoi, cagoncoi; cagondéu, mecagondéu...</i>	Si el parlant fa servir una forma que manté l'article ( <i>me cago en l'ou, l'olla, etc.</i> ), emprarem la grafia estàndard. En cas contrari, optarem per la forma contreta.
<i>si us plau, sisplau</i>	<i>si us plau, sisplau</i>	En cas que s'utilitzi la forma <i>si et plau</i> i s'elideixi la <i>e</i> , la marcarem entre parèntesis: <i>si (e)t plau</i> .
<i>vaja</i> (interjecció)	<i>vaja, vaia</i>	Tot i que en general no distingim ortogràficament les pronúncies palatals [•]/[•], en aquest cas sí que farem la distinció d'una manera paral·lela a <i>vejam/veiam</i> .
<i>vejam</i> (interjecció)	<i>vejam, veiam, viam, aviam; (mall.) meiam, meem...</i>	Formes documentades al <i>DCVB</i>

Taula 4.16 Regularització de formes gramaticalitzades

Finalment, regularitzarem els sintagmes preposicionals *d'això* i *d'allò* gramaticalitzats com a omplidors de parla o mots comodí. Vegeu que a (91a) *d'això* funciona com a sintagma lliure, mentre que a (91b: 2020) i (91c: 430) serveix tant per substituir un nom com per conservar el torn de parla. En aquests dos últims casos farem servir les formes substantivades *daixò* i *dallò* en comptes dels sintagmes preposicionals.

(91a) 1034 MEC: *d'això creieu-se'n la mitat eh*  
(COC: 04)

(91b) 2018 AMM: *era\_*  
2019 *era\_*  
2020 *era a can: daixò\*  
2021 *can::\_*  
2022 *la companyia elèctrica com es diu a can::\_*  
(COC: 03)

(91c) 429 PPP: (... 1.21) *ells tenien un forn elèctric\_*

430	(... 1.10) un daixò_
431	aï\
432	(.. 0.75) un fogó elèctric_
(COC: 07)	

Com a substantius, totes dues formes presenten altres variants (*daixòs, dallòs; daixonses, dallonses, etc.*), que també adaptarem ortogràficament.

#### 4.4.15.5 Regularització de manlleus

La regularització de mots procedents d'altres llengües dependrà del grau d'integració fonètica que presentin. En aquells casos en què la pronúncia es manté igual que en la llengua original o s'intenta reproduir-la, delimitarem el mot o l'expressió amb el codi *{(L2) text afectat}*, que és el que també utilitzarem per marcar fragments de parla enunciats en altres llengües, tal com veurem més endavant, veg. (92). En canvi, si el mot ja està mínimament adaptat a la fonètica de la llengua de transcripció, prescindirem del codi i regularitzarem el mot.

En l'adaptació al català de mots d'origen castellà, seguirem les convencions següents:

a) No assenyalarem la neutralització de la vocal o mitjançant [u] i, per tant, transcriurem *bueno* (no *buenu*), *des de luego* (no *des de luegu*), *algo* (no *algu*), etc. Altres formes semblants: *puesto*, (*col·legi de*) *pago*, *cego*, *bolso*, *centro*, *salvo (que)*, *cuidado*, *barato*, *barco*, *monedero*, *crio*, *quinto*, *curro*, *canguro*, *balonmano*, *baloncesto*.

b) Generalment, representarem el so castellà [•] amb la grafia *h* en mots adaptats fonèticament: *halar (jalar)*, *hinyar (giñar)*. En força casos, però, haurem de mantenir les grafies *j/g* castellanes per no alterar massa l'ortografia del mot: *juerga*, *jefe*, *pijo*, *fijo*, *trajín*, *ojo!* (pronunciat [••••]). Com a excepció anotarem amb el codi *{(L2) text afectat}* la forma *ho*, adaptació de la interjecció *¡jo!*, per evitar l'ambigüitat amb el pronom personal.

c) Representarem el so castellà [•] amb les grafies *s*, *ss* o *ç*, segons convingui, en mots adaptats fonèticament: *consejal*, *cruse*, *en ves de*, *retxassar*, *plaço fijo*.

d) Farem correspondre la grafia castellana *ch* a les catalanes *tx* i *x*, segons la pronúncia, en mots adaptats fonèticament: *trutxa*, *lutxàvem*, *enxufar/entxufar*, *retxassar*, *catxarros*, *txutxeries*.

e) Assignarem el codi {(L2) text afectat} a tots els mots que mantinguin íntegrament la fonètica castellana, sobretot si no presenten reducció vocàlica: {(L2) *desplazamientos*, *hope*, *polvos de talco*, *experimentos*, *adios muy buenas*, *huelga*, *orujo*, *Nochebuena*, *hombre*, *paliza*, *vídeos musicales*, *buzo*, *pomposo*, *moderno*, *vía libre*, *albergue juvenil*, *azafata*, *anda ya*, *traje*, *jolín*, *jolines*, *ojalá*, *pase*, *culebrón*}. Tanmateix, prescindirem d'aquest codi en els marcadors discursius *vale*, *pues* (i *pos*) i *vamos*, molt freqüents en llengua oral.

f) Adequarem a les normes ortogràfiques del català tots els mots adaptats fonèticament: *nòvio*, *collarín*, *en sèrio*, *quarto*, *es repentia*, *arrepenteixo*, *por ahi*, *s'olviden*, *g(a)raig*, *demés*, *s'assentem*, *acostumbrar*, *desperdici*, *pasteleria*, *llamar l'atenció*.

Pel que fa als mots procedents d'altres llengües, emprarem el codi {(L2) text afectat} si el parlant reproduïx la pronúncia original: {(L2) *on line*, *landscape*, *ciao*, *egun*, *suite*}; i també si fa servir estrangerismes freqüents parcialment adaptats: {(L2) *light*, *heavy*, *hobby*, *please*, *software*}. En aquest últim cas, però, si ja existeix una grafia adaptada o una forma convencional força estesa, l'aprofitarem: *friqui* (veg. Ésadir), *web*, *okey*.

Finalment, també delimitarem amb el codi {(L2) text afectat} tot fragment de parla produït en qualsevol altra llengua que no sigui la de la transcripció, tal com s'exemplifica a (92) en castellà, anglès i català.

(92a)	230	JJJ:	diu\
	231		{(L2) pero qué es eso\}
	232		diu\
	233		oh\
	234	(. 0.16)	{(L2) tú m(e) has pedido eso\}
	235		ya me he extrañado\}

- 236 diu\  
 237 {(L2) porque tú normalmente no bebes esto\  
 (COC: 09)
- (92b) 1338 MMA: (.. 0.58) jo hi posat\_  
 1339 (.. 0.75) {(L2) what's capital\_  
 1340 (. 0.27) dallonsis/  
 1341 i era {(L2) what's the capital\  
 1342 i me l'han tatxat\  
 (COC: 01)
- (92c) 123 I01: a:nd\_  
 124 there i- --  
 125 there is anothe:r\_  
 126 (.. 0.98) another bathroom\_  
 127 (. 0.20) {(L2)(P) ah no\  
 128 no t'ho he dit\  
 129 (.. 0.51) there is a bathroom\_  
 (CAP: E3CD01CC)

#### 4.4.16 Convencions tipogràfiques

Com altres models de transcripció discursiva semblants, el sistema del CCCUB obvia algunes convencions de la llengua escrita per evitar ambigüitats entre les normes tipogràfiques generals i les convencions pròpies de la mateixa transcripció, sobretot si és de nivell intermedi o estret. Per tant, tindrem en compte les indicacions següents:

1. No emprarem símbols ni signes de puntuació no recollits pel model aquí descrit (veg. també § 4.5.1) i, si n'utilitzem, no els atribuirem valors ja previstos pel sistema i n'explicarem l'ús en la documentació que acompanya la transcripció.

2. Sempre que sigui possible, evitarem les majúscules, que generalment reservarem per a les inicials dels parlants i



símbols com ara (*INH*) i (*EXH*). Dins el text només posarem en majúscula inicial els noms propis, incloent-hi antropònims, topònims (*serra del Montsant, cap de Creus*), noms d'empreses, institucions i altres organismes (*Gas Natural, Caixa de Catalunya, Universitat de Barcelona, Generalitat, Creu Roja*). També farem servir majúscules en sigles i acrònims (*UGT, ONU, Renfe, Benelux*), per bé que sempre preferirem les formes que convencionalment s'escriuen amb majúscula inicial només (*Unesco*, no *UNESCO*; *Seat*, no *SEAT*) i tractarem com a noms comuns els mots procedents de sigles o acrònims que no es refereixin a organismes (*iva*, no *IVA*; *eso*, no *ESO*; *oenagé*, no *ONG*; *cedé*, no *CD*). I, finalment, aniran en majúscula les entitats religioses (*Déu, Senyor, Esperit Sant*) tant si les trobem aïllades com a l'interior d'expressions ja fixades (*Déu n'hi do; com Déu mana; ai, Senyor*). En cas de dubte, sempre optarem per les minúscules, sobretot en noms genèrics (*ajuntament, departament de recursos humans, hotel Ritz, conselleria de cultura, administració, església*).

En cap cas farem servir cursiva, negreta, versaletes, subratllat ni altres recursos tipogràfics no previstos pel model de transcripció discursiva del CCCUB i, si n'utilitzem —per exemple, per marcar fragments rellevants per a un estudi concret—, ho especificarem en els documents annexos de la transcripció.

3. Transcriurem les xifres, incloent-hi els anys, en lletres, veg. (93). Només farem servir números per identificar els grups tonals, indicar la durada de pauses, riure, fragments inintel·ligibles, etc. i aparellar encavalcaments.

- (93) 1370 JPP: a l'any vuitanta-nou\_  
1371 hi havien dinou mil tres-cents vuitanta-quatre alumnes\_  
1372 (INH)  
1373 i a l'any noranta-set\_  
1374 n'hi ha cent coranta mil\  
1375 (INH)  
1376 (ai)xò sol\_  
1377 ja és important\_  
(COR: POL2)

4. No s'ha establert cap format específic per al text de la transcripció, a banda de la disposició vertical de la pàgina descrita a § 4.2. Per tant, l'edició definitiva de les transcripcions dependrà de l'anàlisi a què vagin destinades i caldrà adaptar aspectes formals del text com ara l'amplada dels marges i el tipus i cos de lletra, entre d'altres, als interessos particulars de cada estudi. Com a orientació, presentem a la taula 4.17 les convencions tipogràfiques que habitualment s'han aplicat a les transcripcions discursives del CCCUB. Com es pot observar, es tracta d'un format molt senzill, pensat sobretot per facilitar la difusió dels textos.

<b>Marges de la pàgina</b>	Format estàndard dels processadors de textos més habituals. Marges superior i inferior: 2,5 cm; marges esquerre i dret: 3 cm
<b>Justificació del text</b>	Esquerra
<b>Interlineat</b>	Senzill
<b>Tabuladors</b>	Situats a 0 cm per a la columna de numeració, 1 cm per a les inicials dels parlants i 2,5 cm per al text; tot i així, per a les transcripcions amb numeració de quatre dígitos o amb inicials de parlants llargues, ens veurem obligats a fer servir tabuladors més amples des de l'inici de la transcripció.
<b>Tipus i cos de lletra</b>	Times New Roman, cos 12
<b>Numeració de pàgines</b>	Marge inferior dret, Times New Roman, cos 12
<b>Identificació de la transcripció</b>	A l'encapçalament, al marge dret

*Taula 4.17* Convencions tipogràfiques de la transcripció discursiva del CCCUB

Recomanem que s'apliqui des de bon principi el format de text que s'hagi escollit, perquè qualsevol modificació en aquest sentit implicarà revisar transcripcions ja acabades i hi ha el perill que un canvi en el cos de lletra, els marges de la pàgina, etc., alteri els tabuladors, la numeració o altres elements del text ja transcrit.

#### 4.4.17 Errades de transcripció

En aquest apartat comentem algunes de les errades més freqüents a l'hora d'aplicar els criteris de transcripció discursiva del CCCUB. Aquí resulta impossible preveure tots els casos conflictius que poden sorgir al llarg de la transcripció i, per tant, només tractarem les errades més habituals a partir d'exemples adaptats (marquem amb un asterisc la crida de les transcripcions incorrectes).

1. Els números a l'esquerra de la transcripció es refereixen als grups tonals —així com a les pauses no atribuïdes a cap parlant i als comentaris del transcriptor—, no a les línies de la pàgina. Per tant, a un grup tonal llarg que ocupa més d'una línia només li correspondrà un número, com a (94a), no dos, cf. \*(94b: 384).

(94a) 383 NAA: (... 2.02) nosatres de moment fins que no tinguem casa no  
decidirem re\  
384 SUS: (.. 0.43) @@  
(COC: 02, adaptat)

\*(94b) 383 NAA: (... 2.02) nosatres de moment fins que no tinguem casa no  
384 decidirem re\  
385 SUS: (.. 0.43) @@

2. Per no atapeir més la transcripció, no és recomanable repetir les inicials del parlant davant de cada grup tonal. N'hi ha prou anotant-les una sola vegada al principi de cada intervenció. També hem d'assegurar-nos que les inicials estan col·locades segons l'ordre d'intervenció dels parlants, sobretot en aquelles

transcripcions que compten amb molts informants. Vegeu que a \*(95a: 1180) no és possible que es repeteixin les inicials de MEC, perquè aquest mateix informant ja ocupa el torn de parla, cf. (95b: 1180).

* (95a)	1177	MEC:	jo et dic una cosa\	(95b)	1177	MEC:	jo et dic una cosa\
	1178		li volia fer un detall\		1178		li volia fer un detall\
	1179		p(e)rò ara s'han casat molta gent\		1179		p(e)rò ara s'han casat molta gent\
	1180	MEC:	[i passo\]		1180		[i passo\]
	1181	ONI:	[quina--]		1181	ONI:	[quina--]
	1182		quina Laura és\		1182		quina Laura és\
	1183	MEC:	la Laura Rebull\ (nom fictici)		1183	MEC:	la Laura Rebull\ (nom fictici)
							(COC: 04)

3. No repetirem el mateix codi al principi de cada grup tonal dins una mateixa intervenció, com a \*(96a: 264-266), sinó que obrirem el codi en el primer grup tonal afectat i el tancarem en el darrer, cf. (96b: 264-266).

\* (96a) 263 NAA: allò que els començava a parlar\_  
 264 {(EV) que si el casament\_  
 265 {(EV) i que si Déu\_  
 266 {(EV) (. 0.26) i que si no sé quantos\_}

(96b) 263 NAA: allò que els començava a parlar\_  
 264 {(EV) que si el casament\_  
 265 i que si Déu\_  
 266 (. 0.26) i que si no sé quantos\_  
 (COC: 02)

4. La segmentació del discurs en grups tonals es regeix principalment per criteris prosòdics, no pas sintàctics i/o semàntics. Així doncs, a l'hora de determinar on comença i on acaba un grup tonal, no ens guiarem per estructures sintàctiques o semàntiques amb sentit complet com són el sintagma i l'oració, sinó sobretot pels aspectes prosòdics que hem descrit a § 4.4.1. Per aquest motiu caldrà revisar transcripcions com la de \*(97a),

en què cada grup tonal es correspon inusualment a un sintagma o una oració, cf. (97b).

<p>* (97a)</p> <p>1 I04: do:ncs_  2 has de creuar_  3 les vie:s_  4 del tren\  5 ECC: mhm\  6 I04: i tirar_  7 cap a l'esquerra_  8 seguint_  9 el mateix recorregut_  10 de les vies\  11 (... 1.00) llavors_  12 arribaràs_  13 a: l'església\  (CAP: C1T104CC)</p>	<p>(97b)</p> <p>1 I04: do:ncs_  2 has de creuar les  vie:s_  3 del tren\  4 ECC: mhm\  5 I04: i tirar cap a  l'esquerra seguint_  6 el mateix recorregut  de les vies\  7 (... 1.00) llavors  arribaràs a:  l'església\  (CAP: C1T104CC)</p>
---	--

Més amunt hem suggerit que en cas de dubte tendirem a dividir un mateix fragment discursiu en dos grups tonals, però això no implica que la segmentació sigui l'opció per defecte en qualsevol circumstància. Abans valorarem detingudament com es combinen els criteris prosòdics que aquí hem descrit (compleció melòdica, ressituació, pausa, etc.) i si, malgrat tot, encara tenim dubtes, segmentarem. Vegeu com a mostra l'extracte de (98a), en què podríem dubtar si l'adverbi *ja*: forma grup tonal amb *p(e)rò bueno*, si és un grup tonal independent o si s'adjunta a la resta de l'oració (*podeu comença:r a...*). En aquest cas el transcriptor no l'ha inclòs en el grup entonatiu següent —com seria esperable des d'un punt de vista sintàctic—, sinó dins el grup tonal anterior, perquè no es produeix ressituació del to i hi ha allargament vocàlic, cf. (98b: 1199).

(98a) DRO: exacte p(e)rò bueno ja: podeu comença:r a mira:r una miqueta:  
(.. 0.33) els preus per tindre una idea...

(98b) 1198 DRO: exacte\  
1199 p(e)rò bueno ja:\_  
1200 podeu comença:r\_  
1201 a mira:r\_  
1202 una miqueta:\_  
1203 (.. 0.33) els preus\_

1204 per tindre una idea\_  
(COR: EMP4)

5. Els símbols d'entonació terminal, truncament de grup tonal i truncament de mot i grup tonal no se situaran mai fora del contingut dels claudàtors d'un encavalcament o les claus d'un codi al final de grup entonatiu, sinó a dins i adherits al text. Evitarem, doncs, transcripcions com la de \*(99a), en què trobem alguns d'aquests símbols fora del contingut dels claudàtors (g. t. 120, 123, 126) o les claus (g. t. 122) o separats del text (g. t. 127), cf. la transcripció correcta de (99b). Sí que són possibles seqüències com la de (99c: 140), en què el parèntesi no ha d'incloure el símbol de manteniment tonal, i (99d: 1666), en què l'encavalcament s'acaba just abans de l'allargament vocàlic.

- \*(99a) 120 NAT: (... 1.49) de lligar el [què:]\  
121 MER: ((a TON)) [no els hi] diguis si són capaços de lligar\  
122 que això avui dia {( @) és una atra cosa}\  
123 MIA: [ja:]\  
124 TON: [oh\  
125 lli]gar no eh\  
126 lli[gar el viatge\  
127 AMM: {{{(F) d'organitzar}\}
- (99b) 120 NAT: (... 1.49) de lligar el [què:]\  
121 MER: ((a TON)) [no els hi] diguis si són capaços de lligar\  
122 que això avui dia {( @) és una atra cosa}\  
123 MIA: [ja:]\  
124 TON: [oh\  
125 lli]gar no eh\  
126 lli[gar el viatge\  
127 AMM: {{{(F) d'organitzar}\}
- (COC: 03)
- (99c) 139 MER: ja ho dèiem\  
140 mà d'autoca:(r)\_  
(COC: 03)
- (99d) 1664 MMM: (... 1.67) i què més ens ha de passar\  
1665 (.. 0.31) no [1 ho sé\  
1666 CME: [1 no sé:]\  
(COC: 07)

6. Quan uns claudàtors d'encavalcament i unes claus de codi contenen exactament el mateix fragment de text, els claudàtors sempre es col·locaran en una posició més externa que les claus, com a (100a: 536), i no a la inversa, cf. \*(100b: 536).

(100a)	535	SUS:	p(e)rò jo quan la	*(100b)	535	SUS:	p(e)rò jo quan la
			vaig veure_				vaig veure_
	536		{{(EV) oi que és		536		{{(EV) [oi que és
			maco/}}				maco/}}
			(COC: 02, adaptat)				

Si un codi conté un encavalcament que coincideix amb la clau d'obertura o de tancament, les claus sempre seran més externes. Per tant, corregirem \*(101a: 1008) i \*(101b: 608) per (101c: 1008) i (101d: 608) respectivament. Farem el mateix en el cas invers, cf. \*(101e) i (101f).

\*(101a) 1005 NAA: (.. 0.33) quan vaig calcular que tindriem arreglat el pis  
[1 ahir\_]  
1006 SUS: [1 @ @]@ @ @  
1007 NAA: [2 xx]  
1008 SUS: [2 {( @ ) amb] tres setmanes\}

\*(101b) 607 HIL: pots comptar com estava fet\_ ((el vestit))  
608 (. 0.29) {(??) fa[ta\]}  
609 SUS: [sem]blava\_  
610 saps un de la Nancy que teníem blau cel/

(101c) 1005 NAA: (.. 0.33) quan vaig calcular que tindriem arreglat el pis  
[1 ahir\_]  
1006 SUS: [1 @ @]@ @ @  
1007 NAA: [2 xx]  
1008 SUS: {( @ ) [2 amb] tres setmanes\}  
(COC: 02)

(101d) 607 HIL: pots comptar com estava fet\_  
608 (. 0.29) {(??) fa[ta\]}  
609 SUS: [sem]blava\_  
610 saps un de la Nancy que teníem blau cel/  
(COC: 02)

\*(101e) 821 ROM: [i he quedat amb el noi aquell per {(??) posar-me a treballar\_}]

(101f) 821 ROM: [i he quedat amb el noi aquell per {(??) posar-me a treballar\_}]  
(COC: 01, adaptat)

En canvi, quan els claudàtors i les claus no afecten el mateix fragment de text o no coincideixen en l'obertura o el tancament, els distribuïrem segons convingui en cada cas. Per exemple, (102) recull un fragment entre claudàtors inserit dins unes claus, veg. (102a: 627); un fragment entre claus dins de claudàtors, veg. (102b: 222), i dos grups entonatius en què la distribució de claudàtors i claus se superposa, veg. (102c: 87-88).

(102a) 625 TON: quan enraonava tothom es--  
626 tothom es tombava i se'l mirava\  
627 MER: {(F) a:: ua:: [ua::] ua:\}  
628 TON: [{(F) au::\}]  
629 T??: (. 0.16) @R(1.19)R@ @  
630 MER: tenia la veu més gruixuda llavors que ara\  
(COC: 03, adaptat)

(102b) 218 JJJ: p(e)rò bueno\  
219 i: que li posa un--  
220 [un go:t\_  
221 amb menta amb\_  
222 {(P) menta} amb licor de menta i\_  
223 i:\  
224 SER: [@@@ xXXx]  
(COC: 09)

(102c) 86 IDA: [tia: i: amb la iaia\  
87 MRC: [àvia\  
88 vols {(L2) orujjo\  
89 LIA: @ @ @  
90 IDA: vols orujo/  
(COC: 09, adaptat)



7. Quan hi pugui haver ambigüitat en el tancament de claus i claudàtors, com a \*(103a: 43), \*(103b: 59) i \*(103c: 1742), en què no sabem quin abast té cada codi o encavalcament, repetirem el codi abans de tancar la clau, cf. (103d: 43, 103e: 59), o el número en el cas dels claudàtors, cf. (103f: 1742).

\*(103a) 42 JUG: jura o promet dir la veritat/  
 43 TE1: (. 0.24) {{??}}{(P) ho} juro\  
 44 JUG: (.. 0.47) digui vostè el seu nom i cognoms\  
 \

\*(103b) 57 MOL: (.. 0.70) {(EV) ui\  
 58 (.. 0.50) {(P) qui truca\  
 59 la mare ens ha dit que no obrim\  
 \

\*(103c) 1741 ????: (.. 0.53) ah és veritat [1 xX(1.01)Xx\  
 1742 IDA: [1 està a [2 dins d'un] calaix\  
 1743 ????: [2 {{??}} home:\]  
 1744 IDA: (. 0.17) a casa teva\  
 \

(103d) 42 JUG: jura o promet dir la veritat/  
 43 TE1: (. 0.24) {{??}}{(P) ho} juro\  
 44 JUG: (.. 0.47) digui vostè el seu nom i cognoms\  
 (COR: CIV1)  
 \

(103e) 57 MOL: (.. 0.70) {(EV) ui\  
 58 (.. 0.50) {(P) qui truca\  
 59 la mare ens ha dit que no obrim\  
 (COR: EDUC2)  
 \

(103f) 1741 ????: (.. 0.53) ah és veritat [1 xX(1.01)Xx\  
 1742 IDA: [1 està a [2 dins d'un 2] calaix\  
 1743 ????: [2 {{??}} home:\]  
 1744 IDA: (. 0.17) a casa teva\  
 (COC: 09)  
 \

Un cop acabada la transcripció, és molt aconsellable repassar el text per assegurar-nos que no queda cap fragment de parla delimitat per claus, claudàtors o parèntesis sense el signe

d'obertura o de tancament i que no hi pot haver cap confusió a l'hora d'interpretar aquests símbols.

8. Sempre que puguem, respectarem l'ordre cronològic del discurs. Vegeu que l'exemple (53b), aquí reprès com a \*(104a), no reflecteix correctament l'ordre temporal en què s'enuncien els grups tonals 357 i 358 i, per tant, l'haurem de corregir per (104b). Recordem, però, que en alguns casos no és possible representar adequadament la seqüència temporal del discurs si no es fa servir el símbol & (veg. § 4.4.7), més propi d'una transcripció estreta.

<p>*(104a) 354 FAN: i:_          355 i per què portes          els cascos          posats/          (... 2.64)          356 (... 2.64)          357 HIL: [perquè està          es]coltant la          ràdio\          358 FAN: [no ho entenc\]</p>	<p>(104b) 354 FAN: i:_          355 i per què portes          els cascos posats/          (... 2.64)          356 (... 2.64)          357 FAN: [no ho entenc\          358 HIL: [perquè està          es]coltant la          ràdio\          (COC: 02, adaptat)</p>
--	---

9. Els codis que afecten un mateix fragment de text s'ordenaran per ordre alfabètic si contenen una lletra, veg. (105a: 821, 823); si contenen un símbol, aquest se situarà en primer lloc, veg. (105b), i entre símbols sempre tindrà preferència el codi *{{(??) text afectat}}*, veg. (105c: 117).

(105a) 821 MJ?: [{{(EV)(F) ai ai ai\}}]  
 822 EUU: [@@@]  
 823 MJJ: [{{(AC)(E)(F)(L2) ánde\}}]  
 824 ánde\  
 825 ánde\}  
 (COC: 05)

(105b) 190 NIA: (... 2.41) [{{(??)(P) què t'ha dit que aniran al Boira\}}]  
 (COC: 06)

(105c) 116 LAU: [{{(??) algun ha dit tacos\}}]  
 117 ????: [{{(??)(@) ai\}}]  
 118 ????: [jo no\}]

(COC: 01, adaptat)

Vegeu també § 4.4.9.1 per a l'ordenació de codis que no afecten el mateix fragment de discurs.

10. Col·locarem els allargaments allí on pertoqui d'acord amb la pronúncia. Per exemple, en la forma verbal *va venir::r* l'allargament se situarà darrere la vocal *i*, no darrere la *err* (*va venir:::*), encara que sigui muda. L'única excepció serà la grafia *h*, *i*, per exemple, transcriurem la interjecció *oh* amb l'allargament darrere la *hac* (*oh:::*) i no darrere la vocal (*o::h*).

11. Procurarem que no quedin tabuladors, salts de línia ni espais sobrers, veg. § 4.2.

12. Sovint trobarem fragments ambigus des del punt de vista lingüístic que inevitablement haurem d'interpretar. Per reduir al mínim la possibilitat d'error, aquí recomanem decantar-se per la solució més probable dins el context comunicatiu en s'inscriu el fet de parla. Per exemple, a (106) la forma pronominal *l'hi* podria correspondre a la combinació *li ho*, en què *ho* reprèn l'oració substantiva amb funció de complement directe *lo que li deuen*, desplaçada al final de l'enunciat, o bé equivaler al datiu singular homòfon *li* si considerem que no hi ha hagut desplaçament.

(106) 1141 CME: però a vostè l'hi han pagat lo que li deuen\  
(COC: 07, adaptat)

En aquest cas el transcriptor, basant-se en la prosòdia, ha interpretat que s'ha produït un desplaçament i ha optat per *l'hi* en lloc de *li*.

Finalment, davant de qualsevol cas conflictiu no previst per aquest manual, aconsellem que es tendeixi a triar l'opció que comprometi menys la lectura de la transcripció. Posem per cas que trobem un fragment de parla especialment difícil d'entendre perquè hi intervenen molts interlocutors alhora:

abans que res haurem de sospesar si val la pena transcriure'l, tenint en compte que hi haurem d'invertir molt d'esforç i segurament obtindrem com a resultat un text espès de llegir i poc rellevant per a l'anàlisi posterior. En aquest cas sembla més enraonat indicar mitjançant un comentari del transcriptor o el símbol *xX(durada)Xx* que no transcrivim el fragment en qüestió.

#### **4.5 NIVELLS DE LA TRANSCRIPCIÓ DISCURSIVA DEL CCCUB**

Les convencions que hem descrit en els apartats precedents corresponen a un nivell de transcripció intermedi-estret, que és el que considerarem més neutre i seguirem per defecte. No obstant això, el sistema de transcripció discursiva del CCCUB preveu dues modalitats de transcripció complementàries: una transcripció ampla, que no incorpora tanta informació com la de nivell intermedi-estret (veg. § 4.5.1), i una transcripció estreta, amb convencions inèdites en la modalitat que hem vist fins ara (veg. § 4.5.2). Depenent dels objectius de la transcripció i de l'anàlisi posterior, adoptarem una modalitat o una altra, i fins i tot podrem combinar-ne més d'una si així ho demana la recerca que volem portar a terme. Res no impedeix, per exemple, que en un estudi sobre la gestió de torns de parla transcrivim el text fent servir una modalitat ampla però incloent-hi convencions més pròpies d'un nivell intermedi-estret, com ara allargaments, claudàtors i pauses breus, per assenyalar els fragments en què els interlocutors competeixen per ocupar el torn de parla o mantenir-lo. Per tant, les dues modalitats de transcripció que desglossem tot seguit no s'han de veure com a sistemes tancats i inamovibles, sinó sobretot com a propostes que caldrà adaptar als objectius de cada estudi particular.

##### **4.5.1 Transcripcions ampla i intermèdia**

Generalment, la transcripció discursiva ampla que preveu el CCCUB manté la segmentació del discurs en grups tonals, però prescindeix de molts dels símbols i codis que s'utilitzen en el nivell intermedi-estret. A continuació descrivim les convencions que caracteritzen aquesta modalitat:

1. *Disposició del text dins la pàgina.* Distribuïrem el text de la transcripció en línies i paràgrafs —no en grups tonals ordenats verticalment—, d'una manera semblant a les obres de teatre. Numerarem cada línia al marge esquerre, identificarem els parlants amb una abreviatura amb majúscula inicial i introduïrem cada torn mitjançant dos punts. Si no podem identificar el parlant que intervé, farem servir el símbol ???, i si tots els interlocutors parlen alhora, T??

2. *Descripció de l'escena.* La presentació del context comunicatiu en què es desenvolupa el fet de parla (incloent-hi els participants, l'escenari físic, etc.) serà esquemàtica i, per tant, no inclourà tants detalls com la descripció basada en el model Speaking.

3. *Aspectes prosòdics.* Per norma general, segmentarem el text en grups entonatius, però representarem les inflexions tonals mitjançant els signes de puntuació següents: descens (.), ascens (?) i manteniment (,). Si no necessitem dividir el discurs en grups entonatius, farem servir els signes de puntuació convencionals (coma, punt, punt i seguit, etc.) segons les normes ortogràfiques, però ho farem constar en els annexos de la transcripció per no confondre el lector.

Representarem el truncament de mot amb guionet (-), el truncament de grup tonal (i mot), amb dos guionets (--), i l'èmfasi, amb majúscules. També podem marcar un fragment de parla amb el codi *{{fenomen) text afectat}* per destacar fenòmens prosòdics rellevants, per exemple, *{{intensitat baixa) no ho diguis a ningú}*. I només anotarem les pauses inferiors a cinc segons, amb (.), o iguals o superiors, amb (..), i no indicarem ni allargaments ni encavalcaments.

4. *Aspectes vocals i gestuals.* Transcriurem el riure amb el comentari (*riure*) i descriurem la resta de fenòmens vocàlics i/o gestuals entre parèntesis simples, per exemple, (*s'escura la gola*) i (*saluda amb la mà*), o bé amb el codi *{{(fenomen) text afectat}* si són simultanis amb la parla, per exemple, *{{(saludant amb la mà) adéu}*.

5. *Regularitzacions.* Sempre que sigui possible, transcriurem amb grafia estàndard i col·locarem les formes regularitzades entre cometes i parèntesis, per exemple, *estiu* (“*istiu*”). Per als sons vocals no verbals, farem servir les convencions de la llengua escrita (veg. § 4.4.10) i assenyalarem els fragments en una altra llengua amb el codi *{{(L2) text afectat}*. Reservarem les majúscules per a la inicial dels noms propis (veg. § 4.4.16), incloent-hi l'abreviatura dels parlants, i per marcar l'èmfasi, per exemple, *ara PLOU!?*

6. *Altres.* Tots els comentaris explicatius del transcriptor aniran entre parèntesis dobles, per exemple, *((se'n va a la cuina))*, *((algú truca a la porta))*. Els fragments inintel·ligibles els indicarem amb el comentari (*inintel·ligible*) i els que resultin dubtosos, amb el codi *{{(??) text afectat}*.

Exemplifiquem les convencions de la transcripció discursiva ampla amb el fragment de (107), extret del COR (EMP4: 138-234). Com podem comprovar, s'ha escurçat molt la descripció inicial de l'escena i la disposició del text és horitzontal, no vertical. S'hi assenyalen els grups tonals, però no els encavalcaments ni els allargaments, i les seqüències tonals terminals són representades per la coma, l'interrogant i el punt. També es limita molt l'ús de codis o se substitueixen per altres recursos (per exemple, l'èmfasi es marca amb majúscules) i es prescindeix dels parèntesis per indicar les elisions. Anotem les formes regularitzades entre parèntesis.

(107) Dues clientes s'informen de les condicions per fer un viatge a l'estiu a la República Dominicana. Data, hora i lloc de la interacció:

dissabte 22 de març de 1998, cap a les dotze del migdia, en una agència de viatges propera a la plaça Universitat de Barcelona. Participants: una venedora de l'agència (Dro) i dues clientes (Aft i Gef).

- 1 Aft: (.) i, bueno. també tinc entès, que hi han, diferents zones, la zona de l'est, (.) La Romana i tot això,
- 2 Dro: jo ara us explico. (.) anem a veure. (.) a veure. (.) agafo un prospecte, si és que n'hi ha, (.) bueno. (.) a veure. (.) vosaltres teniu. (.) a, (.) la zona. (.) de Puerto Plata, Sosúa, tot això d'aquí, (assenyala el catàleg) (.) que és, diguem-ne, la que, (.) la zona, (.) {{??}} llavors,} anem a dir que els hotels són més assequibles, (.) e,
- 3 Aft: més turística.
- 4 Dro: una zona MOLT turística, (.) que hi han MOLTS hotels, (.) i que també hi ha més animació. (.) d'acord? (.) després està, (.) la zona de Samaná. (.) que és una de les zones, més, e, (.) més verges, que hi ha, a tota la illa, (.) però-- molt maca, (.) però-- no és recomanable, (.) per dues ("dugues") noies.
- 5 T??: (.) (riure)
- 6 Dro: és molt recomanable,
- 7 Aft: xungo.
- 8 Dro: per una pare-- no. no per xungo. (.) perquè és molt recomanable per una parella.
- 9 Aft: (.) ah. vale.
- 10 Dro: si jo et fico en un hotel en aquí, (.) tu t'ago-- t'agobiaràs com una ostra.
- 11 Aft: vale.
- 12 Dro: a veure. ojo. si em dieu que, (.) teniu moltes ganes de relax, (.) i només que relax,
- 13 Gef: {(rient) entre poc ni massa.}
- 14 Aft: què vols dir. que és més paisatgístic, i romàntic, i,
- 15 Dro: exactament. i és estar-hi amb un hotel, amb el règim aquest del {(L2) todo incluido,} que si no tens ganes de sortir de l'hotel, no cal que surtis, perquè ho tens tot, (.) però tot ho trobaràs al teu voltant,
- 16 Gef: no. no (inintel·ligible). (riure)
- 17 Aft: (.) no. nosaltre-- ("nosatro--")
- 18 Dro: són parelles.
- 19 Aft: ens interessa també conèixer una mica,
- 20 Gef: clar.
- 21 Dro: (.) llavors ("allavons") aquí, no podeu anar. és, (.) {(desaccelerat) PRECIÓS}, (.) però no podeu anar, perquè us agobiareu. (.)

- perquè tot lo que hi ha al voltant, (.) són parelles. (.) llavors ("llavors") clar.
- 22 T??: (.) (riure)
- 23 Dro: us trobareu molt agobiades. (.) vale? (.) i després està la zona de La Romana, (.) Playa Bávaro,
- 24 Aft: aquesta zona és la que he sentit jo parlar.

Convé insistir que aquests criteris no són inalterables i que es poden adaptar en tot moment als interessos de cada transcripció. A (107), per exemple, hi podríem incloure les pauses molt breus, breus i mitjanes si ens interessés reflectir amb més fidelitat l'idiòlecte de cada interlocutor, o podríem ometre'n els fragments regularitzats si considerem que són prescindibles i més aviat entorpeixen la lectura del text.

Per mostrar més clarament les diferències entre la transcripció discursiva ampla i la transcripció de nivell intermedi-estret, reproduïm a sota el mateix fragment de (107) tal com apareix recollit al COR (EMP4: 138-234). Per motius d'espai, però, ometem la descripció de l'escena, que es pot consultar a Alturo *et al.* (2004).

(108)	Descripció de l'escena a partir del model Speaking	152	(... 1.22) vosaltres teniu\
		153	(. 0.26) a: _
		154	(. 0.37) la zona\
138	AFT: (... 1.08) i: _	155	(. 0.23) de Puerto
139	bueno\		Pla:ta_
140	també tinc entès_	156	Sosúa_
141	que hi han: _	157	tot això d'aquí_
142	diferents zones_		((assenyala el catàleg))
143	la zona de l'est_	158	(. 0.52) que és_
144	(. 0.49) La Roma:na i	159	diguem-ne_
	[tot això_]	160	la que: _
145	DRO: [jo ara] us explico\	161	(. 0.12) la zona_
146	(. 0.49) (a)nem a	162	(. 0.31) {(??) llavors_}
	ve(u)re\	163	(a)nem a dir que els
147	(... 1.52) a ve(u)re\		hotels són més
148	(... 1.52) agafo un prospecte: _	164	assequible:s_
			(. 0.45) e: _
149	si és que n'hi ha_	165	AFT: més turística\
150	(. 0.22) bueno\	166	DRO: una zona {(F) molt}
151	(. 0.92) a ve(u)re\		turística_



167	(. 0.82) que hi han {(F) molts} hotels_	201	relax_ GEF: {( @) entre poc ni massa}
168	(. 0.60) i que també hi ha més animació\	202	AFT: què vols dir\
169	(. 0.67) d'acord/	203	que és més paisatgísti:c_
170	(. 0.77) després està_	204	i romà:nti:c_
171	(. 0.18) la zona de Samaná\	205	[i:]_
172	(. 0.12) que és una de les zones_	206	DRO: [exac]tament\ i és estar-hi amb un hotel_
173	més_	207	
174	e::_	208	am(b) el règim aquest del {(L2) todo incluido_}
175	(. 0.52) més verges_	209	que si no tens ganes de sortir de l'hotel_
176	que hi ha_	210	no cal que surtis_
177	a tota la illa_	211	perquè ho tens tot_
178	(. 0.50) p(e)rò--	212	(. 0.13) p(e)rò tot ho [1 trobaràs al teu] voltant_
179	molt maca_		
180	(. 0.46) p(e)rò--	213	GEF: [1 no\ no x\ @]
181	no és recomanable_	214	AFT: (. 0.28) [2 no\ nosatro- --]
182	(. 0.26) per dugues noies\ 183 T??: (. 0.37) @R(1.18)R@	215	DRO: [2 són pare]lles\ AFT: ens interessa també conèixer una mi[ca:_ 3]
184	DRO: és molt recoma[nable_]	216	GEF: [3 clar]
185	AFT: [xungo]	217	DRO: (. 0.84) allavons aquí_ no podeu anar\ és_ (. 0.41) {(DC)(F) preció:s_}
186	DRO: per una pare- --	218	(. 0.20) p(e)rò no podeu anar_ perquè us agobiareu\ (. 0.13) perquè tot lo que hi ha al voltant_ (. 0.11) són parelles\ (. 0.34) llavons clar\ T??: (. 0.25) @@ [ @R(2.31)R@]
187	no\ no per xungo\ 189 (. 0.25) perquè és molt recomanable per una parella\ 190 AFT: (. 0.17) [ah\ vale]	219	
188		220	
189		221	
190		222	
191		223	
192	DRO: [si jo et fico] en un hotel en aquí_ 193 (. 0.48) tu t'ago- -- 194 t'agobiaràs com una ostrá\ 195 AFT: va:[le] 196 DRO: [a ve(u)]re\ 197 {(L2) ojo\ 198 si em dieu que_ 199 (. 0.11) teniu moltes ganes de relax_ 200 (. 0.38) i només que	224	
		225	
		226	
		227	
		228	
		229	
		230	DRO: [us trobareu molt agobiades]

231	(. 0.28) vale/	233	(. 0.28) [Playa Bávaro_]
232	(.. 0.56) i després està la zona de La Romana_	234	AFT: [aquesta zona] és la que he sentit jo parlar\

En una transcripció discursiva de nivell pròpiament intermedi disposarem el text en format vertical, com a (108), i seguirem la majoria de convencions de la modalitat intermèdia-estreta que hem descrit a § 4.4, però prescindirem d'aquelles que no aportin informació prou rellevant per a l'anàlisi posterior. Per exemple, si escau, no farem servir xifres per indicar la durada de les pauses, que anotarem només amb els símbols (.), (..), (...) i (...), o no assenyalarem els encavalcaments.

#### 4.5.2 Transcripció estreta

La principal diferència entre una transcripció discursiva de nivell intermedi-estret i una transcripció de nivell pròpiament estret és que aquesta última modalitat adopta tot un seguit de convencions destinades a proporcionar una notació més detallada de l'esdeveniment comunicatiu, sobretot en relació amb els aspectes prosòdics dels grups tonals. A continuació desglossem aquestes convencions (veg. també Du Bois *et al.* 1993 i Payrató 1995).

1. *Disposició del text dins la pàgina.* Com la transcripció discursiva de nivell intermedi-estret, la transcripció estreta segmenta el discurs oral en grups entonatius i els distribueix en columna. Ara bé, com hem apuntat més amunt (veg. § 4.4.7), un dels inconvenients d'aquest format és que en cas d'encavalcament intern es fa difícil mantenir l'ordre cronològic de la parla. La transcripció estreta resol aquest problema utilitzant el símbol &, que indica continuïtat de torn, i dividint el grup tonal afectat en dues línies (vegeu-ne exemples a § 4.4.7), d'aquesta manera reproduïm l'encavalcament sense alterar l'ordre cronològic dels enunciats. El mateix símbol ens servirà per reprendre grups entonatius amb una pausa interna molt

breu que l'interlocutor aprofita per intervenir, vegeu-ne un exemple a § 4.4.6.

2. *Descripció de l'escena.* La transcripció discursiva estreta, com la de nivell intermedi-estret, aplicarà el model etnogràfic *Speaking* per descriure el context comunicatiu en què es desenvolupa el fet de parla.

3. *Aspectes prosòdics.* La transcripció estreta permet reflectir alguns fenòmens prosòdics que la transcripció intermèdia-estreta omet, com ara els anomenats *grups tonals menors* o *sintagmes fonològics* (veg. Prieto 2002a: 39, 2002b: 396) i els accents primaris i secundaris. Hem de comptar, però, que recollir aquesta informació demanarà més temps i esforç que els que exigeix una modalitat intermèdia-estreta, per la qual cosa abans de començar a transcriure amb més convencions prosòdiques valorarem detingudament si realment necessitem una notació tan exhaustiva.

Al contrari que els grups tonals majors, que són els que reproduïx la transcripció intermèdia-estreta, els grups tonals menors presenten límits poc definits, perquè no solen coincidir amb pauses i moviments melòdics més o menys pronunciats, sinó amb lleugeres inflexions del to a l'interior de l'enunciat (veg. Prieto 2002a, 2002b: 396). La segmentació del discurs oral en grups tonals menors també dependrà de l'idiolecte del parlant i del registre —que poden afavorir un ritme d'elocució més ràpid o més lent—, així com de factors sintàctics, ja que com més llarg és l'enunciat i més constituents conté, més probabilitats hi ha que estigui format per dos o més grups tonals menors. A més, alguns constituents sintàctics requereixen almenys un grup tonal menor, com ara els vocatius i els elements d'una enumeració. Indicarem la frontera entre grups entonatius menors amb el símbol /, com a (109). En aquest cas trobem un grup tonal major relativament llarg que segmentem en dos grups tonals menors perquè la primera part de l'enunciat presenta un contorn entonatiu descendent que es resstia lleugerament a partir de la preposició *per*.

(109) 189 DRO: (. 0.25) perqu(è) és molt recomanable | per una parella\  
(COR: EMP4, adaptat)

La distinció entre grups entonatius menors i majors està molt lligada als anomenats *accents tonals*, que són inflexions melòdiques ancorades en determinats accents de mot. Tot grup tonal conté un accent prominent o primari, que representarem amb el símbol \* davant la síl·laba afectada, i opcionalment un o més accents secundaris que se subordinen a l'accent principal, i que reproduïrem amb el símbol •. En català, l'accent tonal primari recau generalment en la darrera síl·laba accentuada de la unitat entonativa i es caracteritza sobretot perquè acull moviments de to amb significat pragmàtic. La resta d'accents tonals presenten inflexions menys perceptibles i es reparteixen al llarg de la cadena fònica segons la intenció comunicativa del parlant. Una elocució expressiva, per exemple, tendirà a reduir el nombre d'accents secundaris (veg. Prieto 2002a, 2002b: 398). Així doncs, si a (109) hi afegim els accents tonals, obtindrem la transcripció de (110).

(110) 189 DRO: (. 0.25) per•qu(è) és molt recoma\*nable | per una pa\*rella\  
(COR: EMP4, adaptat)

A efectes d'interpretació pragmàtica, són especialment rellevants els denominats *tons de frontera* (veg. Prieto 2002a: 52, 2002b: 398), moviments melòdics que es produeixen entre l'accent tonal més prominent de l'enunciat, que en català habitualment coincideix amb l'última síl·laba accentuada, i el final del mateix enunciat, que és on es concentren les seqüències tonals terminals. Du Bois *et al.* (1993) proposen reflectir aquestes inflexions combinant els símbols de les seqüències tonals terminals de la manera següent: descens (∖), ascens (/), manteniment (—), descens-ascens (∨) i ascens-descens (∧); i situar-los davant del mot que conté l'accent tonal més prominent sense separar-los del text amb un espai, tal com exemplifiquen per a l'anglès *If you Vthink about it i A lot of it's really Λbad*. Tanmateix, aquí recomanem que aquests

símbols es reservin per a transcripcions estretes en què la informació prosòdica ha de ser particularment detallada, ja que és una notació que altera força la unitat del grup entonatiu. Recordem, a més, que ja disposem de la notació estreta de les seqüències tonals terminals (veg. § 4.4.2, taula 4.1), que ens serveix per representar el moviment melòdic de final d'enunciat amb més precisió que la notació de nivell intermedi-estret.

D'altra banda, la proposta de Du Bois *et al.* (1993) incorpora tot un seguit de símbols i codis de caràcter més interpretatiu que completen els que es refereixen a aspectes més pròpiament prosòdics. En destaquem dos tipus:

a) Els símbols que indiquen si el parlant té la intenció de continuar el discurs (,) o aturar-lo (.) o si apel·la a l'interlocutor (?). Aquesta notació més funcional es combina amb els símbols de les seqüències tonals terminals. Per exemple, a (111a) DRO acaba el grup entonatiu amb to descendent, però mostra la voluntat de mantenir el torn de parla per donar més detalls sobre la informació que li han demanat les interlocutores, tal com expressa la coma abans del símbol de descens tonal; a (111b: 165) AFT fa un incís enmig de l'explicació de DRO sense la voluntat de mantenir el torn, cosa que representem amb el punt; i a (111c: 144) AFT tanca la seva petició d'informació amb un manteniment del to que exhorta DRO a respondre, i per això el signe d'interrogació.

(111a) 145 DRO: jo ara us explico,\  
(COR: EMP4, adaptat)

(111b) 164 DRO: (.. 0.45) e:\_  
165 AFT: més turística.\  
166 DRO: una zona {(F) molt} turística\_  
(COR: EMP4, adaptat)

(111c) 138 AFT: (... 1.08) i:\_  
139 bueno\  
140 també tinc entès\_  
141 que hi han:\_

142           diferents zones\_  
 143           la zona de l'est\_  
 144           (.. 0.49) La Roma:na i tot això?\_  
 (COR: EMP4, adaptat)

b) El codi que identifica els grups tonals de caràcter parentètic, *{{(PAR) text afectat}*, veg. (112a: 149), i el que marca els grups tonals que formen part d'una citació en discurs directe, *{{(DD) text afectat}*, veg. (112b: 587-591).

(112a) 145 DRO: jo ara us explico\  
 146           (.. 0.49) (a)nem a ve(u)re\  
 147           (... 1.52) a ve(u)re\  
 148           (... 1.52) agafo un prospecte:\_  
 149           {{(PAR) si és que n'hi ha\_  
 (COR: EMP4, adaptat)

(112b) 586 DRO: (. 0.16) pues tu em diràs\_  
 587           {{(DD) (.. 0.71) vinga xata\  
 588           (.. 0.36) vaia un puesto que m'ha(s) porta(t)\  
 589           aquí:\_  
 590           només que parelles\_  
 591           {{(L2) vaya un agobio\_  
 592           clar\  
 (COR: EMP4, adaptat)

No obstant això, considerem que són convencions massa interpretatives —no sempre és fàcil determinar si el parlant vol continuar el discurs o no o fins i tot si reproduïx les paraules d'algú altre (veg. Maldonado 1999)— i només les utilitzarem si són rellevants per a l'anàlisi a què va destinada la transcripció.

Finalment, en la modalitat estreta també farem servir, si convé, els codis prosòdics següents:

- a) Ritme d'elocució molt accelerat: *{{(ACC) text afectat}*  
 Ritme d'elocució molt desaccelerat: *{{(DCC) text afectat}*
- b) Intensitat molt forta: *{{(FF) text afectat}*  
 Intensitat molt fluixa: *{{(PP) text afectat}*
- c) Intensitat *in crescendo*: *{{(CR) text afectat}*  
 Intensitat *in decrescendo*: *{{(DR) text afectat}*

4. *Encavalcaments i pauses*. La proposta de Du Bois *et al.* (1993) alinea els claudàtors dels encavalcaments per facilitar la lectura del text, com il·lustra (113: 184-185), però el sistema de transcripció discursiva del CCCUB deixa aquest recurs, que demana força precisió, per a la modalitat estreta.

(113) 183 T??: (.. 0.37) @R(1.18)R@  
 184 DRO: és molt recoma[nable\_]  
 185 AFT: [xungo\  
 186 DRO: per una pare- --  
 (COR: EMP4, adaptat)

Pel que fa a les pauses, en la modalitat estreta emprarem la mateixa notació que en la transcripció de nivell intermedi-estret (veg. § 4.4.6). Tot i així, si necessitem mesuraments més curts, podem optar per les convencions de la proposta estreta de Du Bois *et al.* (1993), recollides a la taula 4.18 (en segons).

<b>Breu</b>	0,2 > pausa	(. durada)
<b>Mitjana</b>	0,2 ≤ pausa < 0,7	(.. durada)
<b>Llarga</b>	0,7 ≤ pausa < 1	(... durada)
<b>Molt llarga</b>	1 ≥ pausa	(... durada)

Taula 4.18 Notació estreta de la durada de les pauses

Òbviament, si fem servir aquests mesuraments, haurem d'advertir-ho en els annexos de la transcripció per evitar confusions amb la notació de pauses de nivell intermedi-estret, que són les que triarem per defecte.

Recordem també que en una transcripció més detallada podem assenyalar que un interlocutor enllaça el seu torn de parla amb el torn anterior —sense pausa perceptible però tampoc encavalcament— mitjançant el símbol (0), col·locat a l'inici del primer grup entonatiu; vegeu-ne un exemple a (47). Reservarem aquest símbol, poc habitual, per a les transcripcions de modalitat estreta.

5. *Regularitzacions*. Respecte a l'adaptació de grafies no estàndards, la transcripció estreta no difereix gaire de la

transcripció de nivell intermedi-estret. Si convé, però, admet un ús menys restringit dels parèntesis, que podrem utilitzar fins i tot en contextos relativament predictibles com (114), en què assenyallem l'elisió de la e tònica.

(114) 181 DRO: no (é)s recomanable\_  
(COR: EMP4, adaptat)

La modalitat estreta també accepta més fàcilment fragments en AFI al costat de la transcripció ortogràfica i, per tant, si cal, inserirem en el text representacions fonètiques mitjançant el codi *(/transcripció fonètica/)*. Aquest recurs pot ser molt útil per resoldre regularitzacions especialment complicades, com ara l'*ojo* de (115: 197), pronunciat [••••], no pas [••••] (veg. § 4.4.15.5) —una altra possibilitat seria fer servir el codi *{(L2) text afectat}*.

(115) 196 DRO: a ve(u)re\  
197 ojo\ (/••••/)  
198 si em dieu que\_  
(COR: EMP4, adaptat)

6. *Aspectes vocals, gestuals i contextuals.* En una transcripció de modalitat estreta tenim l'opció de distingir diferents tipus de riure a partir dels codis previstos pel mateix sistema. Per exemple, podem diferenciar la riallada, que reproduïrem amb el codi *{(F) @R(durada)R@}* de la rialla nasal, molt menys estrident, i representable per *{(nasal) @}*.

Pel que fa als aspectes gestuals, la transcripció estreta compta amb el codi *{(X= fenomen) text X text X}*, que serveix per reflectir fenòmens gestuals esporàdics, no simultanis i repetits al llarg d'un fragment de parla, per exemple, *{(X= mirant al cel) plou/ X o no plou/ X}*.

I també podem recórrer més sovint als comentaris del transcriptor que en una transcripció de modalitat intermèdia-estreta —recordem que a § 4.4.13 en recomanàvem un ús prudencial per no carregar massa el text— i, si convé, disposar-los en una columna paral·lela al text per no dificultar la lectura.



A continuació reprenem el fragment de (107) i (108) en una versió estreta en què anotem les seqüències tonals terminals d'acord amb els símbols de la taula 4.1 (veg. § 4.4.2), així com els grups tonals majors i menors i els accents primaris i secundaris.

(116) Descripció de l'escena a partir del model Speaking (veg. Alturo *et al.* 2004)

138 AFT: (... 1.08) i:→  
 139 \*bueno\  
 140 també tinc en\*tès→  
 141 que \*hi han:→  
 142 diferents \*zones→  
 143 la zona de \*l'est/  
 144 (.. 0.49) La Ro\*ma:na | i [tot a\*ixò/  
 145 DRO: [jo ara] us ex\*plico\  
 146 (.. 0.49) (a)nem a \*ve(u)re\  
 147 (... 1.52) a \*ve(u)re\  
 148 (... 1.52) agafo un pros\*pecte:\  
 149 si és que \*n'hi ha\  
 150 (. 0.22) \*bueno\  
 151 (.. 0.92) a \*ve(u)re\  
 152 (... 1.22) vosaltres te\*niu→  
 153 (. 0.26) a::→  
 154 (.. 0.37) la \*zona→  
 155 (. 0.23) de Puerto \*Pla:ta/  
 156 So\*súa/  
 157 tot això d'a\*quí\ ((assenyala el catàleg))  
 158 (.. 0.52) que \*és→  
 159 di\*guem-ne\  
 160 la que:→  
 161 (. 0.12) la \*zona→  
 162 (.. 0.31) {(??) lla\*vors→}  
 163 (a)nem a •dir que els ho•tels són més asse\*quible:s\  
 164 (.. 0.45) e:→  
 165 AFT: (0) més tu\*rística\  
 166 DRO: una •zona {(F) molt} tu\*rística\  
 167 (.. 0.82) que hi han {(F) molts} ho\*tels\  
 168 (.. 0.60) i que tam•bé hi ha més anima\*ció\  
 169 (.. 0.67) d'a\*cord//

170 (.. 0.77) després es\*tà→  
171 (. 0.18) la •zona de Sama\*nà\  
172 (. 0.12) que és una de les \*zones→  
173 \*més→  
174 e::→  
175 (.. 0.52) més \*verges\  
176 que \*hi ha\  
177 a tota la \*illa\  
178 (.. 0.50) \*p(e)rò--  
179 molt \*maca\  
180 (.. 0.46) \*p(e)rò--  
181 no (é)s recoma\*nable\  
182 (. 0.26) per dugues \*noies↓  
183 T??: (.. 0.37) @R(1.18)R@  
184 DRO: és molt recoma[\*nable→]  
185 AFT: [\*xungo↓]  
186 DRO: per una pa\*re- --  
187 no\  
188 no per \*xungo\  
189 (. 0.25) per•qu(è) és molt recoma\*nable | per una pa\*rella↓  
190 AFT: (. 0.17) [\*ah\  
191 \*vale↓]  
192 DRO: [si jo et •fico] en un hotel en a\*quí→  
193 (.. 0.48) tu t'ago- --  
194 t'agobia•ràs com una \*ostralla\  
195 AFT: \*va:[le\  
196 DRO: [a \*ve(u)]re\  
197 \*ojol\ (/•••/)  
198 si em di\*eu que→  
199 (. 0.11) teniu •moltes ganes de re\*lax\  
200 (.. 0.38) i només que re\*lax\  
201 GEF: {( @) entre poc ni \*massa\  
202 AFT: què vols \*dir//  
203 que és més paisat\*gísti:c/  
204 i ro\*mà:nti:c/  
205 [i:→]  
206 DRO: [exac]ta\*ment\  
207 i és estar-hi amb un \*hotel→  
208 am(b) el •règim a\*quest | del {(L2) todo inclu\*ido}  
209 que si no tens •ganes de sortir de l'ho\*tel→  
210 no cal que \*surtis\  
211 per•què ho tens \*tot\  
212 (. 0.13) p(e)rò tot ho [1 troba•ràs al teu]

213 GEF: [1 no\\  
 214 no x↓ @]  
 DRO: & vol\*tant\\  
 215 AFT: (. 0.28) [2 no\\  
 216 no\*satro- --]  
 217 DRO: [2 són pa\*re]lles\\  
 218 AFT: ens inte•ressa també co•nèixer una \*mi[ca:→ 3]  
 219 GEF: [3 \*clar\\]  
 220 DRO: (.. 0.84) alla•vons a\*quí\\  
 221 no podeu a\*nar↓  
 222 \*és→  
 223 (.. 0.41) {(DC)(F) pre\*ció:s}  
 224 (. 0.20) p(e)rò no podeu a\*nar\\  
 225 perquè us agobia\*reu↓  
 226 (. 0.13) perquè •tot lo que hi ha al vol\*tant\\  
 227 (. 0.11) són pa\*relles\\  
 228 (.. 0.34) llavons \*clar\\  
 229 T??: (. 0.25) @@ [@R(2.31)R@]  
 230 DRO: [us trobareu molt ago\*biades↓]  
 231 (. 0.28) \*vale//  
 232 (.. 0.56) i després es•tà la zona de La Ro\*mana\\  
 233 (. 0.28) [Playa \*Bávaro]  
 234 AFT: [aquesta \*zona] | és la que he sen•tit jo par\*lar\\

### 4.5.3 Recursos complementaris

La bibliografia proposa altres recursos que ens poden ser útils per recollir certs fenòmens discursius en qualsevol de les tres modalitats de transcripció que aquí hem descrit (veg. Du Bois *et al.* 1993, Payrató: 1995). A continuació en destaquem alguns:

- a) Claudàtors dobles, *[[[text afectat]]]*, i triples, *[[[[[text afectat]]]]]*, per ajudar el lector a aparellar encavalcaments.
- b) El símbol d'igualtat (=) per representar allargaments (per exemple, *i=*).
- c) Ús de majúscules en la notació de sons vocals no verbals i comentaris del transcriptor per diferenciar-los

clarament de la resta del text, per exemple, (TOS), ((SURT DE L'HABITACIÓ)).

d) En cas de dubte entre dues possibles transcripcions d'un mateix fragment de parla, anotar-les totes dues —una entre parèntesis. Aquí, però, recomanem triar la solució que considerem més probable i marcar-la amb el codi {{(??) text afectat}}.

e) Indicar la durada d'una exhalació, un gest, etc., fent servir codis, per exemple, {{(0.12) (EXH)}}, {{(0.78) aplaudint}}.

f) Ús de majúscula inicial en el primer mot de cada conjunt de grups tonals amb un contorn prosòdic particular. Per exemple, vegeu que el fragment de (117), extret d'una visita guiada en un museu, es pot dividir en tres unitats discursives, cadascuna amb una prosòdia característica: presentació d'un esdeveniment (g. t. 778-784), amb manteniment i descens del to al final; pregunta retòrica per introduir una explicació sobre aquest mateix fet (g. t. 785), amb manteniment-ascens; i resposta a la pregunta (g. t. 786-790), amb manteniment i descens.

(117) 778 GIA: (.. 0.32) A l'any\_  
779 mil nou-cents trenta-set\_  
780 (. 0.17) Matta\_  
781 (. 0.23) deixa Londres\_  
782 i torna\_  
783 a París\_  
784 per treballar amb Josep Lluís Sert\  
785 Per què\_  
786 (. 0.27) Perquè l'any trenta-set\_  
787 té lloc\_  
788 l'exposició\_  
789 internacional\_  
790 de París\  
(COR: CULT2, adaptat)

Lògicament, es tracta d'una convenció molt interpretativa que només utilitzarem quan sigui pertinent per a l'anàlisi de la transcripció.

g) Parèntesis triangulars, <text afectat>, per assenyalar falsos inicis, com a (118: 1856). Sovint, però, resulta difícil

detectar aquest tipus de grups tonals sense una interpretació aprofundida del text, per la qual cosa destinarem aquesta convenció a estudis en què els falsos inicis es considerin una informació rellevant.

(118) 1854 CME: què ha de tenir quinze anys\  
1855 PPP: no\  
1856 MMM: <en tenia-->  
1857           encara no havia fet la comunió:\  
(COC: 07, adaptat)

*h)* Deixar interlineat doble entre grups tonals per incloure la traducció del text, si n'hi ha. La traducció anirà amb un tipus de lletra diferent (cursiva, cos menor, etc.) per facilitar la lectura del text.

*i)* Fer servir els codis *{(L3) text afectat}*, *{(L4) text afectat}*, etc. en cas que els informants utilitzin dues o més llengües diferents de la llengua base de la transcripció.

*j)* Distribuir en columnes paral·leles dues transcripcions diferents del mateix fragment de parla, com ara una transcripció discursiva i la transcripció fonètica corresponent (vegeu-ne una mostra a Payrató i Alturo 2002, en què els primers cinc minuts de cada conversa també estan transcrits en AFI). En aquest cas haurem de vigilar que la numeració i el text d'una transcripció coincideixin amb la numeració i el text de l'altra. Si una de les transcripcions és fonètica, tenim la possibilitat d'incloure-hi els codis i els símbols propis de la transcripció discursiva, tal com es fa a Payrató i Alturo (2002), o bé ometre'ls si considerem que dificulten massa la lectura.

*k)* El símbol del dòlar (\$) per assenyalar que un grup tonal conté un element rellevant per a l'estudi del fet de parla. Preferentment, anirà a l'inici de grup entonatiu.



## 5. TRANSCRIPCIÓ, REVISIÓ I EDICIÓ

En aquest últim capítol fem algunes recomanacions que poden ser útils tant al llarg de la transcripció com en la fase de revisió i edició del text. Abans, però, cal recordar que anirem fent còpies de seguretat regularment del document amb què treballem per no perdre informació en cas que tinguem alguna incidència amb el sistema informàtic. També ens assegurarem des de bon principi que l'enregistrament és prou rellevant i té una qualitat de so prou bona per ser transcrit. Si no, correm el risc d'haver-nos de plantejar més endavant si continuem transcrivint l'enregistrament sencer, o els fragments que en puguem aprofitar, o si l'hem de descartar del tot i perdre, doncs, la feina que hàgim fet.

Pel que fa al procés de transcripció, una bona estratègia és organitzar-se la feina en etapes per anotar cada tipus de convenció per separat, per exemple: primer segmentar el discurs en grups tonals, després afegir-hi els encavalcaments, a continuació les pauses, etc. Així aconseguim una transcripció més fiable, perquè si volem apuntar totes les convencions alhora, és molt fàcil que se'ns passi per alt més d'un detall, i d'aquesta manera també facilitem que dos o més transcriptors puguin treballar simultàniament en el mateix text, cadascun dedicat a un aspecte diferent. No hi ha cap ordre establert a l'hora d'anotar els diversos tipus de convencions i, per tant, serà cada transcriptor el que decidirà com es distribueix la feina. Sembla lògic, però, començar per la segmentació en grups entonatius i l'anotació de les seqüències tonals terminals, i tot seguit continuar amb els encavalcaments i les pauses —que són les convencions que solen demanar més temps— i la resta de símbols i codis. Igualment, aconsellem que es revisi la transcripció per etapes i que ens fixem primer en un aspecte concret i, en acabat, passem a un altre. Això sí, un cop revisat

cada tipus de convenció per separat, és convenient fer una darrera revisió de conjunt per assegurar-nos que no hi ha incoherències (per exemple, en la distribució dels torns de parla) i no ens hem deixat cap correcció per fer.

En cas que trobem fragments especialment difícils de transcriure, procurarem no aturar-nos-hi gaire, tenint en compte que podrem tornar-hi més endavant, en la fase de revisió. Un cop transcrit tot l'enregistrament o la part que ens interessa més, és quan disposem de més dades contextuals i reconeixem més fàcilment les veus dels informants en cas que només treballem amb so. Aquest coneixement més aprofundit de l'esdeveniment comunicatiu ens ajudarà molt a afinar la transcripció o almenys resoldre alguns punts conflictius que havíem deixat pendants. També hem de tenir present que no hem de traslladar al paper tota la informació que puguem —cosa d'entrada impossible—, sinó sobretot aconseguir una transcripció com més acurada millor que respongui als propòsits de l'anàlisi a què va destinada, i fer-ho sense simplificar massa el text però sense atapeir-lo innecessàriament.

Si es treballa en grup, és molt important anar recollint i comentant les incidències que es produeixen al llarg de la transcripció per unificar criteris —com més transcriptors treballen alhora, més diferències sol haver-hi entre les solucions que s'adopten en cada moment— i no haver de repassar tota la transcripció o conjunt de transcripcions per introduir-hi els canvis que s'hagin acordat *a posteriori*.

Un cop acabada la transcripció, comença tot un procés de revisió i edició del text tan important com la feina anterior. Lluny de ser una tasca secundària, la revisió final esdevé fonamental tant per corregir fragments que no s'han transcrit correctament —perquè no els havíem entès bé, per exemple— com per detectar possibles incoherències que durant el procés de transcripció se'ns han passat per alt, com ara regularitzacions diferents per a una mateixa solució no estàndard. Aconsellem, per tant, que s'encari la revisió del text no pas com un repàs d'última hora per endreçar-ho tot una mica, sinó sobretot com



una manera de polir una feina que ens haurà demanat un esforç considerable.

D'entrada, recomanem que es deixi passar un cert temps entre la transcripció pròpiament dita i la revisió. Si s'ha treballat intensament en un text durant molt de temps, és probable que en alguns fragments acostumem l'orella al que creiem que han dit els informants més que no pas a les paraules que hagin pronunciat realment. Per tant, si és possible, deixarem *reposar* la transcripció per deshabituar-nos del que hem escoltat repetidament durant un temps i tornar al text més endavant per llegir-lo i revisar-lo com si fos la primera vegada que l'escoltem, amb l'orella menys condicionada per les audicions anteriors. Amb aquest procediment sovint és més fàcil acabar d'entendre —o almenys en part— fragments que en un principi semblaven inintel·ligibles i adonar-se d'errades de transcripció, ortogràfics, de tecleig, etc. que s'hagin pogut produir. Mentrestant, dedicarem aquest impàs a tasques més rutinàries, com ara redactar o repassar la descripció contextual que precedeix la transcripció, comprovar que hem recollit tota la informació necessària tant de la interacció com dels participants (data i lloc d'enregistrament, dades dels informants, permisos, etc.) o fer un repàs general per esmenar errades tipogràfiques o de format freqüents, com ara tabuladors mal alineats, numeració alterada, inicials de parlants mal col·locades, etc. Alhora ens assegurarem que hem eliminat tant del text com de l'enregistrament totes aquelles dades que poden comprometre l'anonimat dels informants.

També és molt recomanable que almenys una segona persona revisi la transcripció. L'opció ideal és que un transcriptor amb experiència treballi en el mateix text per garantir que es reproduïx l'esdeveniment comunicatiu amb la màxima fidelitat. Molt sovint, però, això no és possible, i haurèm de buscar una segona opinió que ens ajudi a polir la transcripció, com a mínim en els fragments més problemàtics. D'aquesta manera podrem acabar d'entendre fragments de comprensió difícil o, ben al contrari, assenyalar com a dubtosos punts que en un primer moment ens semblaven força clars.

Per acabar, i pel que fa a l'edició del text, és molt útil fer una lectura acurada de la transcripció sense suport de l'àudio o la imatge —i preferiblement en paper més que no pas en la pantalla de l'ordinador— per corregir les errades tipogràfiques o ortogràfiques que encara hi pugui haver. També és un bon moment per comprovar que seguim les pautes de regularització de solucions no estàndards i de format (tabuladors ben alineats, línies numerades, números de pàgina, etc.).

## CLOENDA

“Transcription is an open-ended process. A transcript changes as the researcher's insights become progressively refined”,  
Jane A. Edwards (2003: 321).

Probablement mai no es pot donar per acabada una transcripció. Cada vegada que confrontem el so o les imatges de l'enregistrament amb el text ja transcrit ens adonarem de petits detalls que se'ns hauran escapat en revisions anteriors: un riure que no havíem anotat, un encavalcament que ha quedat massa curt, un fragment que crèiem intel·ligible i que ara entenem, etc. Per això sempre hi haurem de fer modificacions —si és possible— i deixar el text obert a noves interpretacions i canvis que suggereixin altres transcriptors o els investigadors que l'estudien en profunditat.

Com apuntàvem al principi d'aquest manual, no hem de perdre de vista que tota transcripció no és res més que un filtre que trasllada al paper un esdeveniment comunicatiu que per definició resulta efímer i irrepetible, fins i tot si en tenim imatges enregistrades. No obstant això, tant la provisionalitat com la imperfecció que suposa tota transcripció, així com l'esforç que implica dur-la a terme, es veuen àmpliament compensats per les enormes possibilitats que obre per a la investigació lingüística, sobretot en el camp de la llengua oral, i pels resultats que finalment se n'obtenen.



## BIBLIOGRAFIA

### BIBLIOGRAFIA CITADA

- ALBELDA MARCO, MARTA (2005): "Sistemas de transcripción de los corpus orales del español", dins Carrió Pastor, María Luisa (ed.): *Perspectivas interdisciplinarias de la lingüística aplicada*, Universitat Politècnica de València/Aesla, València, vol. 2, pàg. 381-387.
- ALTURO, NÚRIA; BLADAS, ÒSCAR; PAYÀ, MARTA i LLUÍS PAYRATÓ (ed.) (2004): *Corpus Oral de Registres. Materials de treball*, Publicacions i Edicions de la Universitat de Barcelona, Barcelona.
- ALTURO, NÚRIA; BOIX, EMILI i MARIA PILAR PEREA (2002): "Corpus de Català Contemporani de la Universitat de Barcelona (CUB): A general presentation", dins Pusch, Claus D. i Wolfgang Raible (ed.): *Romanistische Korpuslinguistik – Korpora und gesprochene Sprache. Romance Corpus Linguistics – Corpora and Spoken Language*, Script Oralia, 126, Narr, Tübingen, pàg. 155-170.
- ATKINSON, J. MAXWELL i JOHN HERITAGE (ed.) (1984): *Transcript notation. Structures of social action: Studies in conversation analysis*, Cambridge University Press, Cambridge.
- BILGER, MIREILLE (2002): "Corpus de Français Parlé: recueil et analyse", dins Pusch, Claus D. i Wolfgang Raible (ed.): *Romanistische Korpuslinguistik – Korpora und gesprochene Sprache. Romance Corpus Linguistics – Corpora and Spoken Language*. Script Oralia, 126, Narr, Tübingen, pàg. 45-58.
- BLANCHE-BENVENISTE, CLAIRE i COLETTE JEANJEAN (1987): *Le français parlé. Transcription et édition*, Didier, París.
- BOIX, EMILI i FRANCESC XAVIER VILA (1998): *Sociolingüística de la llengua catalana*, Ariel, Barcelona.
- BOIX-FUSTER, EMILI; ÀLAMO, MARINA; GALINDO, MIREIA i F. XAVIER VILA (2006): *Corpus de varietats socials. Materials de treball*, Publicacions i Edicions de la Universitat de Barcelona, Barcelona.
- BONET, M. EULÀLIA; LLORET, MARIA ROSA i JOAN MASCARÓ (1997): *Manual de transcripció fonètica*, Universitat Autònoma de Barcelona, Bellaterra.
- BRIZ, ANTONIO (1996): "El corpus de conversació coloquial Val.Es.Co", dins Payrató, Lluís; Boix, Emili; Lloret, Maria Rosa i Mercè Lorente (ed.): *Corpus, corpora*, PPU, Barcelona, pàg. 253-295.
- BRIZ, ANTONIO i GRUP VAL.ES.CO (2002): "La transcripció de la lengua hablada: el sistema del grupo Val.Es.Co", *Español Actual*, 77-78, pàg. 57-85.
- BRUCE, JACKSON (1987): *Fieldwork*, MIT Press, Cambridge, Massachusetts.

- BUTTON, GRAHAM i JOHN R. E. LEE (ed.) (1987): *Talk and social organisation, Multilingual Matters*, Clevedon.
- CALSAMIGLIA, HELENA; COTS, JOSEP M.; LORDA, CLARA U.; NUSSBAUM, LUCI; PAYRATÓ, LLUÍS i AMPARO TUSÓN (1997): *La parla com a espectacle. Estudi d'un debat televisiu*, Publicacions de la Universitat Autònoma de Barcelona, Bellaterra.
- CALSAMIGLIA, HELENA i AMPARO TUSÓN (1999): *Las cosas del decir. Manual de análisis del discurso*, Ariel, Barcelona [2a ed., 2007].
- CASTELLÀ, JOSEP MARIA (2004): *Oralitat i escriptura. Dues cares de la complexitat del llenguatge*, Publicacions de l'Abadia de Montserrat, Barcelona.
- CHAFE, WALLACE L. (1979): "The flow of Thought and the Flow of Language", dins Givón, Talmy (ed.): *Syntax and Semantics. Discourse and Syntax*, Academic Press, Nova York/Londres, pàg. 159-181.
- CTILC = INSTITUT D'ESTUDIS CATALANS: Corpus Textual Informatitzat de la Llengua Catalana [en línia], Institut d'Estudis Catalans, Barcelona [consulta: abril 2009]. Disponible a <<http://ctilc.iec.cat/>>.
- CUENCA, MARIA JOSEP (2002): "Els connectors textuais i les interjeccions", dins Solà, Joan; Lloret, Maria Rosa; Mascaró, Joan i Manuel Pérez Saldanya (dir.): *Gramàtica del català contemporani*, Empúries, Barcelona, vol. 3, pàg. 3173-3237.
- \_\_\_\_\_ (2006): *La connexió i els connectors. Perspectiva oracional i textual*, Eumo, Vic.
- CUENCA, MARIA JOSEP i MARTA TORRES VILATARSANA (2008): "Usos de *hombre/home* y *mujer/dona* como marcadores del discurso en la conversación coloquial", *Verba*, 35, p. 235-256.
- DCVB = ALCOVER, ANTONI MARIA i FRANCESC DE BORJA MOLL (1930-1962): *Diccionari català-valencià-balear* [en línia], Institut d'Estudis Catalans, Barcelona [consulta: maig 2009]. Disponible a <<http://dcvb.iec.cat/>>.
- DIEC2 = INSTITUT D'ESTUDIS CATALANS: *Diccionari de la llengua catalana* [2a ed., en línia], Institut d'Estudis Catalans, Barcelona [consulta: maig 2009]. Disponible a <<http://dlc.iec.cat/>>.
- DOMCE = RIERA EURES, MANUEL i MARGARIDA SANJAUME (2002): *Diccionari d'onomatopeies i mots de creació expressiva*, Edicions 62, Barcelona.
- DU BOIS, JOHN W. (1991): "Transcription design principles for spoken discourse research", *Pragmatics*, 1 (1), pàg. 71-106.
- DU BOIS, JOHN W.; SCHUETZE-COBURN, STEPHAN; CUMMING, SUSANNA i DANAE PAOLINO (1991): *Discourse transcription*, Longman, London.
- \_\_\_\_\_ (1993): "Outline of discourse transcription", dins Edwards, Jane A. i Martin D. Lampert (ed.): *Talking data: transcription and coding in discourse research*, Lawrence Erlbaum, Hillsdale, Nova Jersey, pàg. 45-89.
- DURANTI, ALESSANDRO (1997): *Antropología lingüística*, Cambridge University Press, Madrid [trad., 2000].

- EDWARDS, JANE A. (1993): "Principles and Contrasting Systems of Discourse Transcription", dins Edwards, Jane A. i Martin D. Lampert (ed.): *Talking data: Transcription and coding in discourse research*, Lawrence Erlbaum, Hillsdale, Nova Jersey, pàg. 3-31.
- \_\_\_\_\_ (2003): "The Transcription of Discourse", dins Schiffrin, Deborah; Tannen, Deborah i Heidi E. Hamilton (ed.): *The Handbook of Discourse Analysis*, Blackwell, Oxford.
- EDWARDS, JANE A. i MARTIN D. LAMPERT (ed.) (1993): *Talking data: Transcription and coding in discourse research*, Lawrence Erlbaum, Hillsdale, Nova Jersey.
- EHLICH, KONRAD (1993): "A Transcription System for Discourse Data", dins Edwards, Jane A. i Martin D. Lampert (ed.): *Talking data: Transcription and coding in discourse research*, Lawrence Erlbaum, Hillsdale, Nova Jersey, pàg. 123-148.
- FITÓ PARDO, JAUME (2009a): *Multimodalitat i dixi d'espai en joves catalans multilingües*, Universitat de Barcelona, tesi doctoral en preparació.
- \_\_\_\_\_ (2009b): "El gest i la dixi d'espai en textos instructius", *Caplletra*, 46, pàg. 9-41.
- GDLC = ENCICLOPÈDIA CATALANA: *Gran diccionari de la llengua catalana* [en línia], Enciclopèdia Catalana, Barcelona [consulta: maig 2009]. Disponible a <<http://www.enciclopedia.cat/>>.
- GIBBON, DAFYDD; MOORE, ROGER i RICHARD WINSKI (1998): *Spoken Language Systems and Corpus Design*, Walter de Gruyter, Berlín.
- GONZÁLEZ, MONTSERRAT (2004): *Pragmatic Markers in Oral Narrative*, John Benjamins, Amsterdam/Filadèlfia.
- GOODWIN, CHARLES (1993): "Recording Human Interaction in Natural Settings", *Pragmatics*, 3 (2), pàg. 181-209.
- GREEN, JUDITH; FRANQUIZ, MARIA i CAROL DIXON (1997): "The myth of the objective transcript: Transcribing as a situated act", *TESOL Quarterly*, 31 (1), pàg. 172-176.
- GUMPERZ, JOHN J. i NORINE BERENZ (1993): "Transcribing Conversational Exchanges", dins Edwards, Jane A. i Martin D. Lampert (ed.): *Talking data: Transcription and coding in discourse research*, Lawrence Erlbaum, Hillsdale, Nova Jersey, pàg. 91-122.
- HIDALGO, ANTONIO i JULIA SANMARTÍN (2005): "Los sistemas de transcripción de la lengua hablada", *Oralia*, 8, pàg. 13-36.
- LEHMBERG, TIMM i KAI WÖRNER (2008): "Annotation standards", dins Lüdeling, Anke i Merja Kytö (ed.): *Corpus Linguistics. An International Handbook*, Walter de Gruyter, Berlín/Nova York, vol. 1, pàg. 484-501.
- LLISTERRI, JOAQUIM (1999): "Transcripción, etiquetado y codificación de corpus orales", dins Gómez Guinovart, Xavier; Lorenzo Suárez, Anxo; Pérez Guerra, Javier i Alberto Álvarez Lugrís (ed.): *Panorama de la investigación en lingüística informática, Resla (Revista Española de Lingüística Aplicada)*, volum extraordinari 1, pàg. 53-82.

- LORENTE CASAFONT, MERCÈ (2002): "Altres elements lèxics", dins Solà, Joan; Lloret, Maria Rosa; Mascaró, Joan i Manuel Pérez Saldanya (dir.): *Gramàtica del català contemporani*, Empúries, Barcelona, vol. 1, pàg. 831-888.
- MALDONADO, CONCEPCIÓN (1999): "Discurso directo y discurso indirecto", dins Bosque, Ignacio i Violeta Demonte (ed.): *Gramática descriptiva de la lengua española*, Espasa, Madrid, vol. 3, 3549-3596.
- MCNEILL, DAVID (1992): *Hand and Mind. What Gestures Reveal about Thought*, The University of Chicago Press, Chicago/Londres.
- \_\_\_\_\_ (2005): *Gesture and Thought*, The University of Chicago Press, Chicago/Londres.
- MESTHRIE, RAJEND (ed.) (2001): *Concise Encyclopedia of Sociolinguistics*, Elsevier, Amsterdam.
- O'CONNELL, DANIEL C. i SABINE KOWAL (1994): "Some Current Transcription Systems for Spoken Discourse: A Critical Analysis", *Pragmatics*, 4, pàg. 81-107.
- OCHS, ELINOR (1979): "Transcription as Theory", dins Ochs, Elinor i Bambi B. Schieffelin (ed.): *Developmental Pragmatics*, Academic Press, Nova York, pàg. 43-72.
- PAYÀ, MARTA (2003): "Prosody and Pragmatics in Parenthetical Insertions in Catalan", *Catalan Journal of Linguistics*, 2, pàg. 207-227.
- PAYRATÓ, LLUÍS (1988): *Català col·loquial. Aspectes de l'ús corrent de la llengua catalana*, Servei de Publicacions de la Universitat de València, València [3a ed., 1996].
- \_\_\_\_\_ (1995): "Transcripció del discurso coloquial", dins Cortés, Luis (ed.): *El español coloquial. Actas del I Simposio sobre Análisis del Discurso Oral*, Universitat d'Almeria, Almeria, pàg. 45-70. [Versió catalana a Payrató, Lluís (1996): "Transcripció del discurs col·loquial", dins Payrató, Lluís; Boix, Emili; Lloret, Maria Rosa i Mercè Lorente (ed.): *Corpus, corpora*, PPU, Barcelona, pàg. 181-216].
- \_\_\_\_\_ (ed.) (1998): *Oralment. Estudis de la variació funcional*, Publicacions de l'Abadia de Montserrat-Universitat de Barcelona, Barcelona.
- \_\_\_\_\_ (2003): *Pragmàtica, discurs i llengua oral*, Editorial UOC, Barcelona.
- PAYRATÓ, LLUÍS i NÚRIA ALTURO (ed.) (2002): *Corpus de conversa col·loquial. Materials de treball*, Publicacions de la Universitat de Barcelona, Barcelona.
- PAYRATÓ, LLUÍS; BOIX, EMILI; LLORET, MARIA ROSA i MERCÈ LORENTE (ed.) (1996): *Corpus, corpora*, PPU, Barcelona.
- PAYRATÓ, LLUÍS i JAUME FITÓ (ed.) (2008): *Corpus audiovisual plurilingüe*, Publicacions i Edicions de la Universitat de Barcelona, Barcelona.
- POYATOS, FERNANDO (1994): *La comunicación no verbal*, Istmo, Madrid.
- \_\_\_\_\_ (2002): *Nonverbal Communication across Disciplines*, John Benjamins, Amsterdam/Filadèlfia.
- PRIETO, PILAR (2002a): *Entonació. Models, teoria, mètodes*, Ariel, Barcelona.



- \_\_\_\_\_ (2002b): "Entonació", dins Solà, Joan; Lloret, Maria Rosa; Mascaró, Joan i Manuel Pérez Saldanya (dir.): *Gramàtica del català contemporani*, Empúries, Barcelona, vol. 1, pàg. 395-462.
- ROMERO, CATHERINE; O'CONNELL, DANIEL C. i SABINE KOWAL (2002): "Notation Systems for Transcription: An Empirical Investigation", *Journal of Psycholinguistic Research*, 31 (6), pàg. 619-631.
- SACKS, HARVEY; SCHEGLOFF, EMMANUEL A. i GAIL JEFFERSON (1974): "A simplest systematics for the organization of turn-taking in conversation", *Language*, 50, 4, pàg. 696-735.
- SOLÀ, JOAN; LLORET, MARIA ROSA; MASCARÓ, JOAN i MANUEL PÉREZ SILDANYA (dir.) (2002): *Gramàtica del català contemporani*, Empúries, Barcelona.
- TANNEN, DEBORAH (1984): *Conversational style: Analysing talk among friends*, Ablex, Norwood.
- THIBAUT, PIERRETTE i DIANE VINCENT (1990): *Un corpus de français parlé. Montréal 84: historique, méthodes et perspectives de recherche*, Département de Langues et Linguistique de l'Université Laval, Quebec.
- TUSÓN, AMPARO (1995): *Anàlisi de la conversa*, Ariel, Barcelona.
- VALLDUVÍ, ENRIC (2002): "L'oració com a unitat informativa", dins Solà, Joan; Lloret, Maria Rosa; Mascaró, Joan i Manuel Pérez Saldanya (dir.): *Gramàtica del català contemporani*, Empúries, Barcelona, vol. 2, pàg. 1221-1279.
- VENY, JOAN (1995): *Introducció a la dialectologia catalana*, Enciclopèdia Catalana, Barcelona.
- VENY, JOAN i PONS i GRIERA, LÍDIA (1998): *Atles Lingüístic del Domini Català: etnotextos del català oriental*, Institut d'Estudis Catalans, Barcelona.
- VIAPLANA, JOAQUIM (1996): *Dialectologia*, Universitat de València, València [2a ed., 2002].
- VIAPLANA, JOAQUIM i MARIA PILAR PEREA (ed.) (2003): *Textos orals del català sincronitzats*, PPU, Barcelona.
- VINCENT, DIANE; LAFOREST, MARTY i GUYLAINE MARTEL (1995): "Le corpus de Montréal 1995: Adaptation de la méthode d'enquête sociolinguistique pour l'analyse conversationnelle", *Dialangue*, 6, 29-46.

#### RECURSOS EN LÍNIA CITATS

- ADOBE SYSTEMS: *Adobe Premiere Pro CS4*, San Jose (Califòrnia), Adobe Systems [consulta: maig 2009].  
 Disponible a <<http://www.adobe.com/es/products/premiere/>>.
- CHILDES: *Child Language Data Exchange System*, Pittsburgh, Carnegie Mellon University: Brian McWhinney [consulta: maig 2009].  
 Disponible a <<http://childes.psy.cmu.edu/>>.

- ÉSADIR: *Ésadir. El portal lingüístic de la Corporació Catalana de Mitjans Audiovisuals*, Barcelona, Corporació Catalana de Mitjans Audiovisuals [consulta: maig 2009]. Disponible a <<http://esadir.cat/>>.
- INTERNATIONAL STANDARDS FOR LANGUAGE ENGINEERING: *ISLE. International Standards for Language Engineering*, Pisa, Consiglio Nazionale Ricerche, Istituto di Linguistica Computazionale Antonio Zampolli: Paola Baroni, Francesca Bertagna [consulta: maig 2009]. Disponible a <<http://www.ilc.cnr.it/EAGLES/isle/>>.
- KTH ROYAL INSTITUTE OF TECHNOLOGY: *WaveSurfer*, Estocolm, KTH Royal Institute of Technology, Speech, Music and Hearing, School of Computer Science and Communication [consulta: maig 2009]. Disponible a <<http://www.speech.kth.se/wavesurfer/>>.
- LLISTERRI, JOAQUIM: *Universitat Autònoma de Barcelona, Joaquim Llisterra, Grup de Fonètica*, Bellaterra, Universitat Autònoma de Barcelona: Joaquim Llisterra [consulta: maig 2009]. Disponible a <<http://liceu.uab.es/~joaquim/>>.
- MAX PLANCK INSTITUTE FOR PSYCHOLINGUISTICS: *The Language Archiving Technology Portal*, Nimega, Max Planck Institute for Psycholinguistics, Language Archiving Technology [consulta: maig 2009]. Disponible a <<http://www.lat-mpi.eu/tools/elan/>>.
- MCNEILL LAB: *McNeill Lab. Center for Gesture and Speech Research*, Chicago, Universitat de Chicago: David McNeill [consulta: juny 2009]. Disponible a <<http://mcneilllab.uchicago.edu/>>.
- MICHIGAN CORPUS OF ACADEMIC SPOKEN ENGLISH: *Micase. Michigan Corpus of Academic Spoken English*, Ann Arbor, Universitat de Michigan, English Language Institute [consulta: maig 2009]. Disponible a <<http://lw.lsa.umich.edu/eli/micase/index.htm>>.
- OHIO STATE UNIVERSITY: *ToBI*, Columbus, Ohio State University, Department of Linguistics [consulta: maig 2009]. Disponible a <<http://www.ling.ohio-state.edu/~tobi/>>.
- OPTIMOT: *Optimot, consultes lingüístiques*, Barcelona, Generalitat de Catalunya [consulta: maig 2009]. Disponible a <<http://optimot.gencat.cat>>.
- RECERCAT: *Recercat. Dipòsit de la Recerca de Catalunya*, Barcelona, Consorci Centre de Supercomputació de Catalunya/Consorci de Biblioteques Universitàries de Catalunya [consulta: maig 2009]. Disponible a <<http://www.recercat.net/>>.
- SANTA BARBARA CORPUS OF SPOKEN AMERICAN ENGLISH: *The Santa Barbara Corpus of Spoken American English*, Santa Barbara, Universitat de Califòrnia Santa Barbara, Linguistics UCSB [consulta: maig 2009]. Disponible a <<http://www.linguistics.ucsb.edu/research/transcription.html>>.
- SONY CREATIVE SOFTWARE: *The SoundForge Product Family*, Madison (Wisconsin), Sony Creative Software [consulta: maig 2009]. Disponible

a

- <<http://www.sonycreativesoftware.com/products/soundforgefamily.asp>>.
- SOURCEFORGE: *Transcriber. A tool for segmenting, labeling and transcribing speech*, Mountain View (Califòrnia), Sourceforge [consulta: maig 2009]. Disponible a <<http://trans.sourceforge.net/en/presentation.php>>.
- TEXT ENCODING INITIATIVE CONSORTIUM: *TEI. Text Encoding Initiative*, Charlottesville (Virgínia), Universitat de Virgínia, Institute for Advanced Technology in the Humanities [consulta: maig 2009]. Disponible a <<http://www.tei-c.org>>.
- UNIVERSITAT D'AMSTERDAM: *Praat: doing phonetics by computer*, Amsterdam, Universitat d'Amsterdam, Institute of Phonetic Sciences: Paul Boersma, David Weenink [consulta: maig 2009]. Disponible a <<http://www.fon.hum.uva.nl/praat/>>.
- UNIVERSITAT D'EDIMBURG: *NITE XML Toolkit Homepages*, Edimburg, Universitat d'Edimburg, School of Informatics: Jonathan Kilgour [consulta: maig 2009]. Disponible a <<http://groups.inf.ed.ac.uk/nxt>>.
- UNIVERSITAT D'HAMBURG: *EXMARaLDA*, Hamburg, Universitat d'Hamburg, Collaborative Research Center Multilingualism: Thomas Schmidt, Kai Wörner [consulta: maig 2009]. Disponible a <<http://www.exmaralda.org/>>.
- UNIVERSITAT DE CALIFÒRNIA SANTA BARBARA: *Transcription in Action. Resources for the representation of linguistic interaction*, Santa Barbara, Universitat de Califòrnia Santa Barbara, Department of Linguistics: Mary Bucholtz, John W. Du Bois [consulta: maig 2009]. Disponible a <<http://www.linguistics.ucsb.edu/projects/transcription/index.html>>.
- UNIVERSITAT DE FLORÈNCIA: *C-Oral-Rom Project. Integrated reference corpora for spoken Romance languages*, Florència, Universitat de Florència, Lablita: Emanuela Cresti [consulta: maig 2009]. Disponible a <<http://lablita.dit.unifi.it/coralrom/index.html>>.
- UNIVERSITAT DE LLEIDA: *COR. Corpus Oral de Registres*, Lleida, Universitat de Lleida, departament de Filologia Catalana: Antoni Bernadó, Amadeu Viana [consulta: maig 2009]. Disponible a <[http://www.filcat.udl.es/3\\_recerca/unic/cor/index.htm](http://www.filcat.udl.es/3_recerca/unic/cor/index.htm)>.
- UNIVERSITAT DE WISCONSIN-MADISON: *Transana. Qualitative analysis software for video and audio data*, Madison, Universitat de Wisconsin-Madison, WCER, School of Education [consulta: maig 2009]. Disponible a <<http://www.transana.org/>>.
- XCES: *XCES Corpus Encoding Standard for XML*, Poughkeepsie (Nova York), Vassar College, Department of Computer Science/Vandoeuvre-lès-Nancy, LORIA/CNRS, Équipe Langue et Dialogue [consulta: maig 2009]. Disponible a <<http://www.xces.org/>>.

#### BIBLIOGRAFIA COMPLEMENTÀRIA

- BALDRY, ANTHONY i PAUL J. THIBAUT (2006): *Multimodal Transcription and Text Analysis*, Equinox, Londres/Oakville.
- BLANCHE-BENVENISTE, CLAIRE (2002): "Réflexions sur les transcriptions du corpus de français parlé", *Revue Parole*, 22-23-24, pàg. 91-118.
- BLANCHE-BENVENISTE, CLAIRE; BILGER, MIREILLE; ROUGET, CHRISTINE i KAREL VAN DEN EYNDE (1991): *Le français parlé. Études grammaticales*, CNRS Editions, París.
- BOSCO, CRISTINA i BAZZANELLA, CARLA (2002): "Contextualization in spoken language corpora", dins Pusch, Claus D. i Wolfgang Raible (ed.): *Romanistische Korpuslinguistik – Korpora und gesprochene Sprache. Romance Corpus Linguistics – Corpora and Spoken Language*. Script Oralia, 126, Narr, Tübingen, pàg. 19-30.
- BRIZ, ANTONIO (coord.) (1995): *La conversación coloquial (Materiales para su estudio)*, Universitat de València, València.
- BUCHOLTZ, MARY (2000): "The politics of transcription", *Journal of Pragmatics*, 32 (10), pàg. 1439-1465.
- \_\_\_\_\_ (2007): "Variation in transcription", *Discourse Studies*, 9 (6), pàg. 784-808.
- CAPPEAU, PAUL (1997): "Données erronées: quelles erreurs commettent les transcrip-teurs?", *Recherches sur le français parlé*, 14, pàg. 117-126.
- CORBERA, JAUME (dir.) (2006): *Arxiu audiovisual dels dialectes catalans de les Illes Balears. Antologia*, Càtedra Alcover-Moll-Villangóme-z, Palma.
- CROWDY, STEVE (1994): "Spoken corpus transcription", *Literary & Linguistic Computing*, 9 (1), pàg. 25-28.
- DU BOIS, JOHN W. (2003): "Discourse and grammar", dins Tomasello, Michael (ed.): *The new psychology of language: Cognitive and functional approaches to language structure*, Lawrence Erlbaum, Mahwah, Nova Jersey, vol. 2, pàg. 47-87.
- EDWARDS, JANE A. (1995): "Principles and alternative systems in the transcription, coding and mark-up of spoken discourse", dins Leech, Geoffrey; Myers, Greg i Jenny Thomas (ed.): *Spoken English on Computer: Transcript, Mark-up and Application*, Longman, Nova York, pàg. 19-34.
- HAVE, PAUL TEN (1999): *Doing Conversational Analysis. A practical guide*, Sage, Londres/Thousand Oaks/Nova Delhi.
- \_\_\_\_\_ (2002): "Reflections on transcription", *Cahiers de Praxématique*, 39, pàg. 21-43.
- HEPBURN, ALEXA (2004): "Crying: Notes on Description, Transcription, and Interaction", *Research on Language & Social Interaction*, 37, pàg. 251-291.
- JAFFE, ALEXANDRA (2000): "Introduction: Non-standard orthography and non-standard speech", *Journal of Sociolinguistics*, 4 (4), pàg. 497-513.

- \_\_\_\_\_ (2000): "The voices people read: Orthography and the representation of non-standard speech", *Journal of Sociolinguistics*, 4 (4), pàg. 561-587.
- JEFFERSON, GAIL (2004): "Glossary of transcript symbols with an introduction", dins Lerner, Gene H. (ed.): *Conversation Analysis: Studies from the first generation*, John Benjamins, Amsterdam, pàg. 13-31.
- KENNEDY, GRAEME (1998): *An introduction to Corpus Linguistics*, Longman, Nova York/Londres.
- LEE, DAVID: *Bookmarks for Corpus-based Linguistics* [en línia], Hong Kong, Universitat de Hong Kong: David Lee [consulta: maig 2009]. Disponible a <<http://devoted.to/corpora>>.
- LINDSAY, JEAN i DANIEL O'CONNELL (1995): "How do transcribers deal with audio recordings of spoken discourse?", *Journal of Psycholinguistic Research*, 24 (2), pàg. 101-116.
- LLISTERRI, JOAQUIM (1996): "Els corpus lingüístics orals", dins Payrató, Lluís; Boix, Emili; Lloret, Maria Rosa i Mercè Lorente (ed.), *Corpus, corpora*, PPU, Barcelona, pàg. 27-70.
- LÜDELING, ANKE i MERJA KYTÖ (ed.) (2008): *Corpus linguistics: an international handbook*, Walter de Gruyter, Berlín/Nova York.
- MACAULAY, RONALD K. S. (1991): "Coz it izny spelt when they say it. Displaying dialect in writing", *American Speech*, 66 (3), pàg. 280-291.
- MACLEAN, LYNNE M.; MEYER, MECHTHILD i ALMA ESTABLE (2004): "Improving Accuracy of Transcripts in Qualitative Research", *Qualitative Health Research*, 14, pàg. 113-123.
- MCCARTHY, MICHAEL i ANNE O'KEEFFE (2008): "Corpora and spoken language", dins Lüdeling, Anke i Merja Kytö (ed.): *Corpus Linguistics. An International Handbook*, Walter de Gruyter, Berlín/Nova York, vol. 2, pàg. 1008-1023.
- MIETHANER, ULRICH (2000): "Orthographic transcriptions of non-standard varieties: The case of Earlier African-American English", *Journal of Sociolinguistics*, 4 (4), pàg. 534-560.
- MONDADA, LORENZA (2002): "Pratiques de transcriptions et effets de catégorisation", *Cahiers de Praxématique*, 39, pàg. 45-75.
- \_\_\_\_\_ (2007): "Commentary: transcript variations and the indexicality of transcribing practices", *Discourse Studies*, 9 (6), pàg. 809-821.
- NORRIS, SIGRID (2002): "The implication of visual research for discourse analysis: Transcription beyond language", *Visual Communication*, 1 (1), pàg. 97-121.
- O'CONNELL, DANIEL C. i SABINE KOWAL (1999): "Transcription and the issue of standardization", *Journal of Psycholinguistic Research*, 28, pàg. 103-120.
- \_\_\_\_\_ (2000): "Are transcripts reproducible?", *Pragmatics*, 10 (2), pàg. 247-269.
- OOSTDIJK, NELLEKE i LOU BOVES (2008): "Preprocessing speech corpora: Transcription and phonological annotation", dins Lüdeling, Anke i Merja

- Kytö (ed.): *Corpus Linguistics. An International Handbook*, Walter de Gruyter, Berlín/Nova York, vol. 1, pàg. 642-663.
- OSTLER, NICHOLAS (2008): "Corpora of less studied languages", dins Lüdeling, Anke i Merja Kytö (ed.): *Corpus Linguistics. An International Handbook*, Walter de Gruyter, Berlín/Nova York, vol. 1, pàg. 457-483.
- PALLAUD, BERTHILLE (2002): "Erreurs d'écoute dans la transcription de données orales", *Revue Parole*, 22-23-24, pàg. 267-294.
- PSATHAS, GEORGE i TIM ANDERSON (1990): "The practices of transcription in conversation analysis", *Semiotica*, 78, pàg. 75-99.
- SPINOS, ANNA-MARIE R.; O'CONNELL, DANIEL C. i SABINE KOWAL (2002): "An empirical investigation of pauses notation", *Pragmatics*, 12 (1), pàg. 1-9.
- TRAVERSO, VÉRONIQUE (2002): "Transcription et traduction des interactions en langue étrangère", *Cahiers de Praxématique*, 39, pàg. 77-99.
- WITTENBURG, PETER (2008): "Preprocessing multimodal corpora", dins Lüdeling, Anke i Merja Kytö (ed.): *Corpus Linguistics. An International Handbook*, Walter de Gruyter, Berlín/Nova York, vol. 1, pàg. 664-684.
- XIAO, RICHARD (2008): "Well-known and influential corpora", dins Lüdeling, Anke i Merja Kytö (ed.): *Corpus Linguistics. An International Handbook*, Walter de Gruyter, Berlín/Nova York, vol. 1, pàg. 383-456.

## TAULA DE CONVENCIONS

A continuació recollim les convencions del sistema de transcripció discursiva del CCCUB de nivell intermedi-estret, que són les que s'utilitzen per defecte. Marquem amb un asterisc les convencions més pròpies d'un nivell estret de transcripció.

### 1. Aspectes prosòdics

Seqüència tonal terminal ascendent	/
Seqüència tonal terminal de manteniment	—
Seqüència tonal terminal descendent	\
*Seqüència tonal terminal ascendent	↑
*Seqüència tonal terminal lleugerament ascendent	//
*Seqüència tonal terminal de manteniment-ascens	/
*Seqüència tonal terminal de manteniment	→
*Seqüència tonal terminal de manteniment-descens	\
*Seqüència tonal terminal lleugerament descendent	\ \
*Seqüència tonal terminal descendent	↓
Truncament de grup tonal	--
Truncament de mot	-
Truncament de mot i grup tonal	--
*Límit de grup tonal menor	
*Accent principal de grup tonal	*
*Accent secundari de grup tonal	•
Allargament breu	:
Allargament mitjà	::
Allargament llarg	::: (o més)
Intensitat forta	{{(F) text afectat}}
Intensitat molt forta	{{(FF) text afectat}}
Intensitat fluixa	{{(P) text afectat}}
Intensitat molt fluixa	{{(PP) text afectat}}
Ritme d'elocució accelerat	{{(AC) text afectat}}

Ritme d'elocució molt accelerat	<i>{{(ACC) text afectat}}</i>
Ritme d'elocució desaccelerat	<i>{{(DC) text afectat}}</i>
Ritme d'elocució molt desaccelerat	<i>{{(DCC) text afectat}}</i>
Manteniment de l'entonació ascendent	<i>{{(A) text afectat}}</i>
Manteniment de l'entonació descendent	<i>{{(B) text afectat}}</i>
Estrafer la veu	<i>{{(EV) text afectat}}</i>

## 2. Aspectes vocals

Riure no simultani amb el text (un símbol per síl·laba)	@
Riure simultani amb el text	<i>{{(@) text afectat}}</i>
Riure llarg	@RR@
Riure llarg amb durada	@R(durada)R@
Altres fenòmens (somrient, tossint, cantant...)	<i>{{(fenomen) text afectat}}</i>
Inhalació	(INH)
Exhalació	(EXH)
Fragment amb inhalació	<i>{{(INH) text afectat}}</i>
Fragment amb exhalació	<i>{{(EXH) text afectat}}</i>

## 3. Pauses i encavalcaments

Pausa molt breu	( $0.1 \leq p < 0.3$ segons)	(. durada)
Pausa breu	( $0.3 \leq p < 1$ segons)	(.. durada)
Pausa mitjana	( $1 \leq p < 3$ segons)	(... durada)
Pausa llarga	( $p \geq 3$ segons)	(.... durada)
*Grup tonal sense pausa inicial		(0)
Encavalcament		<i>[text afectat]</i>
Encavalcament múltiple		<i>[1 text afectat]</i>
		<i>[1 text afectat]</i>
		<i>[2 text afectat]</i>
		<i>[2 text afectat]</i>



\*Represa de grup tonal &

#### 4. Regularitzacions i comentaris

Elisió	<i>(grafia del so elidit)</i>
Comentari explicatiu del transcriptor	<i>((comentari))</i>
Comentari descriptiu del transcriptor	<i>(comentari)</i>
Fragment dirigit a un parlant determinat	<i>((a INICIALS))</i>

#### 5. Fragments conflictius

Dubtes sobre el text transcrit	<i>{{(??) text afectat}}</i>
Parlant no identificat	<i>???:</i>
Parlants no identificats que intervenen alhora	<i>T??:</i>
Síl·laba intel·ligible	<i>x</i>
Fragment intel·ligible	<i>xXXx</i>
Fragment intel·ligible amb durada	<i>xX(durada)Xx</i>

#### 6. Altres aspectes

Segona llengua	<i>{{(L2) text afectat}}</i>
*Discurs directe	<i>{{(DD) text afectat}}</i>
*Fragment parentètic	<i>{{(PAR) text afectat}}</i>
Fenòmens gestuals	<i>{{(descripció del fenomen) text afectat}}</i>
*Fenòmens gestuals esporàdics	<i>{{(X= fenomen) text X text X}}</i>
*Transcripció fonètica	<i>(/transcripció fonètica/)</i>



## ÍNDEX TEMÀTIC

- Accent tonal, § 4.5.2
- Alfabet Fonètic Internacional, § 3.1
- Allargament, § 4.4.1, 4.4.10, 4.4.17
- Anacrusi, § 4.4.1
- Anàlisi quantitativa/qualitativa, § 3.3
- Anonimat dels informants, § 2.2,  
4.4.13
- Cant, § 4.4.11
- Claudàtors dobles, § 4.5.3
- Codificació de la transcripció, § 4.3
- Codis, § 4.4.17
- combinació de codis, § 4.4.9.1,  
4.4.17
- ordenació de codis, § 4.4.17
- Columnes de text, § 4.2
- Comentari del transcriptor, § 4.4.5,  
4.4.12, 4.4.13
- entre parèntesis dobles, § 4.4.13
- entre parèntesis simples, § 4.4.13
- Connector, § 4.4.3.2, 4.4.3.3
- Context comunicatiu: descripció,  
§ 4.3
- bases de dades, § 4.3
- descripció al CCCUB, § 4.3
- model Speaking, § 4.3
- Convencions tipogràfiques, § 4.2,  
4.4.16
- Corba d'entonació, § 4.4.1
- Corpus Audiovisual Plurilingüe, § 1.3,  
4.3
- Corpus de Català Contemporani de  
la Universitat de Barcelona  
(CCCUB), § 1.3
- Corpus de Varietats Socials, § 1.3,  
4.3
- Corpus Oral de Conversa Col·loquial,  
§ 1.3, 4.3
- Corpus Oral de Registres, § 1.3, 4.3
- Corpus Oral de Registres de la  
Universitat de Lleida, § 4.3,  
4.4.14, 4.4.5.13
- Fenòmens contextuais, § 4.4.13,  
4.5.2
- Discurs oral, § 3.1
- Disposició del text, § 3.2, 4.2, 4.5.1,  
4.5.2, 4.5.3
- Eines de transcripció, § 2.3
- Encavalcament, § 4.4.7
- alineament d'encavalcaments,  
§ 3.2, 4.5.2
- fragment encavalcat intel·ligible,  
§ 4.4.7
- múltiple, § 4.4.7
- numeració, § 4.4.7
- posició, § 4.4.7, 4.4.17
- simple, § 4.4.7
- Enregistrament, § 2
- aspectes ètics, § 2.2
- metodologia, § 2.1
- Errades de transcripció, § 4.4.17

- Estàndards internacionals de transcripció i marcatge, § 3.5
- Exhalació, § 4.4.9.3, 4.5.3
- Fenòmens vocals no verbals, § 4.4.11
- aclariment de gola, § 4.4.11
  - esbufec, § 4.4.11
  - espetec de llengua, § 4.4.9.4
  - esternut, § 4.4.11
  - simultanis amb la parla, § 4.4.11, 4.5.2
  - tos, § 4.4.11
  - xiulet, § 4.4.11
- Fragment dubtós, § 4.4.7, 4.4.8, 4.5.1, 4.5.3
- Fragment en una altra llengua, § 4.4.15.5, 4.5.3
- Fragment intel·ligible, § 4.4.7, 4.4.8
- Fragment parentètic, § 4.5.2
- Gestualitat, § 4.4.12, 4.5.1, 4.5.2
- Gramaticalització, § 4.4.15.4
- Grup tonal
- definició i delimitació, § 4.4.1, 4.4.17
  - grups tonals majors i menors, § 4.5.2
  - llargada, § 4.4.3.1
  - represa, § 4.4.6, 4.4.7, 4.5.2
- Inicials dels informants, § 4.2, 4.4.5, 4.4.13, 4.4.17
- informant no identificat, § 4.4.5
  - intervenció conjunta, § 4.4.5
- Incís, § 4.4.3.2, 4.4.3.3
- Inhalació, § 4.4.9.3, 4.5.3
- Intensitat
- fluixa, § 4.4.9.1, 4.5.2
  - forta, § 4.4.9.1, 4.5.2
  - *in crescendo*, § 4.5.2
  - *in decrescendo*, § 4.5.2
- Interjecció
- transcripció ortogràfica, § 4.4.10, 4.4.15.4
  - transcripció prosòdica, § 4.4.3.2, 4.4.3.3
- Interpretació del text, § 3.3, 4.4.17
- Majúscules, § 4.4.16, 4.5.3
- Marcador discursiu, § 4.4.3.2
- Models de transcripció discursiva, § 3.5
- Mot cua, § 4.4.3.2, 4.4.3.3
- Nivells de transcripció, § 3.1, 4.1
- exemple de transcripció ampla, § 4.5.1
  - exemple de transcripció estreta, § 4.5.2
  - exemple de transcripció intermèdia-estreta, § 4.1, 4.5.1
  - transcripció ampla, § 4.1, 4.5.1
  - transcripció estreta, § 4.1, 4.5.2
  - transcripció intermèdia, § 4.5.1
  - transcripció intermèdia-estreta, § 4.1-4.4
- Numeració de grups tonals, § 4.2, 4.4.17
- Omplidor de parla, § 4.4.3.3, 4.4.9.4
- Onomatopeia, § 4.4.10
- Ordre cronològic, § 4.4.7, 4.4.17
- Parèntesis triangulars, § 4.5.3

- Pausa
- delimitació de grup tonal, § 4.4.1
  - durada i posició, § 3.2, 4.4.6
  - enllaç immediat, § 4.4.6
  - transcripció estreta, § 4.5.2
- Polimorfisme, § 4.4.15.4
- Regularització de grafies no estàndards: criteris generals, § 3.2, 4.4.14
- Regularització del CCCUB, § 3.2, 4.4.15, 4.5.1
- accentuació, § 4.4.15.2
  - articles, § 4.4.15.3
  - assimilació i dissimilació, § 4.4.15.2
  - demostratius, possessius i indefinits, § 4.4.15.3
  - díctics, § 4.4.15.4
  - diftong i hiat, § 4.4.15.2
  - elisió, § 4.4.15.1
  - epèntesi, § 4.4.15.2
  - lèxic, § 4.4.15.4
  - manlleus, § 4.4.15.5
  - morfologia verbal, § 4.4.15.3
  - pronoms forts, § 4.4.15.3
  - pronoms febles, § 4.4.15.3
  - reducció de grup consonàntic, § 4.4.15.1
  - substantius i adjectius, § 4.4.15.3
  - ús de parèntesis, § 4.4.15.1
- Rematització, § 4.4.3.2
- Ressituació del to, § 4.4.1
- Retroalimentador, § 4.4.3.3, 4.4.9.4, 4.4.17
- Revisió i edició, § 5
- Riure, § 4.4.9.2
- simultani amb la parla, § 4.4.9.2
  - múltiple, § 4.4.9.2
  - somriure, § 4.4.9.2
- Seqüència tonal terminal: definició, § 4.4.2, 4.4.3.3, 4.4.17, 4.5.1
- ascens, § 4.4.2, 4.4.3.3, 4.4.17
  - descens, § 4.4.2, 4.4.3.3, 4.4.17
  - manteniment, § 4.4.2, 4.4.3.3, 4.4.17
  - transcripció estreta, § 4.4.2, 4.5.2
  - transcripció funcional, § 4.5.2
- Signes de puntuació, § 4.4.16
- Símbol del dòlar, § 4.5.3
- Símbol d'igualtat, § 4.5.3
- Tabuladors i espais en blanc, § 4.2
- Tematització, § 4.4.3.2, 4.4.3.3
- Tempo* accelerat, § 4.4.9.1, 4.5.2
- Tempo* desaccelerat, § 4.4.9.1, 4.5.2
- To de frontera, § 4.5.2
- Traducció del text, § 4.5.3
- Transcripció: definició, § 3.1
- discursiva, § 3.4, 3.6
  - discursiva del CCCUB, § 4.1
  - objectius, § 3.3
  - principis metodològics, § 3.2, 3.5
- Transcripció fonètica, § 4.5.2
- Transliteració, § 3.1
- Truncament
- de grup tonal, § 4.4.2
  - de mot, § 4.4.4
  - de mot i grup tonal, § 4.4.2
- Unitat de transcripció, § 3.2

— grup tonal, § 4.4.1  
— torn de parla, § 4.4.5  
— paraula, § 4.4.4  
Variació tonal  
— to alt, § 4.4.9.1

— to baix, § 4.4.9.1  
— veu estafeta, § 4.4.9.1  
Vocatiu, § 4.4.3.2, 4.4.3.3  
Xifres, § 4.4.16  
Xiuxiueig, § 4.4