

**LA COMPLEXITAT LINGÜÍSTICA  
EN EL DISCURS ORAL I ESCRIT:  
densitat lèxica, composició oracional  
i connexió textual**

**per Josep M. Castellà Lidon**

**Direcció: Lluís Payrató**

**Setembre del 2001**

**Dipòsit legal: B.14058-2002  
ISBN: 84-699-8511-6**

**Tesi doctoral  
Programa de Doctorat *Lingüística Aplicada (lèxic i discurs)*  
Institut Universitari de Lingüística Aplicada  
Universitat Pompeu Fabra**



**Per a la Rosa Casado, la meva dona.**



**“Ciència i poesia són, tanmateix, amor, car és l’amor qui fa entenedores les coses.” Guerau de Liost (1922)**



# ÍNDEX

<b>AGRAÏMENTS</b>	.....	<b>pàg. 19</b>
<b>RESUM DE LA RECERCA</b>	.....	<b>pàg. 23</b>
<b>1- PRESENTACIÓ</b>	.....	<b>pàg. 25</b>
<b>2- MARC TEÒRIC</b>	.....	<b>pàg. 33</b>
<b>2.1- EL CONTRAST LLENGUA ORAL / LLENGUA ESCRITA:</b>		
<b>PERSPECTIVA HISTÒRICA</b>	.....	<b>pàg. 33</b>
2.1.1- Oralitat i escriptura	.....	pàg. 33
2.1.2- La tradició en els estudis lingüístics	.....	pàg. 36
2.1.3- La necessitat de l'estudi de la llengua oral i de la llengua escrita	.....	pàg. 40

**2.2- EL MODE DEL REGISTRE ..... pàg. 47**

2.2.1- Registre, canal i mode ..... pàg. 47

2.2.2- La gradació llengua oral / llengua escrita ..... pàg. 52

2.2.3- Altres paràmetres que incideixen en el mode del registre .. pàg. 62

**2.3- LA COMPLEXITAT LINGÜÍSTICA EN LA LLENGUA  
ORAL I EN LA LLENGUA ESCRITA ..... pàg. 73**

2.3.1- La complexitat estructural de la llengua escrita (I):

W. Chafe ..... pàg. 73

2.3.2- La complexitat estructural de la llengua escrita (II):

altres autors ..... pàg. 76

2.3.3- La complexitat estructural de la llengua oral (I):

M.A.K. Halliday ..... pàg. 87

2.3.4- La complexitat estructural de la llengua oral (II):

altres autors ..... pàg. 93

2.3.5- Sobre la noció mateixa de complexitat lingüística ..... pàg. 103

**3- METODOLOGIA ..... pàg. 109**

**3.1- OBJECTIUS I JUSTIFICACIÓ DE LA RECERCA ..... pàg. 109**

3.1.1- Àmbit de la recerca ..... pàg. 109

3.1.2- Necessitat i originalitat de la recerca ..... pàg. 111

3.1.3- Objectius i pregunta de recerca ..... pàg. 113

3.1.4- Hipòtesis de recerca ..... pàg. 115

**3.2- LA VARIABLE DEPENDENT ..... pàg. 119**

3.2.1- Les variables de la recerca ..... pàg. 119



3.2.2- El gènere discursiu .....	pàg. 121
3.2.3- La conversa informal .....	pàg. 125
3.2.4- La classe magistral .....	pàg. 127
3.2.5- La prosa acadèmica .....	pàg. 130
<b>3.3 - LES VARIABLES INDEPENDENTS .....</b>	<b>pàg. 134</b>
3.3.1- Variables independents de l'estudi .....	pàg. 134
3.3.2- Trets lingüístics no inclosos en l'estudi .....	pàg. 137
<b>3.4- LES VARIABLES DE CONTROL .....</b>	<b>pàg. 142</b>
3.4.1- Validesa interna i externa: representativitat de la mostra ..	pàg. 142
3.4.2- Registres i seqüències textuais dominants .....	pàg. 151
3.4.3- Textos, individus i situacions de producció .....	pàg. 155
<b>3.5- LES DIMENSIONS DE LA MOSTRA .....</b>	<b>pàg. 160</b>
3.5.1- Extensió idònia dels textos .....	pàg. 160
3.5.2- Extensió de la mostra .....	pàg. 163
<b>3.6- ALTRES ASPECTES METODOLÒGICS .....</b>	<b>pàg. 165</b>
3.6.1- Sobre la transcripció i sobre la preparació dels textos ...	pàg. 165
3.6.2- Sobre la fiabilitat en la codificació .....	pàg. 166
3.6.3- Sobre la significació de les dades: el test $\chi^2$ (chi quadrat) ..	pàg. 171
3.6.4- Sobre els denominadors i sobre les repeticions contigües ..	pàg. 176
3.6.5- Sobre la recerca bibliogràfica .....	pàg. 178

<b>4- RESULTATS DE LA RECERCA (I): DENSITAT LÈXICA, ESTIL NOMINAL I ESTIL VERBAL</b>	<b>..... pàg. 181</b>
<b>4.1- LA DENSITAT LÈXICA, L'ESTIL NOMINAL I L'ESTIL VERBAL</b>	<b>..... pàg. 181</b>
<b>4.2- ELS NOMS</b>	<b>..... pàg. 185</b>
4.2.1- Fonamentació conceptual i hipòtesis	..... pàg. 185
4.2.2- Unitats, procediments i criteris de codificació	..... pàg. 187
4.2.3- Resultats i discussió	..... pàg. 193
4.2.4- Conclusions de l'apartat	..... pàg. 206
<b>4.3- ELS VERBS</b>	<b>..... pàg. 207</b>
4.3.1- Fonamentació conceptual i hipòtesis	..... pàg. 207
4.3.2- Unitats, procediments i criteris de codificació	..... pàg. 209
4.3.3- Resultats i discussió	..... pàg. 213
4.3.4- Conclusions de l'apartat	..... pàg. 227
<b>4.4- ELS ADJECTIUS</b>	<b>..... pàg. 228</b>
4.4.1- Fonamentació conceptual i hipòtesis	..... pàg. 228
4.4.2- Unitats, procediments i criteris de codificació	..... pàg. 229
4.4.3- Resultats i discussió	..... pàg. 232
4.4.4- Conclusions de l'apartat	..... pàg. 242
<b>4.5- ELS ADVERBIS</b>	<b>..... pàg. 243</b>
4.5.1- Fonamentació conceptual i hipòtesis	..... pàg. 243
4.5.2- Unitats, procediments i criteris de codificació	..... pàg. 248

4.5.3- Resultats i discussió .....	pàg. 249
4.5.4- Conclusions de l'apartat .....	pàg. 254
<b>4.6- ELS MOTS LEXICALS I ELS MOTS GRAMATICALS .....</b>	<b>pàg. 255</b>
4.6.1- Fonamentació conceptual, unitats, procediments i criteris de codificació .....	pàg. 255
4.6.2- Resultats i discussió .....	pàg. 256
4.6.3- Conclusions de l'apartat .....	pàg. 262
<b>4.7- ELS SINTAGMES PREPOSICIONALS .....</b>	<b>pàg. 263</b>
4.7.1- Fonamentació conceptual i hipòtesis .....	pàg. 263
4.7.2- Unitats, procediments i criteris de codificació .....	pàg. 264
4.7.3- Resultats i discussió .....	pàg. 268
4.7.4- Conclusions de l'apartat .....	pàg. 281
<b>4.8- LES NOMINALITZACIONS .....</b>	<b>pàg. 283</b>
4.8.1- Fonamentació conceptual i hipòtesis .....	pàg. 283
4.8.2- Unitats, procediments i criteris de codificació .....	pàg. 287
4.8.3- Resultats i discussió .....	pàg. 290
4.8.4- Conclusions de l'apartat .....	pàg. 297
<b>4.9- LA VARIACIÓ LÈXICA, LA REPETICIÓ LÈXICA I LA LONGITUD DE MOT .....</b>	<b>pàg. 298</b>
4.9.1- Fonamentació conceptual i hipòtesis .....	pàg. 298
4.9.2- Unitats, procediments i criteris de codificació .....	pàg. 302
4.9.3- Resultats i discussió .....	pàg. 305
4.9.4- Conclusions de l'apartat .....	pàg. 311

<b>4.10- LA DENSITAT LÈXICA, L'ESTIL NOMINAL I L'ESTIL VERBAL: VISIÓ DE CONJUNT</b>	<b>.....</b>	<b>pàg. 312</b>
4.10.1- Densitat lèxica: dades i interpretació globals	.....	pàg. 313
4.10.2- Estil nominal i estil verbal: dades i interpretació globals	.....	pàg. 315
4.10.3- Comparació català-anglès	.....	pàg. 317
4.10.4- Resum dels trets característics de cada gènere	.....	pàg. 325
4.10.5- Posició relativa de la classe magistral	.....	pàg. 328
4.10.6- Distribució dels resultats: bandes i correspondències	.....	pàg. 331
4.10.7- Estil nominal i estil verbal: exemples de textos	.....	pàg. 334
4.10.8- Conclusions de l'apartat	.....	pàg. 338

## **5- RESULTATS DE LA RECERCA (II): COMPOSICIÓ ORACIONAL I CONNEXIÓ TEXTUAL** ..... **pàg. 341**

### **5.1- LA COMPOSICIÓ ORACIONAL I LA CONNEXIÓ TEXTUAL** ..... **pàg. 341**

5.1.1- Sintaxi i cohesió: l'articulació de la linealitat lingüística	..	pàg. 341
5.1.2- La clàusula en la llengua oral i en la llengua escrita	.....	pàg. 353

### **5.2- LA COMPOSICIÓ ORACIONAL (I): LES CONSTRUCCIONS DE SUBORDINACIÓ** ..... **pàg. 365**

5.2.1- Fonamentació conceptual i hipòtesis	.....	pàg. 365
5.2.2- Unitats, procediments i criteris de codificació	.....	pàg. 376
5.2.3- Resultats i discussió	.....	pàg. 387
5.2.4- Conclusions de l'apartat	.....	pàg. 413

<b>5.3- LA COMPOSICIÓ ORACIONAL (II): LES CONSTRUCCIONS DE COORDINACIÓ</b>	<b>pàg. 416</b>
5.3.1- Fonamentació conceptual i hipòtesis	pàg. 416
5.3.2- Unitats, procediments i criteris de codificació	pàg. 422
5.3.3- Resultats i discussió	pàg. 430
5.3.4- Conclusions de l'apartat	pàg. 444
<b>5.4- LA COMPOSICIÓ ORACIONAL (III): LES CONSTRUCCIONS D'INTERORDINACIÓ</b>	<b>pàg. 446</b>
5.4.1- Fonamentació conceptual i hipòtesis	pàg. 446
5.4.2- Unitats, procediments i criteris de codificació	pàg. 451
5.4.3- Resultats i discussió	pàg. 473
5.4.4- Conclusions de l'apartat	pàg. 490
<b>5.5- LA CONNEXIÓ TEXTUAL (I): LES CONJUNCIONS I ELS MARCADORS O CONNECTORS PARENTÈTICS EN USOS EXTRAORACIONALS</b>	<b>pàg. 491</b>
5.5.1- Fonamentació conceptual i hipòtesis	pàg. 491
5.5.2- Unitats, procediments i criteris de codificació	pàg. 501
5.5.3- Resultats i discussió	pàg. 524
5.5.4- Conclusions de l'apartat	pàg. 544
<b>5.6- LA CONNEXIÓ TEXTUAL (II): LES INTERJECCIONS CONNECTIVES I ELS OMLIDORS</b>	<b>pàg. 546</b>
5.6.1- Fonamentació conceptual i hipòtesis	pàg. 546
5.6.2- Unitats, procediments i criteris de codificació	pàg. 553
5.6.3- Resultats i discussió	pàg. 563
5.6.4- Conclusions de l'apartat	pàg. 578

<b>5.7- LA COMPOSICIÓ ORACIONAL I LA CONNEXIÓ</b>	
<b>TEXTUAL: VISIÓ DE CONJUNT</b> .....	<b>pàg. 580</b>
5.7.1- Estil nominal i estil verbal: dades i interpretació globals ..	pàg. 580
5.7.2- Composició oracional i connexió textual: dades i interpretació globals .....	pàg. 583
5.7.3- Comparació català-anglès .....	pàg. 590
5.7.4- Resum dels trets característics de cada gènere .....	pàg. 597
5.7.5- Posició relativa de la classe magistral .....	pàg. 604
5.7.6- Distribució dels resultats: bandes i correspondències ...	pàg. 608
5.7.7- Composició oracional i connexió textual: exemples de textos .....	pàg. 609
5.7.8- Conclusions de l'apartat .....	pàg. 617
<b>6- CONCLUSIONS</b> .....	<b>pàg. 619</b>
<b>7- BIBLIOGRAFIA</b> .....	<b>pàg. 627</b>

## AGRAÏMENTS

Una obra que s'elabora durant un període llarg de temps ha de tenir, gairebé per força, multitud de contribuïdors. La present tesi, naturalment, no és pas una excepció a aquest principi. En aquests anys han estat moltíssims els companys i amics a qui he preguntat dubtes i he demanat informació, que m'han fet conèixer referències concretes i m'han facilitat l'accés o la realització de determinades tasques, o, simplement, que m'han escoltat en alguna època en què els problemes plantejats demanaven de ser verbalitzats, per intentar de donar-los llum nova en el diàleg creatiu entre col·legues.

En el començament, els projectes solen tenir els perfils difusos. Recordo amb nitidesa, però, una classe de doctorat, l'any 1990 a la Universitat de Barcelona, en què M. Paz Battaner va presentar l'obra d'un autor per a mi aleshores desconegut, D. Biber, que va resultar ser la guspira inicial de tot el meu treball. Recordo també el dia que Montserrat Vilà, en una taula rodona al Centre Borja de Sant Cugat del Vallès, va afirmar amb convicció que calia encetar la descripció de la llengua oral formal. Després, hi va haver les converses amb Miquel Llobera, Anna Camps —a Elx, el 1993—, Margarida Cambra, Javier Lahuerta, Mercè Pujol, Sebastián Bonilla, M. Josep Cuenca —en successives visites a Barcelona i a València—, Amparo Tuson, Joan Solà, Vicent Salvador —aquests dos darrers, a Frankfurt, el 1994, arran del congrés de l'AILLC— i, especialment, Anna Cros, amb la qual vàrem parlar llargament de la seva tesi i de la meva, amb un destacat objectiu comú: el discurs del professor a l'aula. Més endavant, ja dins del Programa de Doctorat *Lingüística Aplicada (lèxic i discurs)*, a l'IULA de la Universitat Pompeu Fabra, em varen orientar Toni Badia —el meu tutor—, Mercè Tricàs i M. Teresa Turell.

El 1995, un any després d'haver tingut l'encert de demanar a Lluís Payrató que en fos el director, la tesi va prendre un camí més definit i va començar a adoptar la forma que ha tingut finalment. De llavors ençà, també m'han ajudat en un moment o un altre moltes

persones —Marta Anton, Remei Arnaus, Carme Bach, Gabrielle Berkenbusch, Cristine Bierwach, Glòria Bordons, Teresa Cabré, Esteve Clua, Pepe Contreras, Elisabet Costa, Neus Galí, Anna Gili, Montse González, M. Josep Marín, Dolors Ollé, Manel Ollé, Anna Oller, Francesc Piferrer, Montserrat Ribas, Jaume Subirana, Mireia Trenchs, Frederic Udina i Lluís de Yzaguirre—, alhora que han seguit fent-ho moltes de les que he esmentat anteriorment. Helena Calsamiglia, Josep M. Fontana i Joan Julià varen tenir la gentilesa de convidar-me a explicar el meu treball durant el procés d'elaboració, al grup de recerca, al curs de doctorat i al curs de postgrau de què, respectivament, eren responsables. Aquestes presentacions i les que es van organitzar per dues vegades a l'IULA, em van ser útils per posar a prova la feina feta fins al moment i per rebre valuosos suggeriments dels assistents.

He d'agrair en gran manera la col·laboració dels professors universitaris que es varen prestar a ser conillets d'Índies —els autèntics protagonistes anònims d'aquest treball—, em varen deixar entrar a les seves classes per enregistrar-los i em varen facilitar textos acadèmics escrits per ells mateixos. He d'agrair, també, als responsables del *Corpus de Català Contemporani de la Universitat de Barcelona (CUB)* —en particular del *Corpus Oral de Conversa Col·loquial (COC)*— que m'hagin permès d'adoptar-ne una part per a la meua mostra, i a l'esforçat equip que va fer-ne els enregistraments i les transcripcions, que hi dedicés el temps pensant que arribaria un dia que un investigador trauria profit de la seva feina.

He contret un deute especial amb les persones que, ja en la fase final, han fet la lectura d'algun o de tots els capítols de l'obra: Rosa Casado, Anna Cros, Miquel M. Gibert, M. Josep Cuenca i M. Teresa Turell. A les dues darreres, en concret, perquè les indicacions que m'han fet, cadascuna sobre la seva especialitat, m'han estat de gran utilitat i m'han estalviat alguna relliscada aparatosa.

Finalment, l'agraïment a Lluís Payrató, que m'ha aportat l'orientació imprescindible en les fases crucials del procés. La seva experiència, els seus vastos coneixements en aquest camp de recerca i la seva natural serenitat m'han donat la confiança que necessitava —la



seguretat en un mateix i en la feina feta que, en una tasca tan llarga, és tan fàcil de perdre.

I a Rosa Casado, la meva estimada companya de camí, a qui he d'agrair el costat incondicional que m'ha fet i, sobretot, la paciència gairebé benedictina amb què ha suportat la circumstància que, per causa meva, li ha tocat de viure.

A tots els qui he esmentat i als qui involuntàriament hagi pogut ometre, la meva gratitud per l'ajuda i la comprensió que m'han ofert. No cal dir que els errors que, de segur, apareixeran en el text que segueix, són imputables únicament i exclusiva a qui escriu aquestes ratlles, i, alhora, que molts dels encerts que hi pugui haver seran deutors del bon criteri dels col·legues i amics que m'hi han ajudat. Moltes gràcies a totes i a tots.

Mira-sol, 1 d'agost del 2001.



## **RESUM DE LA RECERCA**

L'objectiu d'aquesta tesi, desenvolupada dins de l'àmbit de l'estudi de la variació funcional en el llenguatge, és l'anàlisi comparativa de tres gèneres de discurs en llengua catalana, pel que fa a tres aspectes de l'estructuració i la complexitat lingüística: la densitat lèxica, la composició oracional i la connexió textual. Els gèneres escollits per a l'anàlisi han estat: a/ la conversa informal, com a forma arquetípica de llengua oral; b/ la prosa acadèmica, com a representant de la llengua escrita; i c/ la classe magistral, com a gènere intermediari entre els dos modes, oral i escrit.

La motivació de la recerca ha estat doble. D'una banda, s'ha volgut proporcionar la lingüística catalana amb una descripció, fins ara inexistente, del contrast lingüístic entre oralitat i escriptura. D'altra banda, s'ha pretès fer una aportació a la investigació internacional en aquest camp, tant per intervenir en el debat originat sobre la complexitat sintàctica i, especialment, sobre la presència d'oracions subordinades en cada mode, com per contribuir a una ampliació de la recerca al discurs —en concret, als elements de connexió textual—, que alguns tractadistes han reclamat i que ha estat generalment absent de les anàlisis efectuades fins al moment.

Després de les tasques corresponents de definició inicial i d'ubicació en el marc teòric, de constitució i de transcripció de la mostra, de codificació dels textos, i de còmput i anàlisi de les dades en els trets lingüístics considerats —amb la necessària resolució dels problemes metodològics que originen específicament el treball amb corpus i el maneig de textos orals—, les conclusions principals han estat les següents:

1- La llengua oral té un estil bàsicament verbal de construcció dels enunciats. La llengua escrita, en canvi, el té més nominal i presenta una major densitat i variació lèxiques.

2- La llengua oral i la llengua escrita empren un nombre similar de nexes subordinants, fins i tot amb un lleuger avantatge per a la llengua oral formal expositiva (classe magistral), però, respecte del nombre total de clàusules, és la llengua escrita la qui prefereix clarament la subordinació per a la interrelació clausal. Més enllà de l'àmbit estricte de l'oració, la llengua oral opta proporcionalment més per la connexió textual.

3- La complexitat lingüística, per tant, no és patrimoni d'un dels dos modes, oral o escrit, sinó que es manifesta a cadascun en direccions i amb naturaleses diferents.

Les conclusions reparteixen l'encert entre els dos principals punts de vista, previs a la present recerca, sobre la complexitat de la llengua oral i la llengua escrita, reflectits sobretot en els treballs contraposats de W. Chafe i de M.A.K. Halliday. El conjunt dels resultats ha mostrat que les posicions encontrades troben l'explicació —i, doncs, la possibilitat de reconciliació— en les divergències entre les diverses recerques pel que fa a la naturalesa dels textos analitzats, al caràcter de les unitats lingüístiques seleccionades i, especialment, a la tria del denominador —mot o clàusula— emprat per al còmput. Oralitat i escriptura s'han mostrat, per tant, com les dues cares, sinuoses i polièdriques, igualment complexes i necessàriament complementàries, de la realitat viva que és el llenguatge.

# 1- PRESENTACIÓ

“SENYOR JOURDAIN: I les coses que un hom diu quan parla, què és això?

EL MESTRE DE FILOSOFIA: Prosa.

SENYOR JOURDAIN: Com! Quan jo dic, «Nicolaua, porteu-me les sabatilles i porteu-me l'estrenyecaps», és prosa?

EL MESTRE DE FILOSOFIA: Sí, senyor.

SENYOR JOURDAIN: En bona fe, fa més de quaranta anys que parlo en prosa sense saber-ho poc ni gaire, i us porto la major gratitud del món per haver-m'ho ensenyat.”

Molière, *El burgès gentilhome*, 1670.<sup>1</sup>

El Senyor Jourdain, personatge principal d'*El burgès gentilhome* de Molière, se sorprèn de descobrir que ha estat parlant en prosa tota la vida. Si pogués llegir aquesta tesi, s'adonaria que, de fet, no hi ha parlat mai i que les diferències entre la llengua parlada i l'escripta són moltes més de les que es podria imaginar, tantes que conformen dues varietats lingüístiques diferents i relativament allunyades entre si.

La confusió entre la llengua oral i la prosa només es pot explicar en el context d'una cultura que no estudia ni s'interessa pel llenguatge quotidià. Com diu W. Ong, en un fragment que citaré més extensament en el seu moment, “*literacy is imperious*”. En les nostres societats alfabetitzades, l'escriptura es va arrogar, ja fa segles, una posició preeminent en la concepció del llenguatge, de manera que en els estudis gramaticals i en la valoració social la paraula escrita ha gaudit d'un prestigi enorme i la llengua oral ha estat menystinguda o ignorada. Va caler arribar al segle XX, amb l'aparició del

---

<sup>1</sup> Fragment citat per P. Kay (1977: 29) i per S. Greenbaum i G. Nelson (1995: 1,17). Traducció catalana de Josep Carner.

magnetòfon i d'altres instruments tècnics per a l'enregistrament, perquè la lingüística s'interessés pel llenguatge parlat, que encara avui constitueix una àrea d'investigació incipient.

La primera vegada que es veu la transcripció d'un discurs oral pot suposar fàcilment un dels impactes més grans en l'experiència d'una persona dedicada al món de la llengua. En el meu cas, la investigació de l'oralitat va esdevenir, des d'aleshores, un dels interessos científics primordials. I el treball que presento, en el camp del contrast llengua oral / llengua escrita, vol ser un pas en aquesta direcció.

El fet que em va proporcionar el tema central de recerca va ser la discrepància observada entre diversos autors pel que fa a la complexitat del discurs oral i l'escrit. La concepció tradicional, confirmada per la majoria de lingüistes —d'entre els quals he triat, com a exponent més representatiu, W. Chafe—, és que la llengua escrita és més complexa sintàcticament i lèxicament, mentre que l'oral és més simple en aquests apartats perquè incorpora elements com l'entonació o la comunicació no verbal, que compensen la falta d'elaboració lingüística.

Alguns investigadors, però, exposen un altre punt de vista. Consideren que la llengua oral i l'escrita tenen el mateix grau de complexitat lingüística, però en direccions diferents: l'oral és més complexa gramaticalment i l'escrita ho és lèxicament. Aquesta afirmació, que té l'interès de resultar sorprenent per si mateixa, és, a més, la posició mantinguda per un lingüista de la rellevància de M.A.K. Halliday i pels seus seguidors.

Davant de la magnitud de la contradicció, s'han començat a fer els estudis de variació necessaris per comprovar qui està més encertat i com s'explica que hi hagi opinions tan divergents sobre el mateix tema. La meva tesi forma part d'aquesta línia de recerca i, per això, hi analitzaré tres aspectes fonamentals de la complexitat lingüística en discursos orals i escrits en llengua catalana: la densitat lèxica, la composició oracional i la connexió textual.

L'objectiu, en concret, és el còmput i l'anàlisi de l'aparició d'uns trets lingüístics (lèxics, sintàctics i textuals) en tres gèneres de discurs: la conversa informal, com a forma arquetípica de llengua oral; la classe magistral, com una forma intermèdia d'oralitat amb molta influència de l'escriptura; i la prosa acadèmica, com a representant de la llengua escrita. La comparació entre els tres gèneres ha de permetre d'efectuar les operacions següents:

A/ Caracteritzar cada gènere en els trets lingüístics analitzats, individualment i en relació amb els altres dos gèneres.

B/ Determinar el comportament dels trets lingüístics seleccionats en el contrast entre llengua oral i llengua escrita en llengua catalana.

C/ Cercar una resposta a la contradicció plantejada entorn de la complexitat lingüística —lèxica i sintàctica— en el mode del registre.

El text que segueix s'organitza en sis grans apartats que, al costat d'aquesta presentació, constitueixen els set capítols de què consta:

## **Capítol 1- Presentació.**

## **Capítol 2- Marc teòric**

Fruit d'una revisió crítica de la bibliografia, hi ofereixo, en primer lloc, els elements teòrics i epistemològics necessaris per a l'anàlisi del mode del registre; en segon lloc, una panoràmica de les visions globals del contrast entre oralitat i escriptura, amb la presentació dels paràmetres que s'han considerat més explicatius de la distinció i amb una atenció especial a la noció de gradació o *continuum*; i, en tercer lloc, una anàlisi més aprofundida de les raons oposades, esgrimides pels diversos estudiosos, sobre l'estructuració i la complexitat lingüística en la llengua oral i en la llengua escrita.

### **Capítol 3- Metodologia**

Lluny de constituir un apartat breu i accessori, els aspectes metodològics han merescut una posició central en aquesta recerca i ha calgut un capítol sencer per exposar-los. De fet, és habitual que en una tesi experimental s'hagi de dedicar espai als objectius, les hipòtesis, les variables, les característiques de la mostra, etc. Però, a més, en aquest cas hi ha hagut tres factors especials que han fet encara més imprescindible la presentació i la justificació dels procediments emprats:

a/ La recerca sobre la llengua oral, amb els problemes corresponents de transcripció, convencions, etc.

b/ El treball mitjançant una mostra extensa i recursos de tipus informàtic, amb alguns dels problemes plantejats per la lingüística computacional i la lingüística de corpus.

c/ L'anàlisi d'algunes unitats lingüístiques encara en curs de definició en els camps respectius, amb els problemes lògics de falta de criteris segurs i unànimement acceptats per a la codificació.

### **Capítol 4- Resultats de la recerca (I): densitat lèxica, estil nominal i estil verbal**

La presentació dels resultats obtinguts en la recerca s'ha fet bastant llarga a causa del nombre considerable de xifres i de taules, acompanyades de la discussió dels resultats, d'exemples il·lustratius i de comentaris sobre la casuística, que permeten d'interpretar les dades numèriques amb més propietat. Per aquest motiu, ha calgut dividir-la en dos capítols. En el primer hi ha els trets de tipus més lexical que afecten la densitat lèxica, l'estil nominal i l'estil verbal, i que, en la mesura que intervenen en l'afaiçament d'aquests estils, tenen també clares repercussions oracionals i discursives. Aquests trets lèxics han estat poc discutits, ja que en general constitueixen un terreny d'acord entre la majoria d'autors, però resulten, a més de l'estricta interès descriptiu que tenen, bastant explicatius d'alguns dels fenòmens que es produeixen en el capítol següent.



## **Capítol 5- Resultats de la recerca (II): composició oracional i connexió textual**

Aquest és, lògicament, el capítol més extens de la tesi, ja que conté alguns dels trets principals (subordinació, coordinació, interordinació i connexió textual) que intervenen en la conformació de l'estructura oracional i, en part, de la discursiva. A més, és aquí on es posen en joc els aspectes més polèmics, és a dir, els trets en què es produeixen les divergències sobre la complexitat que he assenyalat.

## **Capítol 6- Conclusions de la recerca**

Es tracta d'un capítol breu en el qual resumeixo i presento més assagísticament els resultats i les conclusions que he anat desgranant de manera precisa i extensa en els dos capítols anteriors. D'altra banda, al final de cadascun dels apartats que componen els capítols 4 i 5, hi ha uns petits subapartats de conclusions, per facilitar el seguiment resumit i progressiu de tot l'aparell quantitatiu i de les conclusions que se'n desprenen a cada pas.

## **Capítol 7- Bibliografia**

Com és propi de qualsevol recerca, el darrer capítol conté la llista completa de les referències bibliogràfiques efectuades al llarg del treball, tant de les obres citades directament com de les remissions aparegudes dins de les citacions d'altri.

## **Annex- Textos de la mostra**

En CD-Rom annex, hi ha la mostra completa dels textos orals i escrits utilitzats en l'anàlisi, amb les convencions de transcripció, diverses dades sobre els paràmetres individuals i contextuals de producció, i la necessària informació sobre la procedència dels textos —salvaguardant, això sí, l'anonimat dels individus seleccionats.

Hi ha, finalment, algunes decisions i convencions relatives a la redacció de la tesi que cal destacar per a un guiatge efectiu de la lectura:

a- Després de força reflexió, he optat per redactar aquest text en primera persona del singular —i no del plural, com és tradicional en l'àmbit català. Al darrere d'aquesta decisió no s'hi ha de veure una adhesió a la manera de fer anglosaxona, sinó un desig de modernització amb l'ús d'un estil menys formal, potser per convertir la solemnitat —o, fins i tot, la pompositat— del gènere “tesi” en una forma de comunicació més pròxima i més humana (em plantejo fins a quin punt l'anomenat *plural de modèstia* no pot acabar essent, si no *majestàtic*, sí una mena de senyal de distinció pretensiosa i gens modesta del nostre grup professional). A més, d'entre els escrits de l'àmbit academicocientífic, n'hi deu haver pocs de tan profundament personals com una tesi. De totes maneres, dins d'aquest estil, per defugir també el possible excés de protagonisme o d'invidualisme que es pugui associar a l'ús del singular, he evitat l'aparició directa del pronom personal *jo* i he restringit a la mínima imprescindible la dels adjectius i pronoms possessius *meu / meva / meus / meves*.

b- He denominat *capítol* la porció de text encapçalada per un sol nombre (p. ex., “2- Marc teòric”), *apartat* la subdivisió del capítol que en té dos (p. ex., “2.1- El contrast llengua oral / llengua escrita: perspectiva històrica”), i *subapartat*, *punt* o, abreviadament, *§*, la subdivisió de l'apartat, que duu tres nombres al davant (p. ex., “2.1.1- Oralitat i escriptura”). He organitzat el text de manera que no hi hagi més que aquests tres nivells de jerarquia i, quan ha calgut fer altres subdivisions internes al subapartat, he optat per les lletres (A, B, C...; a, b, c...) o per nombres isolats, fora de la classificació decimal (1, 2, 3...).

c- M'he referit generalment als nombres amb xifres (p. ex., “74326 ocurrencies”) i no amb lletres (p. ex., “setanta-quatre mil tres-centes vint-i-sis ocurrencies”), òbviament quan apareixen en les llistes o en les taules de resultats, però també quan són dins del redactat, en el cos del text. Aquest criteri no és el més habitual en el món de la filologia —els amants de les paraules solem decantar-nos per les lletres—, però m'ha tombat l'evidència que les xifres es perceben amb més claredat, especialment quan

n'apareixen diverses dins d'un mateix paràgraf o oració i, a més, s'han de comparar entre si. D'altra banda, he optat per la convenció de col·locar un punt davant dels decimals d'un nombre (35.4 es llegeix “trenta-cinc punt quatre”) i de no col·locar cap signe, ni punt ni coma, ni deixar cap espai per indicar la separació entre els milers i la resta de xifres d'una quantitat (2345 es llegeix “dos mil tres-cents quaranta-cinc”). Amb això m'he allunyat tant del sistema tradicional entre nosaltres (coma davant dels decimals: 35,4 i punt darrere els milers: 2.345) com del sistema invers, més habitual a Nord-amèrica (punt davant dels decimals: 35.4 i coma darrere els milers: 2,345).

d- He esmentat els autors que cito al llarg del redactat amb les inicials prèvies al cognom —p. ex., D. Biber (1988) o S.A. Thompson (1984)—, en contra del criteri d'economia i d'impersonalitat que se sol aplicar a aquesta mena de referències. La veritat és que quan cito el treball d'algú m'agrada de pensar, en la línia d'humanització que ja he esmentat, que al darrere hi ha una persona que l'ha concebut i l'ha redactat, i la inclusió de les inicials va en aquesta direcció. He acceptat, però, l'estalvi gràfic quan la referència va entre parèntesis, sovint al final d'una citació, perquè en aquests casos, generalment, el nom amb les inicials ha aparegut poc abans.

e- Com a regla general, fet les citacions en la llengua original dels autors corresponents. En algun cas, però, he treballat amb la traducció catalana o, més sovint, castellana, d'autors anglòfons, francòfons, germanòfons, etc. Quan aquest fet s'ha produït, a la bibliografia final sempre he esmentat en primer lloc la traducció que he utilitzat, i he especificat —a continuació, entre parèntesis i amb la indicació pertinent— l'obra original. També ho he indicat d'aquesta manera quan un treball —llibre o article—, ha estat publicat anteriorment en un altre lloc: he col·locat primer la meua font i, seguidament i entre parèntesis, la referència de la primera publicació. En qualsevol cas, però, la data que he donat entre parèntesis al costat del nom de l'autor és sempre la de la primera publicació del treball, cosa que permet de situar-lo cronològicament de manera adequada respecte d'altres treballs del mateix autor o d'un altre. Amb aquesta mateixa intenció, la majoria de les llistes d'autors que he esmentat dins del cos del text estan ordenades amb un criteri cronològic.

No voldria acabar aquesta introducció sense agrair per endavant l'atenció i l'esforç dels doctes i amables lectors que estan començant a llegir aquest tesi, ni sense demanar-los la seva benevolència pels errors i les inexactituds en què, malgrat l'aplicació, l'empeny i l'afany de minuciositat, hagi pogut incórrer, en un treball que ha representat per al doctorand la tasca de recerca científica i de redacció especialitzada més llarga, més laboriosa, més absorbent i, ahora, més satisfactòria que mai no havia abordat.

## 2. MARC TEÒRIC

### 2.1- EL CONTRAST LLENGUA ORAL / LLENGUA ESCRITA: PERSPECTIVA HISTÒRICA

#### 2.1.1- ORALITAT I ESCRITURA

**“Literacy is imperious. It tends to arrogate to itself supreme power by taking itself as normative for human expression and thought.”**

W.J. Ong, “Writing is a Technology that Restructures Thought”, 1992, pàg. 293.

L’aparició de l’escriptura, fa uns cinc mil anys, va comportar una revolució tecnològica en l’àmbit del llenguatge, que amb el pas dels anys havia de tenir grans conseqüències en la cultura de les societats alfabetitzades (Goody, 1977: § I-V i Ong, 1982: § I-IV). La llengua escrita permetia la transmissió dels missatges a través del temps i de l’espai, donava permanència a les paraules i desplaçava del sentit de l’oïda al sentit de la vista el pes central de l’activitat comunicativa.

El llenguatge oral, auditiu, és envoltant i intern a l’individu (mondèg interior), de manera que en una mentalitat no alfabetitzada es fa difícil, per no dir impossible, distingir entre la llengua i la persona mateixa. La vista, en canvi, focalitza, crea una distància entre el subjecte —escriptor o lector— i l’objecte escrit (el *text* en el sentit llatí de TEXTUM ‘trama, teixit’), i possibilita de concebre el llenguatge com a matèria d’estudi. Aquest procés es veu accentuat en les societats occidentals per l’ús de l’alfabet,

que en relacionar cada so amb una grafia reforça el caràcter analític de la mirada. En la Grècia clàssica, l'ús de l'escriptura alfabètica coincideix, possibilita i, en part, provoca, la invenció d'instruments tan bàsics per a nosaltres com el registre, l'arxiu, la prosa, l'exposició, l'argumentació, la gramàtica, la retòrica, la història, la filosofia i la ciència<sup>2</sup>.

Segons J. Goody (1977 i 1987), el pensament abstracte i descontextualitzat només es desenvolupa plenament en individus alfabetitzats. L'escriptura ha estat un instrument poderós de transformació del pensament i de la societat, i ha fet de la lògica analítica i sil·logística la nostra manera habitual de pensar<sup>3</sup>. L'escriptura és el que va permetre, per tant, el pas del mite al *logos* i va donar lloc al tipus de reflexió i de conceptualització que conformen la visió que avui tenim del món. És per això, per l'enorme importància que

---

<sup>2</sup> W.J. Ong (1982: 19) sosté que la retòrica, encara que es basi en l'expressió oral, no s'hauria donat en una societat sense escriptura i que, de fet, és una tècnica expositiva i argumentativa que segueix la lògica del pensament alfabetitzat. D'altra banda, la filosofia neix amb l'escriptura però mostra des dels inicis (Sòcrates, Plató) un gust per la conversa com a forma de raonament i d'aprenentatge. En el segle XVI, un assagista com M. de Montaigne també la defensa enèrgicament:

“El ejercicio más fructífero y natural de nuestro espíritu es, a mi juicio, la conversación. Encuentro su práctica más dulce que cualquier otra actividad de nuestra vida; ésta es la razón por la que, si se me fuerza a elegir, creo que antes consentiría perder la vista que el oído o el habla. Los atenienses, e incluso los romanos, tenían este ejercicio como un gran honor en sus academias. (...) El estudio de los libros es una actividad mortecina y floja que no estimula, mientras que la plática enseña y ejercita a la vez.” (Montaigne, 1595: 173-174)

Malgrat això, la conversa a què es refereix M. de Montaigne és molt lògica i formal —relacionada amb el que a continuació anomenarem *oralitat secundària*: la dels qui saben escriure. Una atenció real per la conversa quotidiana no es dona fins al segle XX amb l'anomenat “segon” Wittgenstein i amb J.L. Austin i J.R. Searle, de l'escola analítica d'Oxford (vg. J.M. Castellà, 1992: 77-83 i 95-101).

<sup>3</sup> J. Goody explica aquesta revolucionària influència de la llengua escrita sobre el pensament de la manera següent:

“(…) Watt y yo hemos tratado de describir alguna de las realizaciones que habíamos visto como estando estrechamente vinculadas al advenimiento de la escritura y en particular a la invención del sistema alfabético que hizo posible la vasta expansión de la capacidad de escribir. Nosotros sugeríamos que la lógica, «nuestra lógica», en el sentido restringido de un instrumento de los procedimientos analíticos (...) parecía ser una función de la escritura, desde el punto en que era el asentamiento del habla lo que capacitaba claramente al hombre para separar palabras, manipular su orden y para desarrollar formas silogísticas de razonamiento; estas últimas fueron vistas como específicamente escritas más que orales, aún haciendo uso de otro gráfico puramente aislado, la letra, como un medio de indicar la relación entre los elementos constituyentes.” (Goody, 1977: 21)

Cal remarcar, però, que J. Goody (1977) insisteix que no es poden fer dicotomies estrictes entre les societats orals i les escrites, ja que, encara que una capacitat d'abstracció avançada només es pugui desenvolupar en l'àmbit de l'escriptura i de l'educació formal, el món de l'oralitat també admet determinades formes d'abstracció.

adquireix en una cultura com la nostra, que l'escriptura s'ha erigit dins de la mentalitat alfabètica en una condició de normalitat, que W.J. Ong (1992) considera enganyosa:

“The term «illiterate» itself suggests that persons belonging to the class it designates are deviants, defined by something they lack, namely literacy. Moreover, in high-technology cultures —which, more and more, are setting the style for cultures across the world— since literacy is regarded as so unquestionably normative and normal, the deviancy of illiterates tends to be thought of as lack of a simple mechanical skill. Illiterates should learn writing as they learned to tie shoe-laces or to drive a car. Such views of writing as simply a mechanical skill obligatory for all human beings distort our understanding of what is human if only because they block understanding of what natural mental processes are before writing takes possession of consciousness. These views also by the same token block understanding of what writing itself really is.” (Ong, 1992: 293)

Els estudis que he citat de J. Goody i de W.J. Ong prenen en consideració el que s'anomena *oralitat primària*, que és la pròpia de les societats sense escriptura, en contrast amb l'*oralitat secundària*, que es dóna en les societats alfabetitzades:

“(…) llamo «oralidad primaria» a la oralidad de una cultura que carece de todo conocimiento de la escritura o de la impresión. Es «primaria» por el contraste con la «oralidad secundaria» de la actual cultura de alta tecnología, en la cual se mantiene una nueva oralidad mediante el teléfono, la radio, la televisión y otros aparatos electrónicos que para su existencia dependen de la escritura y la impresión. Hoy en día la cultura oral primaria casi no existe en sentido estricto puesto que toda cultura conoce la escritura y tiene alguna experiencia de sus efectos. No obstante, en grados variables muchas culturas y subculturas, aun en un ambiente altamente tecnológico, conservan gran parte del molde mental de la oralidad primaria.” (Ong, 1982: 20)

Així, fent una extensió d'aquesta idea, es pot considerar que la llengua oral dels individus catalanoparlants i, en general, el registre col·loquial de les anomenades “llengües de cultura”, és una mostra de l'oralitat secundària, al costat de la pervivència d'alguns trets de l'oralitat primària. Els textos orals que analitzaré en aquesta tesi, en conseqüència, s'han de contemplar dins d'aquesta combinació dels dos tipus d'oralitat.

## 2.1.2- LA TRADICIÓ EN ELS ESTUDIS LINGÜÍSTICS

**“Traditionally, grammar has always been the grammar of written language; and it has always been a product grammar.”**

M.A.K. Halliday, *An Introduction to Functional Grammar*, 1985, pàg. xxiii.

Com assenyala M.A.K. Halliday en la citació precedent, l'objecte d'estudi de la gramàtica al llarg dels segles ha estat la llengua escrita. Aquest fet no ha de sorprendre, ja que la mateixa etimologia del mot —el grec GRAMMATIKÉ, de GRAMMA 'lletra'— mostra la relació originària entre els estudis gramaticals i el text escrit. La mentalitat alfabètica de les cultures escrites és tan accentuada que tenim ben arrelada en la nostra societat la idea dominant que la llengua és bàsicament l'escriptura, allò consignat en els diccionaris i en les gramàtiques, mentre que la parla és considerada una forma secundària, degradada, inculta i incorrecta. Aquesta concepció que subordina l'oralitat ha influït decisivament en els estudis gramaticals tradicionals, que han tingut la llengua escrita com a únic centre d'interès i que han fet de la prescripció l'objectiu fonamental.

Les principals escoles lingüístiques del segle XX van modificar radicalment els postulats teòrics —encara que, com veurem, no van variar els pràctics. F. de Saussure i L. Bloomfield, fundadors de l'estructuralisme europeu i nord-americà, respectivament, van considerar la llengua oral com la forma primordial de llenguatge i van atorgar a l'escrit un caràcter subsidiari. Vegem, per exemple, aquesta citació ben eloqüent de L. Bloomfield (1933):

“Writing is not language, but merely a way of recording language by means of visible marks.” (Bloomfield, 1933: 21).

D. Biber (1988) forneix altres cites similars a l'anterior, en què lingüistes de tendències diverses afirmen el caràcter secundari de la llengua escrita:

**“E. Sapir:** writing is ‘visual speech symbolism’ (1921: 19-20)



**P. Postal:** ‘writing is a crude way of representing linguistic structure rather than a sign system with a direct relation to the world’ (1966: 91)

**Ch. Fillmore:** written communication is ‘derivative of the face-to-face conversational norm’ (1981: 53)” (Biber, 1988: 6)

La proclamació programàtica de la primacia de la llengua oral dins de l’estructuralisme va provocar el brillant desenvolupament de la fonologia i la fonètica. Aquesta línia de treball, però, no va tenir una correspondència en els altres nivells de l’anàlisi lingüística. Així, en els estudis de morfologia, de sintaxi i de semàntica estructural, la llengua de base és l’escripta, i l’oral, considerada caòtica i poc explicable, no és objecte de consideració.

La gramàtica generativotransformacional, per la seva part, no estableix cap preponderància teòrica de la llengua oral o de l’escripta, sinó que postula l’estudi de la competència lingüística d’un *parlant-oient ideal*, que és definit per N. Chomsky (1965) de la manera següent:

“Lo que concierne primariamente a la teoría lingüística es un hablante-oyente ideal, en una comunidad lingüística del todo homogénea, que sabe su lengua perfectamente y al que no afectan condiciones sin valor gramatical, como son limitaciones de memoria, distracciones, cambios del centro de atención e interés, y errores (característicos o fortuitos) al aplicar su conocimiento de la lengua al uso real.” (Chomsky, 1965: 5)

Malgrat aquesta definició, les formes lingüístiques que s’analitzen en els treballs de gramàtica generativa, obtingudes a partir de la intuïció i la introspecció de l’investigador, són les típiques d’una llengua escrita estàndard o neutra i s’hi sol descartar tant allò que és col·loquial com allò que és literari, en la consideració que són usos particulars del llenguatge que transgredeixen les regles de la gramàtica<sup>4</sup>.

---

<sup>4</sup> Aquest punt de vista mena a una anàlisi descontextualitzada i abstracta del llenguatge, necessària per a determinats fins teòrics, que difereix en gran manera de l’enfocament de la gramàtica funcional:

“Many actual instances can be analyzed in two or more different ways, none of which can be ruled as impossible (...). Especially in a literary text it is to be expected that we will find clauses with multiple grammars; but this is not a distortion of the system —it is a richer use of its natural resources.” (Halliday, 1985c: 54)

El fet és, doncs, que tant si la teoria destaca un dels dos modes de llenguatge per damunt de l'altre —l'escrit o l'oral, sigui quin sigui—, com si en prescindeix i apel·la a una entitat superior més abstracta, a la pràctica es produeix una simplificació excessiva. Oralitat i escriptura són dos fenòmens diversos, cadascun amb una complexitat pròpia, i l'estudi del llenguatge que en prioritza un, en realitat està amagant aspectes de la complexitat d'ambdós. Així, l'afirmació del caràcter primigeni de la llengua oral ha resultat, de fet, en una manca d'interès per l'estudi de les diferències entre oralitat i escriptura. D. Biber (1988) i M.A.K. Halliday (1985a i 1987) expliquen aquest fenomen:

“Assuming this secondary, derivative nature of written language, there was no motivation within structural linguistics for comparison of speech and writing. (...) Even within structural linguistics, researchers have not been entirely consistent regarding the primacy of speech. In particular, there has been a gap between theory and practice in recent syntactic research. In theory, writing is disregarded as secondary and derivative from speech. In practice, however, speech is also disregarded as unsystematic and not representative of the true linguistic structure of a language. This view is especially prominent within the generative-transformational paradigm, where grammatical intuitions are the primary data to be analyzed. Although these intuitions are typically collected by means of verbal elicitation, they are in many respects more like writing than speech.” (Biber, 1988: 6-7).

“The linguists’ professional commitment to the primacy of speech did not, however, arise from or carry with it an awareness of the properties of spoken discourse. It arose from the two sources of diachronic phonology (the study of sound change) and articulatory phonetics (the study of speech production), which came together in twentieth century phonological theory.” (Halliday, 1987: 56)

“(…) Some linguists —or rather, perhaps, philosophers of language— have tended to take over the folk belief, typical of a written culture, according to which spoken language is disorganized and featureless, while only writing shows a wealth of structure and a purity of pattern. This is then ‘demonstrated’ by transcriptions in which speech is reduced to writing and made to look like a dog’s dinner.” (Halliday, 1985a: xxiv)<sup>5</sup>.

---

<sup>5</sup> Uns anys abans, H.A. Gleason, Jr. (1965) ja havia anunciat el perill d'analitzar la transcripció de la llengua oral amb mentalitat de llengua escrita, tot destacant la sorpresa inicial que suposa la visió per primera vegada d'una transcripció de l'oral:

“Most Americans are unaware of the depth and extent of the contrasts between written and spoken English. If an extempore but formal speech is taped and transcribed without editing, the problem becomes immediately evident. Competent appraisal of the speech when heard from the tape and when seen on the paper will be dramatically different. Generally the reaction to the transcription

En aquest panorama, l'estudi científic de la llengua oral, més enllà de la fonologia, no ha començat fins al desenvolupament de la sociolingüística, l'anàlisi del discurs, la lingüística del text, l'etnografia de la parla i la pragmàtica<sup>6</sup>. Totes aquestes disciplines entronquen en major o menor mesura amb la recuperació de la retòrica clàssica, que havia anat perdent-se fins que va caure pràcticament en desús<sup>7</sup>. La retòrica, com a tècnica del discurs efectiu per a la persuasió d'auditoris, és el precedent més il·lustre de l'actual interès per la llengua oral<sup>8</sup>.

---

will be strong rejection or total disbelief. We are not accustomed to seeing spoken deliberative English in unedited written form, and it is so strange when we do see it that we are unable to judge it on its own terms." (Gleason, 1965: 367)

<sup>6</sup> Cal matisar que la lingüística del text té, moltes vegades, una propensió més gran cap a la llengua escrita. Per posar només un exemple ben recent, que és una bona mostra de la "mentalitat escrita" en molts dels estudis sobre el llenguatge, en el treball de M.A. Martín Zorraquino i J. Portolés (1999) [dins la *Gramàtica descriptiva de la llengua espanyola*, Ignacio Bosque i Violeta Demonte (eds.)], es fa una classificació dels marcadors discursius en què, després d'un llarg recorregut per diversos marcadors típics de la llengua escrita, apareixen en un sisè i darrer apartat els "*marcadores conversacionales*", i se'n comença l'explicació corresponent amb l'expressió, ben grafocèntrica, "*la conversación constituye una forma de comunicación peculiar...*" (1999: 4143).

<sup>7</sup> La retòrica va viure l'època d'esplendor en la Grècia i la Roma clàssiques. A l'Edat Mitjana, formava part del *Trivium* (retòrica, gramàtica i filosofia), però durant els segles XVII i XVIII va ser progressivament abandonada en benefici de la gramàtica, fins que va quedar reduïda a les figures retòriques de la literatura que han arribat al nostre sistema escolar. La "data oficial" de defunció de la retòrica és l'any 1885, quan va ser suprimida com a assignatura d'ensenyament secundari pel ministeri francès d'educació (Medina, 1991: 45).

<sup>8</sup> L'existència dels estudis retòrics al llarg dels segles podria semblar contradictòria amb el fet que la llengua oral no hagi estat objecte d'atenció dels gramàtics fins al segle XX. De fet, es tracta d'una qüestió d'àmbit d'estudi. La retòrica és una tècnica aplicada a la construcció de discursos i no un estudi de tipus gramatical. M.A.K. Halliday (1985a) aclareix la relació inicial entre retòrica i gramàtica:

"(...) It seems that in its earliest origins classical Greek grammar was a grammar of speech —the first attempts at syntax were tied to rhetoric, to an explanation of what it is that makes spoken discourse effective. But Aristotle took grammar out of rhetoric into logic; and since then it has been mainly a grammar of written discourse." (Halliday, 1985a: xxiii)

### 2.1.3- LA NECESSITAT DE L'ESTUDI DE LA LLENGUA ORAL I DE LA LLENGUA ESCRITA

**“Speaking and writing are alike —and different. Just how like and how different has never been adequately stated. (...) Little has ever been written of their differences; much has been implied of their likeness; more than the facts warrant.”**

C.H. Woolbert, “Speaking and Writing —a Study of Differences”, 1922, pàg. 271.<sup>9</sup>

**“(...) we know that written and spoken structures differ to some extent, and that there are differences in the underlying grammatical structure, but we have no accurate idea of their magnitude. At some points, they are probably overestimated, at others equally certainly underestimated. Perhaps some have hardly even been noticed.”**

H.A. Gleason, Jr., *Linguistics and English Grammar*, 1965, pàg. 368.

La contradicció de la tradició lingüística que he assenyalat ha estat també tractada per altres autors que, amb matisos i amb orientacions diferents, han reclamat un tractament específic per a aquest àmbit de recerca. Vegeu, a continuació, les paraules d’A. Culioli (1983), W. Grabe (1984), W. Chafe (1992) i A. Teberosky (1998), que confirmen els punts de vista de D. Biber i M.A.K. Halliday citats anteriorment:

“Le français parlé n’intéressait guère, si ce n’est (...) pour montrer (...) que le français parlé n’était pas une forme de langue sur laquelle on pouvait travailler (bien plus, que l’on se devait de tenir à l’écart), car le français parlé était une succession d’amorces avortées, de ratages, de phrases en suspens, qui paraissaient défier l’analyse (...) et surtout qui venaient introduire le désordre dans le beau jardin syntaxique en cours d’élaboration.” (Culioli, 1983: 291-292)

“Until very recently, most linguists have generally assumed the primacy of oral language (Bloomfield 1933, Fillmore 1981, Gardner 1984, Givón 1979a) with little consideration for the apparent differences that exist between oral and written language.” (Grabe, 1984: 102)

---

<sup>9</sup> Fragment citat per R.C. O’Donnell (1974: 102).

“It is a curious fact that we know relatively little about how spoken language works. It is curious in light of the powerful linguistic tradition, extending from Boas through Sapir and Bloomfield, that emphasized—in fact over-emphasized—the primacy of spoken over written language. It is curious in light of the many grammars that have been written of languages that are only spoken, or languages for which the influence of writing has been negligible. (...) In spite of these developments, it remains true that the constraints and regularities inherent in speaking—what we might call the grammar of spoken language—are so little understood. (...) Sapir and Bloomfield seem never to have recognized any real difference between speaking and writing, beyond expressing their belief that the latter is derived from and completely dependent on the former. (...) Ironically, this attitude seems to have led both these linguists to assume that whatever might be written down might also be taken as a perfectly good example of speaking.” (Chafe, 1992: 17-18)

“La descripción lingüística de oralidad y escritura, a lo largo de la historia, ha dado lugar a una gran contradicción: por una parte, considerar el lenguaje hablado como natural y primigenio, la escritura como una modalidad derivada de lo oral y la práctica de escribir su mera transcripción; y atribuir, por otra parte, al lenguaje hablado las características de erróneo, incompleto, no normativo. Apoyada en la oposición entre oral y escrito, esta descripción implicaba la desvalorización de uno y otro de los polos.” (Teberosky, 1998: 9)

Dins d'aquesta situació, lògicament, la que ha sortit més mal parada ha estat la llengua oral. És per això que cal destacar la necessitat de l'estudi de l'oralitat. No es pot continuar negligint el fet incontrovertible que la llengua oral és primigènia, tant en la història de la humanitat com en la vida de cada individu, i que és quantitativament la més usada per la major part de la població en la vida diària, fins i tot en les societats alfabetitzades. A més, milions de persones han viscut sense utilitzar mai la llengua escrita i, encara en l'actualitat, una part molt important de la humanitat utilitza exclusivament la llengua oral per comunicar-se. I, encara avui dia, la majoria de llengües que es parlen en el món no disposen d'un codi ortogràfic i no han produït mai una literatura escrita.

L'oralitat, lluny de ser secundària, és la forma bàsica de llenguatge que es desenvolupa d'una manera espontània en la infància i que està, per tant, indistriciblement lligada a allò que ens fa persones<sup>10</sup>. I, paradoxalment, aquesta forma de llenguatge és la menys

---

<sup>10</sup> K. Perera (1984) fa l'observació següent sobre la relació entre llenguatge oral i llenguatge escrit pel que fa a la immanència:

coneguda i la menys estudiada<sup>11</sup>. És important, doncs, aprofundir en l'anàlisi de la llengua oral i és presumible que l'estudi que se'n faci pugui aportar elements essencials tant per al coneixement del llenguatge com per a la comprensió de la naturalesa comunicativa i social de l'espècie humana.

La investigació de la llengua oral i de la relació que manté amb la llengua escrita ens aboca a una qüestió crucial per a la lingüística: si les gramàtiques que s'han elaborat fins ara es basen en l'escrit, caldria elaborar-ne de basades en l'oral i veure si podrien partir dels mateixos principis que les descripcions existents, de les quals serien meres variants complementàries, o bé si s'arribaria a la constitució d'uns artefactes de mena diferent. Sobre la necessitat de l'elaboració d'aquesta gramàtica i sobre les característiques més o menys diferenciades que tindria argumenten D. Crystal (1987), J.-P. Sueur (1990: 148),

---

"(...) It is clear, I think, that writing is not 'natural' in the way that speech is. Speaking is as fundamental a part of being human as walking upright, but writing is an optional extra —exclusive to human beings, certainly, but not a defining characteristic of the species." (Perera, 1984: 160)

J. Miller (1995), per la seva banda, argumenta amb encert que la conversa i la *narrativa* —en el sentit de 'narració oral breu' que s'utilitza típicament en l'anàlisi del discurs britànica i nord-americana— són gèneres bàsics del repertori humà:

"None the less, it is spontaneous conversation and narrative that children are exposed to when they learn their first language and that most adults use most of the time. These facts are good reasons for regarding spontaneous conversation and narrative as basic in the human linguistic repertoire." (Miller, 1995: 116-117)

<sup>11</sup> En aquest sentit, resulta extraordinàriament reveladora la idea, establerta ja per Aristòtil (Anna Gili, comunicació personal), que allò que és ontològicament anterior és epistemològicament posterior, i viceversa. Efectivament, l'oralitat és, sens dubte, ontològicament —i ontogenèticament, i filogenèticament— anterior, i ha patit un endarreriment epistemològic, respecte de la llengua escrita, d'almenys 3000 anys. D'altra banda, la poca atenció per la vessant oral de les activitats comunicatives ha afectat també camps contigus a la lingüística com ara la psicologia del llenguatge, dins de la qual s'han estudiat més les habilitats escrites (llegir i escriure) que les orals (escoltar i parlar). Vegem el comentari que en fan els psicòlegs F. Ferreira, J. Henderson, M. Anes, P. Weeks i D. McFarlane (1996):

"Most of the useful linguistic information that children receive during the several years preceding literacy is obtained through the auditory modality, and even as adults, the majority of our linguistic experience comes from spoken language. Given its predominance, one might expect that the study of how spoken language is understood would be far more advanced than the study of reading. However, this is clearly not so. Little is known about the processing of spoken language, particularly that of connected speech (i.e., phrases, sentences, and texts), in contrast to the large body of work on the processing of visual language." (Ferreira, Henderson, Anes, Weeks i McFarlane, 1996: 324)

J. Lerot (1991: 148), R. Carter i M. McCarthy (1995: 141) i A. Narbona (1995), el primer i el darrer dels quals cito a continuació:

“In informal speech, in particular, constructions can lack the careful organization we associate with the written language. It is not that conversation lacks grammar: it is simply that the grammar is of a rather different kind, with sentences being particularly difficult to demarcate.” (Crystal, 1987: 94)

“Nadie ha elaborado una auténtica gramática de la variedad coloquial. Hay quien piensa que cuando se lleve a cabo, quedará al descubierto una lengua notablemente distinta de la descrita por los gramáticos. (...) Si se ha reconocido repetidamente (...) que la sintaxis oracional, base y meta del análisis en los modelos estructurales, ha sido precisamente el campo menos y peor investigado, no se entiende muy bien que una y otra vez se intente la caracterización lingüística del español coloquial —que requiere, además, como se ha dicho, la adopción de un punto de vista no lineal y supra o transoracional— desde unos presupuestos insuficientes y en gran medida inadecuados.” (Narbona, 1995: 39)

És, efectivament, tan menor l'atenció que ha rebut la llengua oral col·loquial que fins i tot costa imaginar com seria aquesta *gramàtica de l'oral* i en quina mesura necessitaria unes bases teòriques pròpies, diferenciades de les de l'escrit. En aquest sentit, la present tesi forma part del conjunt enorme de treballs de tipus empíric que caldrà fer per començar la descripció gramatical de l'oralitat abans de poder elaborar-ne una teoria sòlida i ben constituïda<sup>12</sup>.

El fet que la llengua escrita hagi estat la matèria bàsica de la lingüística, però, tampoc no ha permès que l'escriptura en general hagi rebut una atenció adequada. Així, al mateix

---

<sup>12</sup> Per exemple, un aspecte essencial d'aquesta gramàtica de l'oral que aquí només puc apuntar, és el de les vacil·lacions, trencaments, falsos inicis, anacoluts i reformulacions típiques de la parla espontània. Aquests fenòmens, que han estat tractats moltes vegades pejorativament, com una mancança —ho han fet així fins i tot estudiosos de la lingüística del text, com H. Mederos (1988: 10)—, només poden ser “errors” o “incompletituds” des de l'òptica de la llengua escrita. Diversos autors (Hildyard i Olson, 1982: 31-32; Chafe, 1985: 113; Narbona, 1988: 166; Fox i Thompson, 1990: 297-315; Viana, 1990: 84-86; Goodwin, 1994: 198-217; Gülich i Kotschi, 1994: 31-36; Kukkonen, 1995: 317 i Vigara, 1996: 19) han postulat que aquestes vacil·lacions formen part indissoluble de la riquesa i la creativitat en l'arquitectura de la conversa, en la qual la sintaxi és interactiva i l'estructura de les oracions es construeix sobre la marxa i, sovint, en col·laboració entre els interlocutors. El fet més interessant d'aquests fenòmens és que no són percebuts com a error pel receptor, si no és que els veu per escrit o enregistrats, és a dir, amb la distància de l'analista i amb la mentalitat de l'escriptor o del gramàtic. Les vacil·lacions i els anacoluts només poden constituir un problema quan es donen en un gran nombre i, en tot cas, no són un obstacle més gran en la comunicació oral que, per exemple, l'excés de densitat informativa en la comunicació escrita —és a dir, el cas habitual del lector que es veu obligat a rellegir un fragment “espès” per provar de capir-lo millor.

temps que els autors que hem vist denuncien la migrada orientació dels estudis lingüístics cap a la llengua oral, n'hi ha que reclamen una *lingüística de l'escriptura* que contempli determinats aspectes de la complexitat del text escrit (especialment literari) que generalment no es tenen en compte. Un autor en aquesta línia és J. Culler (1987), l'argumentació del qual comença analitzant l'enfocament bàsicament alfabètic sobre el qual s'ha construït la lingüística:

“La idea de oración ambigua —*Vi a Pedro al entrar en el hotel*— parece deber algo al modelo de la escritura: la idea de una oración como secuencia de palabras estipulada fuera de todo contexto y despojada de toda curva de entonación. (...) Las ambigüedades, de las que pretende dar cuenta la lingüística, parecen ser con frecuencia ambigüedades de una oración escrita presentada como ejemplo en un trabajo de lingüística más que ambigüedades de los enunciados, donde las diferencias entre patrones de entonación y entre significados intencionales impedirían a los oyentes confundir lo que parecen secuencias totalmente distintas. (...) La gramática generativo-transformacional parece tener más sentido como lingüística de la escritura que como lingüística del habla.” (Culler, 1987: 182)

A continuació, J. Culler destaca que hi ha característiques essencials de l'escriptura mateixa que també han estat deixades de banda i relegades a la marginalitat, i considera que són precisament aquests els aspectes en què hauria de fonamentar-se la “nova lingüística”, un estudi del llenguatge realment centrat en l'especificitat de l'escriptura<sup>13</sup>:

“Una lingüística de la escritura puede proceder prestando atención (...) a las posibles relaciones y patrones relegados a los márgenes de la lingüística. Los actos de relegación, que a menudo suponen una curiosa imposición, por las buenas o por las malas, de una ideología del lenguaje, pueden ser excelentes puntos de partida para una nueva lingüística. (...) Ante todo, tendría que atender a lo que queda fuera de la lingüística corriente pero proporciona gran parte del material de la crítica literaria. (...) La tarea de la lingüística ha

---

<sup>13</sup> La demanda de J. Culler ha tingut, almenys parcialment, la seva resposta en l'anàlisi del discurs i, sobretot, en la lingüística del text, disciplines que s'han aplicat àmpliament a l'estudi dels textos literaris (en són una bona mostra els treballs en llengua castellana de T. Albadalejo i d'A. García Berrio). En un altre lloc, el mateix J. Culler assenyala que la lingüística del text pot tenir molt d'interès per als estudis literaris, encara que no sigui, com ell mateix diu, la panacea:

“Moreover, if linguistics succeeds in moving beyond the sentence and characterizing in a systematic way the linguistic relations that produce coherent discourses or texts, the results will be of interest to the critic and particularly to the student of poetics; but they will not be a panacea.” (Culler, 1982: 21-22)



sido dividir lo significativo de lo no significativo, excluyendo esto último de la lingüística, pero si esta región fronteriza es central en el lenguaje y en su funcionamiento —y textos como *Finnegans Wake* dan a entender que lo es<sup>14</sup>— entonces se debe revisar su geografía (...)” (Culler, 1987: 191-192)<sup>15</sup>

En resum, la lliçó és clara: quan la lingüística s’orienta més cap a un mode i obvia l’altre, es troba en una contradicció que acaba no satisfent ningú, ni tan sols els qui es dediquen a l’estudi del mode prioritzat. En aquest aspecte, la tesi present s’inscriu en una línia de recerca que vol trobar les especificitats del mode oral i del mode escrit del llenguatge, tot prestant al discurs oral una atenció de la qual no havia gaudit en èpoques anteriors i comparant-lo amb la llengua escrita en un pla d’igualtat<sup>16</sup>. Com afirma D. Biber (1988: 9) en un paràgraf que resumeix d’una manera especialment encertada la present opció i que serveix com una bona conclusió per a aquesta argumentació, “*they are simply different*”:

“(...) Neither speech nor writing needs to be considered primary to the exclusion of the other. Rather, both require analysis, and the linguistic comparison of the two modes becomes an important question. Of course, in terms of human development, speech has primary status. (...) Once a culture has developed written communication, however, there

---

<sup>14</sup> Es tracta de l’obra *Finnegans Wake* de J. Joyce, que es caracteritza per l’exploració lingüística i per un llenguatge complex al qual J. Culler es refereix diverses vegades al llarg del seu article.

<sup>15</sup> J. Culler (1987), encara que estigui sobretot interessat per la lingüística de l’escriptura, acaba donant algunes pistes ben valuoses per a l’estudi de l’oralitat i, en concret, de la prosòdia:

“Una lingüística que tratase de abarcar todas las curvas de vacilación y énfasis, los tonos de voz que funcionan como operadores modales para indicar seguridad, agresividad, modestia, etc., así como todas las variantes dialectales del habla que son portadoras de información social, sería claramente más complicada que la que tratase ante todo de asignar las descripciones gramaticales correctas a oraciones consideradas como objetos ideales. Semejante lingüística, teniendo en cuenta las cualidades significativas potenciales del habla —el habla como gesto, diríamos—, se estaría enfrentando de hecho a cierta textualidad de la voz: diferencias significativas potenciales no fácilmente reducibles a sistemas de convención.” (Culler, 1987: 183)

<sup>16</sup> Una raó important per a aquesta elecció és també que l’anàlisi lingüística al nivell textual demana de tenir en compte la diversitat i, per tant, la distinció entre discurs oral i discurs escrit, com assenyala R. de Beaugrande (1984):

“It is now recognized that writing and speech are distinct systems: the norms and functions of each system fit a given situation better than those of the other. (...) There is hardly a dimension of textuality or text processing that does not reflect some significant difference between the spoken and the written modalities”. (Beaugrande, 1984: 256-257)

is no reason to regard writing as secondary within that context. (...) From this perspective, neither can be said to be primary; they are simply different.” (Biber, 1988: 8-9)

Com s’ha vist en les citacions inicials d’aquest punt, ja l’any 1922 C.H. Woolbert assenyalava l’escàs coneixement que es tenia de les diferències entre llengua oral i llengua escrita. Al 1965, H.A. Gleason Jr. repetia més o menys el mateix. I quasi vuitanta anys després de la primera d’aquestes afirmacions, als inicis del segle XXI, els viaranys i les contradiccions de la ciència lingüística han fet que la situació no hagi canviat prou —o, almenys, no prou decisivament— i la imperiosa necessitat de fer recerca en aquest camp continua gairebé intacta.

## 2.2- EL MODE DEL REGISTRE

### 2.2.1- REGISTRE, CANAL I MODE

**“Language is not realized in the abstract: it is realized as the activity of people in situations, as linguistic events which are manifested in a particular dialect and register.”**

M.A.K. Halliday, A. McIntosh i P. Strevens, “The Users and Uses of Language”, 1964, pàg. 151.

El marc teòric més usual per a l'estudi de l'oposició oral / escrit és la noció sociolingüística de *registre*<sup>17</sup>. Vegeu, a continuació, unes citacions de M.A.K. Halliday, A. McIntosh i P. Strevens (1964); J. Ure i J. Ellis (1974); M.A.K. Halliday i R. Hasan (1976) i M.A.K. Halliday (1978), que resulten pertinents per a la definició d'aquest concepte:

“Under the second question come these subdivisions: varieties according to users (that is, varieties in the sense that each speaker uses one variety and uses it all the time) and varieties according to use (that is, in the sense that each speaker has a range of varieties and chooses between them at different times). The variety according to users is a DIALECT; the variety according to use is a REGISTER.” (Halliday, McIntosh i Strevens, 1964: 151)

---

<sup>17</sup> Segons J. Ure i J. Ellis (1974: 116) i I. Marí (1994: 24), el terme *registre* va ser utilitzat per primera vegada per T.B. Reid l'any 1956 i ha estat àmpliament utilitzat i difós per lingüistes, sociolingüistes i antropòlegs. M.A.K. Halliday ha estat un dels autors que més ha contribuït a definir-lo. D'altra banda, la denominació *registre* entra en conflicte amb altres termes com *gènere*, *tipus de text* i *estil*, que comentaré més endavant (vg. § 3.2.2).

“Un registro es una subdivisión de una lengua dada. Es una ‘variedad situacional’, constituida por una selección de preferencias, de entre el total de las opciones lingüísticas que ofrece esa lengua específica.” (Ure i Ellis, 1974: 116)

“The linguistic features which are typically associated with a configuration of situational features —with particular values of the field, mode and tenor— constitute a REGISTER.” (Halliday i Hasan, 1976: 22)

“La noción de registro es a la vez muy simple y muy importante: se refiere al hecho de que la lengua que hablamos o escribimos varía de acuerdo con el tipo de situación. (...) La noción de registro constituye así una forma de predicción: dado que conocemos la situación, el contexto social de utilización del lenguaje, podemos predecir mucho respecto del lenguaje que se producirá, con probabilidades razonables de estar en lo cierto. (...) Lo que necesitamos conocer respecto de un contexto de situación a fin de predecir las características lingüísticas que probablemente se vinculan a él se ha resumido en tres encabezados: el ‘campo del discurso’, el ‘tenor del discurso’ y el ‘modo del discurso’.” (Halliday, 1978: 46-48)

Tal com s’indica en les definicions precedents, hi ha uns paràmetres contextuals (situacionals i culturals) que condicionen l’aparició d’uns determinats trets lingüístics en els textos. Els registres són *els conjunts de característiques lingüístiques* associades a aquests *paràmetres contextuals*, és a dir, als diversos usos socials del llenguatge. La distinció oral / escrit, en la mesura que és una característica relativa al canal de comunicació, forma part del *mode del registre*:

“The ‘context of situation’ of the text is described in terms of field —the nature of the social semiotic activity—, tenor —the statuses and mutual roles of the interactants— and mode —the rhetorical and communicative channels. (...) This features determine the choice of ‘register’; that is, the kinds of meanings that are likely to be exchanged.”<sup>18</sup> (Halliday, 1985c: 53)

“The FIELD is the total event, in which the text is functioning, together with the purposive activity of the speaker or writer; it thus includes the subject-matter as one element in it. The MODE is the function of the text in the event, including therefore both the channel taken by the language —spoken or written, extempore or prepared— and its genre, or rhetorical mode, as narrative, didactic, persuasive, ‘phatic communion’ and so

---

<sup>18</sup> En aquest fragment i el següent, M.A.K. Halliday referma la idea, que tornaré a esmentar en el punt 3.2.2, segons la qual els *canals retòrics* —expressió amb la qual es refereix als gèneres i als tipus de text— també formen part del mode del registre.

on. The TENOR refers to the type of role interaction, the set of relevant social relations, permanent and temporary, among the participants involved. Field, mode and tenor collectively define the context of situation of a text.” (Halliday i Hasan, 1976: 22)

”This leads to ‘mode of discourse’, since this refers to the medium or mode of the language activity, and it is this that determines, or rather correlates with, the role played by the language activity in the situation. The primary distinction on this dimension is that into spoken and written language, the two having, by and large, different situational roles.” (Halliday, McIntosh i Strevens, 1964: 153)

El mode del registre, d’acord amb la mateixa definició de M.A.K. Halliday, A. McIntosh i P. Strevens, es refereix no només al canal físic de la comunicació, sinó que engloba un bon nombre de factors situacionals que formen part de la *language activity*. Cal no confondre, per tant, les nocions de *canal* i de *mode*:

**Canal:** sistema físic a través del qual es realitza la comunicació (p. ex.: *canal acústic, canal gràfic, canal telefònic, canal televisiu, etc.*).

**Mode:** estil lingüístic i discursiu que està determinat per les condicions físiques de producció i de recepció del text (p. ex.: *mode oral, mode escrit*<sup>19</sup>).

La distinció entre canal i mode és reflectida de manera gràfica pels romanistes alemanys P. Koch i W. Oesterreicher (1990), en un estudi comparatiu del registre col·loquial en tres llengües romàniques (francès, italià i espanyol). Aquests autors distingeixen entre *konzeption* (= mode o estil) i *medium* (= canal), en un quadre de doble entrada:

---

<sup>19</sup> Els mots *oral* i *escrit* tenen un ús bastant indefinit, fins i tot quan són usats en l’àmbit de la lingüística, perquè poden indicar el canal de transmissió o bé el mode o estil lingüístic que hi està associat. M.A.K. Halliday assenyala aquesta ambigüitat:

“(…) The categories of ‘written’ and ‘spoken’ are themselves highly indeterminate —they may refer to the medium in which the text was originally produced, or the medium for which it was intended, or in which it is performed in a particular instance; or not to the medium at all, but to other properties of a text which are seen as characteristic of the medium.” (Halliday, 1987: 66)

La distinció proposada entre canal i mode pot resoldre la indefinició descrita, especialment si s’usen els termes *fònic* i *gràfic* en relació amb el canal i *oral* i *escrit* es reserven preferentment per a l’especificació del mode.

## KONZEPTION (MODE)

		<b>Oral</b>	<b>Escrit</b>
<b>MEDIUM (CANAL)</b>	<b>Codi gràfic</b>	faut pas le dire	il ne faut pas le dire
	<b>Codi fònic</b>	/fopal'di:r/	/ilnðfopal'di:r/

(Koch i Oesterreicher, 1990: 5)

El canal, doncs, com a part de la configuració contextual, és només un dels paràmetres que incideixen en el mode del registre, al costat d'altres amb els quals constitueix el conjunt de les *condicions físiques o situacionals de producció i de recepció* del text. Aquests altres paràmetres poden ser els següents:

- Espontaneïtat / preparació (o planificació / no planificació).
- Monogestió / poligestió (o discurs monologat / discurs dialogat).
- Unidireccionalitat (= no hi ha possibilitat de resposta per part del receptor) / multidireccionalitat (= hi ha llibertat de torns de paraula).
- Simultaneïtat en l'espai / no simultaneïtat en l'espai.
- Simultaneïtat en el temps / no simultaneïtat en el temps (o recepció en temps real / recepció diferida).

La combinació d'aquests factors fa aparèixer uns grups de característiques lingüístiques que s'anomenen *mode o registre oral* i *mode o registre escrit*. Oral i escrit són, per tant, estils discursius resultants de la influència d'una suma de factors<sup>20</sup>. S'han elaborat força llistes d'aquests factors i dels trets que caracteritzen la llengua oral i la llengua escrita.

---

<sup>20</sup> Hi ha encara més elements que incideixen en les tries del mode i que comentaré més endavant: les altres característiques del registre (camp, to i tenor), el gènere, el tipus de text, etc.

R. Horowitz i S.J. Samuels (1987), en una revisió de la recerca en aquest camp, n'ofereixen una de particularment sintètica i completa<sup>21</sup>:

<b>THE ORAL-WRITTEN DICHOTOMY</b>	
<b>Oral language – <i>Talk</i></b>	<b>Written language - <i>Text</i></b>
<i>Face-to-face conversation</i>	<i>Face-to-text</i>
<i>Reciprocity speaker-listener</i>	<i>Limited reciprocity author-reader</i>
<i>Narrativelike</i>	<i>Expositorylike</i>
<i>Action-Event-Story oriented</i>	<i>Idea-Argument oriented</i>
<i>Here and now in given space-time</i>	<i>Future and past not space-time bound</i>
<i>Informal</i>	<i>Formal</i>
<i>Primary discourse</i>	<i>Secondary discourse</i>
<i>Natural communication</i>	<i>Artificial communication</i>
<i>Interpersonal</i>	<i>Objective and distanced</i>
<i>Spontaneous</i>	<i>Planned</i>
<i>Sharing of situational context</i>	<i>No common situational context</i>
<i>Ellipsis</i>	<i>Explicitness in text</i>
<i>Structureless</i>	<i>Highly structured</i>
<i>Cohesion through paralinguistic cues</i>	<i>Cohesion through lexical cues</i>
<i>Simple predication</i>	<i>Multiple predication</i>
<i>Repetition</i>	<i>Succinctness</i>
<i>Simple linear structures</i>	<i>Complex hierarchical structures</i>
<i>Paratactic patterns</i>	<i>Hypotactic patterns</i>
<i>Limited subordination</i>	<i>Multiple subordination</i>
<i>Fleeting</i>	<i>Permanent</i>
<i>Unconscious</i>	<i>Conscious-restructures consciousness</i>

(Horowitz i Samuels, 1987b: 9)

<sup>21</sup> En català, D. Cassany (1987: 31-35) en presenta una llista més extensa, en la qual distingeix entre trets textuais i trets contextuais.

## 2.2.2- LA GRADACIÓ LLENGUA ORAL / LLENGUA ESCRITA

**“There are relatively few absolute and clearcut categories in language; there are many tendencies, continuities, and overlaps.”**

M.A.K. Halliday, “Dimensions of Discourse Analysis: Grammar”, 1985c, pàg. 54.

El títol mateix del quadre de R. Horowitz i S.J. Samuels (1987) que acabo d’oferir indica una *dicotomia* entre els dos modes, oral i escrit. La confrontació dicotòmica, que com assenyala M.A.K. Halliday (1985) és explicativa de molt pocs fenòmens dins del llenguatge, ha estat molt emprada en aquest camp de recerca però no és la visió més adequada<sup>22</sup>. En realitat, la relació que mantenen oralitat i escriptura *no és dicotòmica sinó gradual*. Així, les característiques de la columna de l’esquerra del quadre R. Horowitz i S.J. Samuels, atribuïdes sovint a la llengua oral en general, s’han de limitar a la conversa informal espontània, que es considera la forma prototípica de la llengua oral, i les característiques de la columna de la dreta, s’han d’atribuir a alguns gèneres escrits, però no pas a tots<sup>23</sup>.

---

<sup>22</sup> D. Biber (1988) assenyala aquesta idea, citant al seu torn D. Hymes (1974):

“Most previous studies have treated linguistic variation in terms of dichotomous distinctions rather than continuous scales. There is no reason, however, to expect that dimensions of variation should be dichotomous. Hymes (1974: 41) points out that «the fact that present dimensions consist so largely of dichotomies (...) shows how preliminary is the stage at which we work». The situation changed little in the decade since Hymes wrote.” (Biber, 1988: 22)

<sup>23</sup> M.J. Cuenca (comunicació personal) assenyala, en una visió d’inspiració clarament cognitivista, que la gradualitat és un fenomen central en la naturalesa del llenguatge:

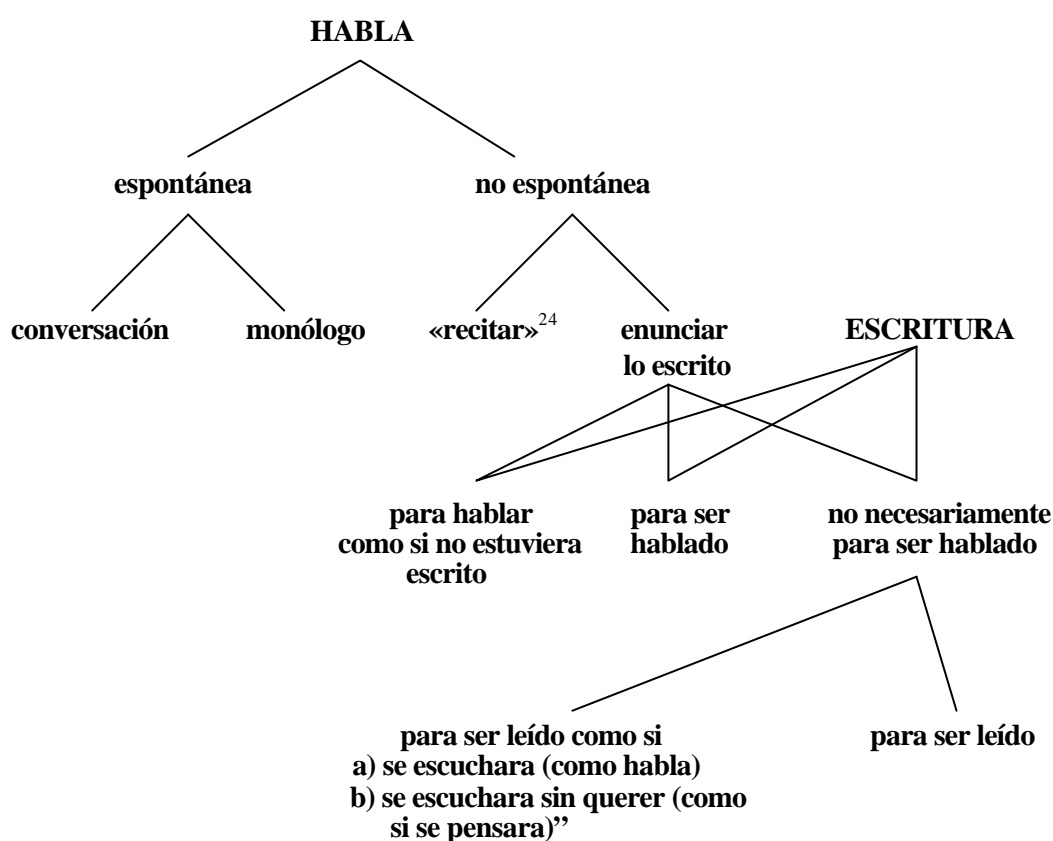
“No només les relacions entre oralitat i escriptura són graduals. Gairebé totes les relacions ho són. D’aquí la importància d’adoptar un enfocament prototípic, no com una excepció o una solució d’urgència, sinó com la cosa més normal, com a part central de la metodologia. Si no és així, et veus sempre en situacions problemàtiques que, d’aquesta manera, deixen de ser un problema per ser la manifestació d’un tret inherent del pensament humà. La definició per feixos de trets dóna una explicació més satisfactòria de tot plegat i, a més, explica que es puguin haver proposat teories aparentment contradictòries del mateix fet.” (Cuenca, correu-e, 18.08.2000)

Per la seva banda, M.T. Cabré (també comunicació personal), des de la perspectiva terminològica i dels llenguatges d’especialitat —i amb un enfocament, de fet, també cognitiu— considera que el tret *especialitzat / no especialitzat* no és de caràcter gradual, sinó que correspon a la inclusió o no d’un text en les complexes estructures conceptuals que constitueixen les ciències i les tècniques. A parer meu, el tret



L'oposició oralitat / escriptura, lluny de ser dicotòmica, dóna lloc a múltiples possibilitats combinatòries, que han estat apuntades, en un quadre esdevingut ja clàssic, per M. Gregory i S. Carroll (1978):

**“Distinciones sugeridas en la dimensión de variaciones de situación, clasificadas como la relación entre el usuario y el medio utilizado.**



(Gregory i Carroll, 1978: 82)

---

*especialitzat / no especialitzat* pot ser dicotòmic només en la mesura que es consideri un paràmetre contextual —com, per exemple, el canal físic de transmissió, acústic o gràfic— però la seva manifestació en el text, en combinació amb altres factors contextuals, hauria de ser gradual. La qüestió del *continuum* és, per tant, un aspecte important en la descripció del llenguatge, que pot provocar, com veiem, una discussió teòrica ben interessant.

<sup>24</sup> M. Gregory i S. Carroll escriuen entre cometes aquest mot de la graella, sense cap indicació del perquè i sense que sembli fàcil d'entendre'n la intenció.

La descripció de M. Gregory i S. Carroll, que té el mèrit de fer palesa la diversitat en el mode del registre, pot ser objecte d'algunes matisacions. En primer lloc, és més una presentació de les combinacions que es donen entre el discurs escrit i la possible execució oral (*escrit per ser dit com si no fos escrit, escrit no necessàriament per ser dit*, etc.), que no pas una descripció completa del *continuum* oral / escrit. En segon lloc, una aplicació exacta del quadre deixaria fora *els discursos escrits provinents d'un discurs oral*, com els apunts de classe, l'acta de reunió, les entrevistes publicades, etc. i, d'altra banda, *els discursos orals no (totalment) espontanis*, sovint formals i expositius, com una conferència, una classe magistral, una homilia, un míting, un debat parlamentari, una tertúlia radiofònica o televisiva, etc., que no poden considerar-se monòlegs espontanis i tampoc no són la recitació ni l'enunciació d'un escrit<sup>25</sup>. Una llista més detallada d'aquests discursos o gèneres intermedis entre la llengua oral i l'escrita és la que es proposa a J.M. Castellà (1996a: 262-263)<sup>26</sup>.

Les possibles mancances en l'esquema de M. Gregory i S. Carroll (1978), que no solen posar-se gaire de relleu, queden subsanades en l'obra de J. Munby (1978), que planteja 11

---

<sup>25</sup> M. Gregory i S. Carroll, en la seva explicació, consideren que una conferència estaria inclosa en la categoria *monólogo espontáneo*, tenint en compte que "(...) desde luego, hay un continuo de mayor a menor espontaneidad en esta específica relación con el medio." (Gregory i Carroll, 1978: 72). Aquesta precisió no apareix reflectida, però, en el quadre. De fet, potser el punt més flac de la proposta és la distinció *espontani / no espontani* que presideix la proposta de classificació i que, com comentaré més endavant (vg. § 2.2.3), resulta difícil de definir i de concretar.

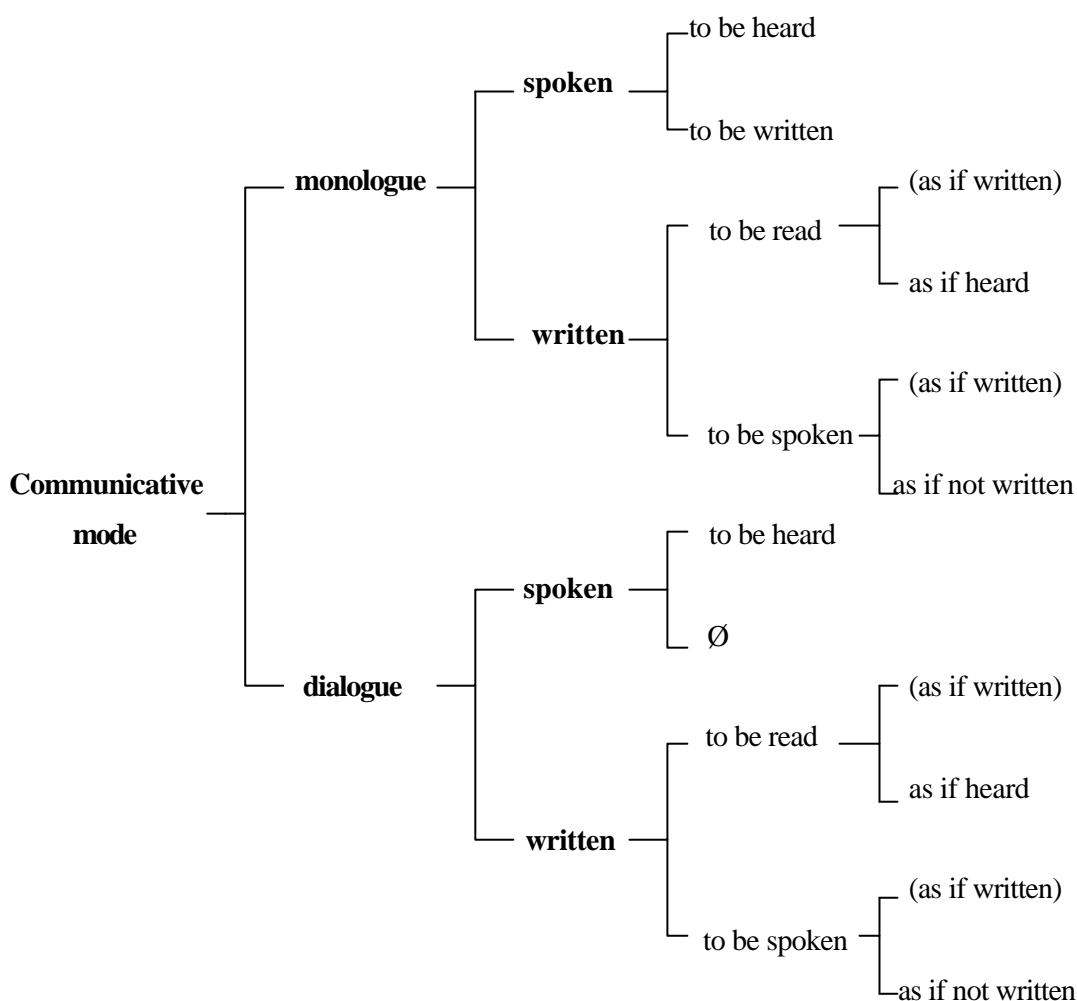
<sup>26</sup> Hi ha estudis de publicació recent que analitzen alguns dels gèneres intermedis. W. Vagle (1991), en un article amb el títol ben indicatiu del tema que estem tractant ("Radio Language - Spoken or Written?"), analitza els discursos radiofònics i els situa entre l'oralitat i l'escriptura, amb la precisió que han d'ocupar un lloc especial respecte del *continuum*, com a representants d'una oralitat secundària, que ell anomena "*a new non-literacy*", superposada a la llengua escrita de les societats dels mitjans de comunicació. D. Du Bartell (1995), per la seva part, estudia el novíssim registre de la comunicació a través d'ordinadors (*e-mail*, etc.) i hi observa també característiques pròpies dels dos modes, oral i escrit, al costat d'algunes que creu que són exclusives del canal informàtic:

"The linguistic and paralinguistic 'peculiarities' of CMC (Computer-Mediated Communication), the spoken and written-like characteristics in a written medium, result from the machine architecture, or in other words, from the constraints imposed by the computer medium. The constraints have caused the development of conversational and lexico-grammatical properties unique to CMC discourse." (Du Bartell, 1995: 238)

Aquest camp ha començat a ser investigat recentment en català en els treballs de M. Torres (1999) i L. Payrató i M. Torres (2000), en els quals també es remarca la fusió de característiques de l'oralitat i de l'escriptura.

subapartats dins del mode del registre, que cobreixen satisfactòriament qualsevol situació comunicativa<sup>27</sup>. Vegem-ho a continuació en el seu quadre i en l'explicació que l'acompanya:

“Figure 11 Classification of communicative mode

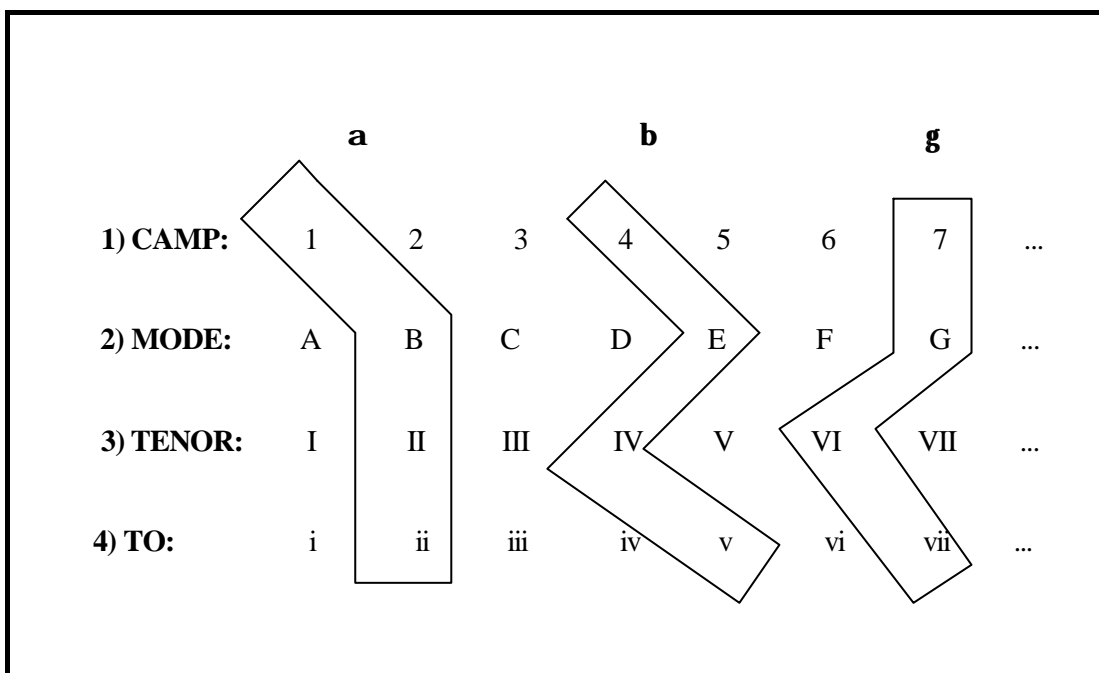


The most common subcategories are *monologue, spoken to be heard*, the mode for a lecture and the expository part of most class-teaching; *monologue, written to be read* (as if written), as exemplified by texts of books of non-fiction or journals; and *dialogue, spoken to be heard*, as exemplified by conversation and discussion. Less common but still applicable to some likely inputs are *monologue, spoken to be written*, the mode in which dictation would be

<sup>27</sup> De fet, l'esquema de M. Gregory i S. Carroll ha tingut molta més fortuna que el de J. Munby —el qual, paradoxalment, és més complet— i ha estat reproduït, sovint mimèticament, en obres de tipus general i especialitzat, de sociolingüística i de didàctica de la llengua, sense una reflexió sobre l'abast i el grau de concreció que se li pot atorgar.

given to secretaries; *monologue, written to be spoken* (as if written), the mode for news bulletins and airport announcements; *dialogue, written to be read* (as if written), the mode for correspondence; and *dialogue / monologue, written to be spoken as if not written*, the modes used for the scripts for plays, films, and television interviews (dialogue, and certain kinds of public speeches (monologue). The remaining three modes are unlikely to be in much demand: *dialogue, written to be spoken* (as if written), as exemplified by the written reporting of a conversation which is to be read aloud as documentary evidence in court; and *dialogue / monologue, written to be read as if heard*, as exemplified by passages of direct speech in novels (...).” (Munby, 1978: 76-77)

Veiem, doncs, que, més enllà d’aquestes precisions, tant un esquema com l’altre mostren una notable potencialitat combinatòria entre oralitat i escriptura, i confirmen que la dimensió oral / escrit en l’eix del mode del registre no és pas una oposició dicotòmica, sinó un *continuum* o fenomen gradual. La concepció de gradació, que ja és en la proposta dels registres des de l’inici, n’afecta tots els apartats, com es pot veure d’una manera gràfica en el quadre següent, proposat per L. Payrató (1988):



(Payrató, 1988: 47)

Així doncs, un registre convencional —col·loquial, administratiu, etc.— i, fins i tot, un text qualsevol, és el resultat de una configuració particular de punts diferents del *continuum* en cadascun dels eixos que el defineixen —camp, mode, tenor i to<sup>28</sup>. La idea de gradualitat resulta, per tant, central en el model i és suggerida per molts dels autors que han tractat recentment l'oposició llengua oral / llengua escrita. Per exemple, A. Tuson (1991: 16-17), M. McCarthy (1993: 170-173) i C. Blanche-Benveniste (1994: 63), parlen de la presència de l'oralitat en l'escriptura i de la interrelació entre ambdós modes. D. Tannen (1982b, 1987 i 1992) planteja els elements de proximitat entre la conversa quotidiana i el discurs literari<sup>29</sup>, idea que també formula R. Lakoff (1982), que afirma que només alguns dels seus escriptors preferits són *literaris* —en el sentit de *lletrats*<sup>30</sup>—, ja que la majoria combinen estratègies típiques de l'oralitat i de l'escriptura en les seves obres<sup>31</sup>. En el

---

<sup>28</sup> Hi ha una certa diversitat en la enumeració de les característiques constitutives dels registres. És habitual, per exemple, de veure el *tenor* de M.A.K. Halliday subdividit en *tenor interpersonal* i *tenor funcional* o, com aquí, en *tenor i to*.

<sup>29</sup> D. Tannen (1982b) defensa aquesta idea de la manera següent:

“Comparative analysis of spoken and written versions of a narrative demonstrates (1) that features which have been identified as characterizing oral discourse are also found in written discourse, and (2) that the written story combines syntactic complexity expected in writing with features which create involvement expected in speaking. (...) Imaginative literature exhibits typically written, integrated prose; but it simultaneously elaborates strategies associated with speaking, in order to create involvement. Thus imaginative literature, far from being maximally different from ordinary conversation, elaborates and refines features which are spontaneously produced in everyday conversation.” (Tannen, 1982b: 1-19)

<sup>30</sup> R. Lakoff (1982) juga en el títol del seu article (“Some of my Favorite Writers are Literate: The Mingling of Oral and Literate Strategies in Written Communication”) amb el contingut de la paraula *literate* en anglès, que pot estar relacionada amb *literacy* i significar ‘el qui ha estat educat per llegir i escriure’ o amb *literature* i significar ‘el qui escriu textos literaris (poesia, novel·la, etc.)’.

<sup>31</sup> M. McCarthy (1993) reprèn la qüestió, esmenta també D. Tannen i arriba a afirmar que els textos escrits no són menys interactius que els orals —potser seria millor dir que tenen un tipus d'interactivitat diferent. M. McCarthy acaba amb un petit joc lingüístic:

“In this chapter I have looked at the occurrence of spoken discourse markers in written texts taken from journalistic, advertising and literary sources, and argued that they play a major role in our judgement of the degree of spokenness present in the text. I offer it as a small contribution to the general argument that written text is no less interactive than spoken text, that terms such as ‘spoken’ and ‘written’ can often be misleading, as can ‘monologue’ and ‘dialogue’, and that literary texts no less than advertising texts exploit creatively the very stuff of everyday conversation, a point nowhere put more persuasively (...) than in Tannen (1989). Well, that’s all, then. And don’t want say any more. So, that’s that.” (McCarthy, 1993: 180)

Amb “Tannen (1989)”, M. McCarthy es refereix a D. Tannen (1987): “Repetition in Conversation: Towards a Poetics of Talk”, article aparegut a *Language*, núm. 63, i tornat a publicar dos anys després

nostre entorn, A. Briz i E. Serra (1997) també assenyalen el caràcter continu del contrast oral / escrit i posen en joc les nocions de canal i de mode que he presentat anteriorment:

“No hay duda de las diferencias polares entre una transmisión oral y una transmisión escrita, es decir, entre la vocalidad o producción y recepción de un mensaje por el canal fónico y la producida u obtenida por el canal gráfico, en cuanto al tiempo de ejecución, a las posibilidades y formas de planificar, de construir el mensaje, al espacio de emisión y recepción, a la perdurabilidad, etc. Pero esa oposición tajante se convierte en un *continuum* gradual cuando nos referimos, no ya al medio o canal, sino a los modos o realizaciones de lo oral y de lo escrito. Así, junto a los extremos, *lo oral* y *lo escrito*, encontramos manifestaciones o reflejos diversos de *lo oral en lo escrito*, que convenimos en llamar en abstracto *oralidad*, y de *lo escrito en lo oral* o *escrituridad*.” (Briz i Serra, 1997: 1)<sup>32</sup>

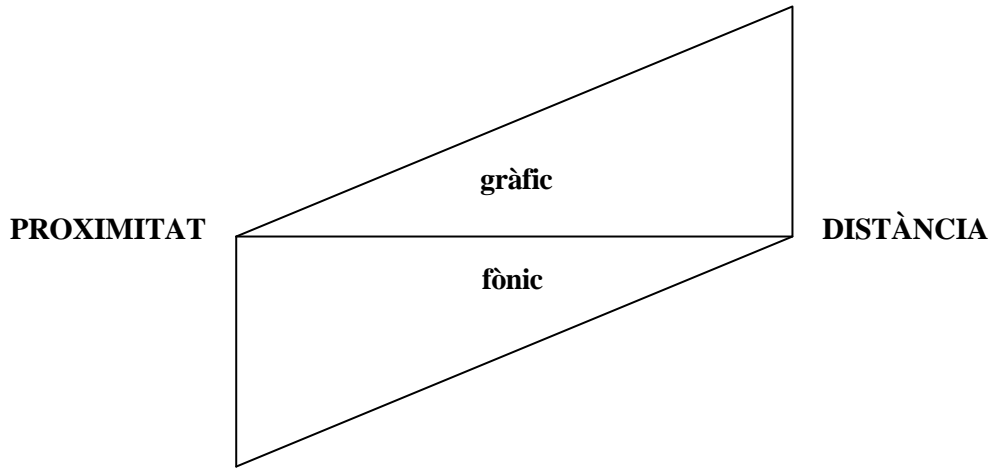
Una visió particularment interessant de la gradualitat és la que ofereixen P. Koch i W. Oesterreicher (1990). En la seva obra, ja citada anteriorment, presenten una representació gràfica que inclou les diferències de canal i les d'altres paràmetres, agrupats sota els conceptes de *proximitat* i de *distància*, física i social, entre emissor i receptor (vg. pàgina següent). Com es pot observar en la figura, la major part dels discursos produïts a través d'un canal gràfic tenen les característiques pròpies de la distància i la major part dels discursos produïts a través d'un canal fònic tenen les característiques pròpies de la proximitat; però també se'n poden donar de gràfics amb trets de proximitat i de fònics amb trets de distància. El quadre sembla altament explicatiu de la gradació i el paràmetre *proximitat / distància*, en combinació amb el factor bàsic del canal (*gràfic / fònic*), queda prioritzat com a explicatiu de la diversitat de l'estil oral i l'estil escrit.

---

dins del volum de la mateixa autora *Talking voices: Repetition, Dialogue and Imagery in Conversational Discourse*, tal com especifico en la bibliografia.

<sup>32</sup> En aquesta línia, V. Lamíquiz (1996) estableix una distinció entre *texto oral* y *texto coloquial* que també té com a fonament les nocions de canal i de mode, però hi barreja altres aspectes quan oposa al col·loquial un registre amb la peculiar —i esbiaixada— denominació de *texto cuidado*.

**Das Nahë / Distanz-Kontinuum und die Konzeptionell-medialen Affinitäten**  
(El *continuum* Proximitat / Distància i l'afinitat Mode-Canal)



**PROXIMITAT**

**DISTÀNCIA**

**a) Condicions de comunicació:**

- Privacitat.
- Intimitat.
- Implicació emocional.
- Lligam amb la situació.
- Proximitat física.
- Diàleg.
- Espontaneïtat.
- Fixació del tema.

**a) Condicions de comunicació:**

- Publicitat.
- Distància afectiva.
- No implicació emocional.
- No lligam amb la situació.
- Distància física.
- Monòleg.
- Reflexió.
- Llibertat de tria del tema.

**b) Estratègies de comunicació:**

- Llenguatges no verbals (gestos, etc.)
- Escassa planificació.
- Provisionalitat.
- Agregació.

**b) Estratègies de comunicació:**

- Llenguatge verbal.
- Esforç de planificació.
- Caràcter definitiu.
- Integració.

(Koch i Oesterreicher, 1990: 12)<sup>33</sup>

<sup>33</sup> La traducció de l'alemany ha estat meua, amb l'ajuda imprescindible de Gabrielle Berkenbusch, que va ser, a més, qui em va fer conèixer l'obra d'aquests autors.

La idea de gradació o *continuum*, que d'una manera general és de gran aplicació, per a l'oposició oral / escrit demana algunes precisions. La gradació en altres eixos del registre (pe. ex., formal / informal) apareix com una successió de possibilitats, sense aparent solució de continuïtat entre els dos pols de la gradació. En la dimensió oral / escrit, en canvi, sembla que la combinació de diversos paràmetres contextuals pot deixar alguns espais buits en el *continuum*. Hi ha escrits que poden aproximar-se molt a les característiques de l'oral col·loquial, però seria estrany que s'hi identifiquessin del tot. En el teatre o en el cinema, per exemple, els diàlegs volen imitar la llengua parlada, però, com assenyalen L. Payrató (1988) i A. Narbona (1988), mai no arriben a equiparar-s'hi:

“Pel que fa als registres escrits, és evident que es poden apropar al col·loquial, però també que no s'hi identifiquen mai. El text escrit «més col·loquial» que poguéssim imaginar — un sainet d'Eduard Escalante, una «gatada» de Serafí Pitarra, etc.— no contindrà mai per exemple, les vacil·lacions pròpies de l'organització del discurs oral improvisat. Si ho fes, el públic ho interpretaria com una mostra que els actors... no recorden el paper! La diferència entre un text escrit per ser dit (com si no fos escrit) i la parla espontània és que el primer està *planificat* i el segon no. El text «falsament» oral és més compacte, més cohesiu. La parla espontània és més vacil·lant i menys sistemàtica.” (Payrató, 1988: 52)

“En consecuencia, puede afirmarse que la incorporación de lo coloquial en lo literario no se puede conseguir nunca con autenticidad. Siempre se lleva a cabo una adaptación, que implica simplificación o reducción y manipulación por parte del autor; por un lado, se prescinde del contexto del registro coloquial real; por otro, se vale de esta modalidad de uso para unos propósitos (realismo, caracterización de personajes, etc.) que en el empleo corriente no se dan.” (Narbona, 1988: 151)<sup>34</sup>

---

<sup>34</sup> Aquesta discontinuïtat en la gradació pot relacionar-se amb l'explicació, de caràcter més global, que fa R. Hughes (1996) sobre els processos de “traducció” de l'oralitat a l'escriptura i viceversa. Per a aquesta autora, el pas de l'una a l'altra no és automàtic sinó comparable amb la transmutació entre dues matèries de naturalesa diferent que requereixen un procés ben sofisticat:

“(...) the transference of language items from speech into writing and vice versa is by no means a neutral one. The tendency is to think of the process as similar to, say, the translation of a computer file from one wordprocessing system to another. This model (...) is a misleading one. A better model is one which hinges on the notion of transmutation, for example, the formation of ice cubes from water. A child soon learns the relationship between water and ice, and knows the level of identity between them; nonetheless, the substances on either side of the process of freezing are quite different, and are fit for quite different purposes. Literate people, who are the describers and analysts of language, can freeze and unfreeze language very easily, so easily in fact that we often lose sight of the sophisticated nature of the process.” (Hughes, 1996: 6)



Uns altres textos escrits que s'aproximen a la llengua oral són els esborranys, els apunts, les notes preses sobre la marxa, etc., que hi comparteixen el caràcter improvisat i això té conseqüències sobre l'estructuració i la sintaxi. De totes maneres, no tinc coneixement de cap estudi que n'hagi comprovat el grau de proximitat i caldria veure si hi són realment molt propers o si, més probablement, tenen unes característiques genèriques pròpies que els distancien d'una manera suficient tant de l'oralitat com de l'escriptura prototípica.

Per l'altre extrem del *continuum*, el discurs oral més formal i expositiu, si no és la lectura en veu alta d'un escrit, difícilment té les característiques de la llengua escrita (sempre presenta, fins i tot en els individus més destres, almenys algun anacolut, alguna vacil·lació, una construcció menys compacta lèxicament o més repetitiva, etc.). Així, semblaria que, en la figura de P. Koch i W. Oesterreicher (1990), les puntes extremes de cada triangle podrien no donar-se en la realitat. D'altra banda, la lectura en veu alta de textos escrits seria més explicable com una *superposició* o combinació dels dos modes que no com un punt determinat en un *continuum* ininterromput. De fet, parlar de naturalesa gradual o escalar no té per què implicar la continuïtat en el sentit literal del mot.

En resum, en un *continuum* rectilini o bé amb algunes interrupcions, la idea de gradació s'imposa com la interpretació més acurada de la relació entre oralitat i escriptura, i permet de superar la visió dicotòmica, excessivament simplificadora —i, per tant, enganyosa— de la realitat. És dins d'aquesta noció de gradació que s'emmarca la comparació que efectuaré entre tres gèneres discursius —especialment en el cas del gènere intermedi, la classe magistral— i la interpretació dels resultats obtinguts.

---

L. Payrató (comunicació personal) apunta que en aquest pas d'un text oral a un d'escrit el fenomen bàsic, més que una *transmutació* aparentment neutra, és la *interpretació*, amb bases convencionals, per part del subjecte transcriptor (p. ex, quan s'interpreta que una pausa és prou llarga o prou significativa per ser una coma, un punt, etc.).

### 2.2.3- ALTRES PARÀMETRES QUE INCIDEIXEN EN EL MODE DEL REGISTRE

**“The variation among texts within speech and writing is often as great as the variation across the two modes. No absolute spoken / written distinction is identified in the study.”**

D. Biber, *Variation across Speech and Writing*, 1988, pàg. 24.

El quadre de P. Kock i W. Oesterreicher (1990) presentat en el punt anterior es basa en la consideració d'un paràmetre (*proximitat / distància*) com a essencial en la gradació del mode del registre, més explicatiu que les oposicions gràfic / fònic i oral / escrit. Altres investigadors també han tendit a assenyalar un paràmetre com a més prominent<sup>35</sup>. D. Tannen, en dos articles, un dels quals es titula precisament “*The Oral / Literate Continuum in Discourse*” (1982a), postula que, al costat de la distinció oral / escrit, n'hi ha una altra de més potent:

“(...) These are some of the research areas in which I have found useful the notion of an oral / literate continuum, or, more precisely, a continuum of relative focus on interpersonal involvement vs. message content.” (Tannen, 1982a: 15)

“I have suggested that previous work on oral and literate tradition and spoken versus written language has led to two hypotheses. The first, that written language is decontextualized whereas spoken is context-bound, seems to grow out of the types of spoken and written discourse that were examined: face-to-face conversation on the one hand and expository prose on the other. I suggest that the identified differences result not so much from the spoken and written modes as from relative focus on interpersonal involvement typically found in conversation and relative lack of focus on involvement in

---

<sup>35</sup> Al marge dels autors que comentaré a continuació, també puc esmentar S.A. Thompson (1984: 91), que creu que l'oposició formal / informal resulta més explicativa que la parella oral / escrit: “... *in the area of clause combining, the formal / informal distinction, rather than the written / spoken one, may be most relevant*”. També hi ha el cas de P. Kay (1977), que destaca el paràmetre *autònom / no autònom* i el relaciona amb les nocions de *codi elaborat* i *codi restringit*, respectivament, de B. Bernstein, que comentaré més endavant (vg. § 2.3.2). P. Kay (1977: 24), en una afirmació de més abast, arriba a la conclusió que “*the direction of linguistic evolution is from a nonautonomous to an autonomous system of communication.*”

favor of a focus on information or message typically found in expository prose.” (Tannen, 1985: 143)

És a dir, per a D. Tannen, l’oral prototípic està altament implicat en la situació i, més en concret, en la relació interpersonal entre emissor i receptor, mentre que l’escrit prototípic està focalitzat en el contingut informatiu del missatge. I, com a conseqüència, un discurs oral que se centri en el contingut o un discurs escrit que es decanti cap a la funció interpersonal, donaran lloc a formes intermèdies en la gradació.

L’oposició *interpersonal / informatiu*, de fet, és generalment acceptada com una dimensió crucial de la variació dels textos, pròpia del tenor funcional del registre<sup>36</sup>. G. Brown i G. Yule (1983b) l’han tractada sota els rètols *funció interaccional* i *funció transaccional*, i també l’han considerada com un paràmetre incident en la gradació en el mode del registre<sup>37</sup>:

“We could say that primarily interactional language is primarily *listener-oriented*, whereas primarily transactional language is primarily *message-oriented*. (...) If we generalise across the uses of written language in our society (still excepting literature) we find that the fundamental function common to most uses of the written language is the transmission of information, whether recording information about what is past, or recording the intentions of writers about what is to happen in the future. We shall call this information-transferring function of language the *transactional* function of language. (...) We have said that the primary function of written language is *transactional*, to convey information (though there are relatively major exceptions, like literature, and relatively minor interactional ones like love letters). We have said that the primary function of spoken language is interactional, to establish and maintain social relations. However, an important function of spoken language is primarily transactional —to convey information. We shall schematically represent these functions in figure 1.1.

---

<sup>36</sup> Aquesta oposició coincideix amb el factor *Implicat o interactiu vs. Informatiu*, el principal obtingut per D. Biber en el seu treball de 1988 (vg. més endavant, en aquest mateix punt).

<sup>37</sup> G. Brown i G. Yule reprenen una distinció que té les arrels en els estudis de teoria de la comunicació de l’Escola de Palo Alto, en els quals s’estableix que tots els episodis comunicatius es poden arrencar en dos eixos: la interacció i la informació (L. Payrató, comunicació personal).

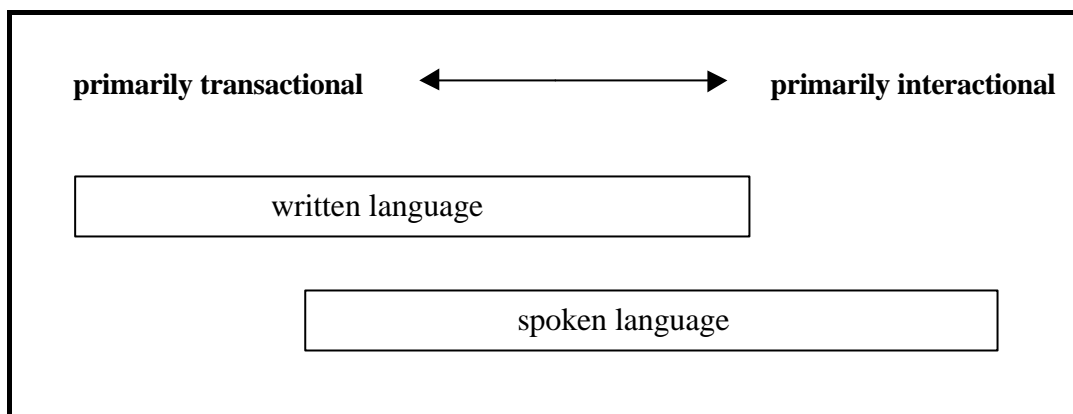


Figure 1.1 *Functions of language*” (Brown i Yule, 1983b: 11-23)

El treball d’E.O. Keenan / E. Ochs (1977 / 1979)<sup>38</sup> és també un exemple de la tendència a considerar que hi ha una altra dimensió fonamental al costat de —o per comptes de— l’oposició oralitat / escriptura. Aquesta autora, en un estudi encaminat a descriure la parla dels infants en relació amb la dels adults, arriba a la conclusió que el paràmetre més explicatiu és el *grau de planificació* del discurs i que les estratègies típiques del llenguatge infantil coincideixen amb les del discurs adult no planificat:

“(…) That is, we suggest that, when speakers have not planned the form of their discourse, they rely more heavily on morphosyntactic and discourse skills acquired in the first 3-4 years of life. The counterpart of this suggestion is that more planned language use draws on knowledge that is acquired or learned (Krashen, 1976,1977)<sup>39</sup> later in life. In the case of our own society, the suggestion is that more planned uses of language draw upon knowledge transmitted through formal education. This knowledge includes use of complex syntactic structures (e.g. complementation, cleft constructions, certain types of relative clauses, passives) and more formal discourse devices (e.g. use of textual cohesion and transitional terms such as *for example, that is, furthermore, on the other hand*, and use of topic sentences to open paragraphs in written discourse). (...) At the heart of our notion of planning is the idea of *forethought*. (...) 1. *Unplanned discourse* is discourse that lacks

<sup>38</sup> Com comentaré en el seu moment (vg. § 3.6.5), aquest escrit va aparèixer sota títols diferents i signat per autores diferents amb dos anys d’interval. D’altra banda, la referència que ha tingut més difusió és una tercera publicació, E. Ochs (1983), que consisteix en el mateix escrit, sense cap canvi en el cos del text, aparegut sis anys després de la primera edició.

<sup>39</sup> E.O. Keenan / E. Ochs (1977 / 1979) es refereix a Stephen D. Krashen (1976), “Linguistic Puberty”, comunicació presentada a la *NAFSA Conference*, San Diego, Califòrnia, 7 de maig de 1976, que no m’ha estat possible de localitzar. La referència de S. Krashen (1977), en canvi, ha estat més accessible i l’he pogut incloure en la bibliografia.

forethought and organizational preparation. 2. *Planned discourse* is discourse that has been thought out and organized (designed) prior to its expression.” (Keenan / Ochs, 1977: 4-6 / 1979: 53-55)

Complementàriament a aquestes idees, E.O. Keenan / E. Ochs (1977 / 1979) afirmen que el desenvolupament lingüístic progressa gradualment de l'ús d'estratègies pragmàtiques a la construcció de les oracions amb una sintaxi elaborada. En resum, hi ha una relació directa entre l'edat de l'individu, el grau de planificació del discurs i la forma lingüística, més simple o més complexa, que presentarà finalment aquest discurs.

La proposta d'E.O. Keenan / E. Ochs (1977 / 1979) té l'interès de destacar la importància del tret planificat / no planificat en la construcció dels discursos<sup>40</sup>. Sembla, però, que la seva aposta per la planificació com a paràmetre essencial en la diferenciació dels discursos pot ser discutible, com es veurà a continuació.

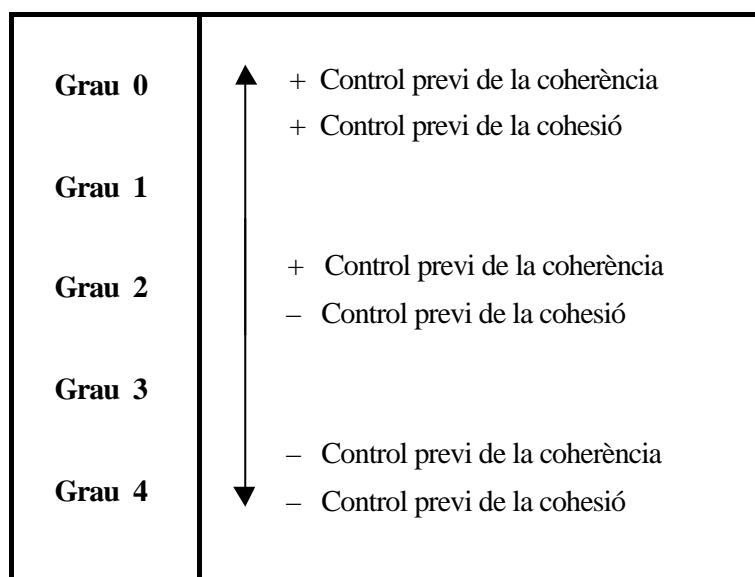
En primer lloc, el concepte de planificació, que sembla similar al d'espontaneïtat utilitzat per molts altres autors —p. ex., el quadre de M. Gregory i S. Carroll (1978) comentat anteriorment—, es presenta incòmode per a l'investigador i difícil de mesurar. Podria dependre del temps de planificació, del nombre de vegades que el parlant ha produït aquell discurs amb anterioritat, de si ha llegit sobre el tema, de si té el costum de parlar en aquell estil de llenguatge, etc. En darrer terme, com que la planificació pot ser una

<sup>40</sup> Aquest és un paràmetre que pot tenir interès especialment per a una gradació, necessàriament aproximativa, del grau d'espontaneïtat dels discursos orals. Una proposta d'aquesta gradació apareix a J.M. Castellà (1992: 128):

<b>GRAU 0</b>	Rituals.	Fórmules de salutació i comiat; ritus religiosos, acadèmics, parlamentaris, etc.
<b>GRAU 1</b>	Lectura o memorització d'un text.	Noticiaris de ràdio i TV; teatre; cinema; recitació; cançó; conte llegit; lectura d'una conferència; etc.
<b>GRAU 2</b>	Textos amb esquema previ.	Conferència; homilia; entrevista (l'entrevistador); presentador d'un concurs; tertúlia; conte explicat; etc.
<b>GRAU 3</b>	Textos improvisats de tema elaborat prèviament.	Conferència; entrevista (l'entrevistat); classe; declaracions d'un polític; intervencions a reunions professionals, juntes, assemblees; etc.
<b>GRAU 4</b>	Textos espontanis	Conversa en confiança (el desenvolupament).

operació interna de l'individu i, per tant, no observable, no hi hauria més remei que demanar-li si ha preparat poc o gaire el discurs i és presumible que l'emissor, en moltes ocasions, no pogués ni respondre-ho d'una manera clara<sup>41</sup>.

En segon lloc, el grau de planificació no sembla afectar de la mateixa manera els diversos aspectes de la producció del discurs. Així, en un discurs oral preparat sol haver-hi més planificació dels aspectes de coherència informativa global (estructura, ordenació, selecció de les idees) que no dels de cohesió lineal i construcció lexicogramatical<sup>42</sup>. De fet, el control total del lèxic i de la sintaxi implicaria la recitació de memòria —o la lectura en veu alta— de tots els enunciats del text. Coherència i cohesió es combinen amb l'espontaneïtat i donen lloc a diversos graus de control previ per part de l'emissor:



(Adaptació de Castellà, 1992: 129)<sup>43</sup>

<sup>41</sup> Un comentari en aquesta línia sobre el grau d'espontaneïtat pot trobar-se a J.M. Castellà (1996a: 263-264).

<sup>42</sup> La distinció entre els conceptes de coherència i de cohesió no és contemplada per tots els autors. Aquí l'empro tal com la defineixo a J.M. Castellà (1992: 57-58).

<sup>43</sup> Els graus que s'indiquen aquí es corresponen amb els del quadre que acabo de donar en nota a la pàgina anterior.

Quan un discurs oral ha estat produït per un individu diverses vegades, pot haver-hi una certa tendència a fixar algunes estructures, alguns exemples, etc. Però el més habitual, en la majoria dels discursos orals preparats, és que no es faci una memorització de cada mot, sinó de les idees més generals.

Així doncs, al contrari del que suggereixen E.O. Keenan / E. Ochs (1977 / 1979), no sembla que la planificació hagi de ser el factor determinant del tipus de relacions cohesives i sintàctiques emprades en un text, i que, en tot cas, si s'hi observen unes coincidències, hauran de ser causades per algun altre paràmetre, que apareix al mateix temps que la planificació i que no ha estat considerat. D. Tannen (1982b) fa també una observació crítica, força concloent, dels arguments d'E.O. Keenan / E. Ochs, tot destacant el paper de les convencions de gènere:

“Moreover, differences in features of the type which Ochs attributes to planning seem to be far more responsive to changes in genre and context than to changes in planning time. For example, if one is about to deliver a formal lecture, one indeed has time to plan. However, one may just as well plan to have a talk with a spouse or friend; yet endless planning time will not result in a discourse that sounds like a lecture. Similarly, following a formal lecture, one may be asked an unanticipated question and be required to answer on the spot, producing discourse that sounds much more like a lecture than like yesterday's planned talk at home. Some people, at least, are able to produce prose that sounds either academic and 'planned', or informal and spontaneous, in the time it takes to pass a pen over paper or hit typewriter keys. What seems to determine their choice of words and structures is their sense of what is appropriate to the context.” (Tannen, 1982b: 6)

Efectivament, com ja he assenyalat anteriorment, cal tenir en compte, al costat del registre, els trets associats amb el gènere i el tipus de text, és a dir, amb les convencions sobre les formes arquetípiques que solen adoptar els discursos en situacions socials —o per a finalitats socials— també arquetípiques. D. Tannen, a partir de l'anàlisi de narracions orals i escrites, insisteix en la importància del factor *gènere*:

“Comparative analysis of spoken and written versions of a narrative demonstrates (1) that features which have been identified as characterizing oral discourse are also found in written discourse, and (2) that the written short story combines syntactic complexity expected in writing with features which create involvement expected in speaking. (...) The

difference between features of language which distinguish discourse types reflects not only —and not mainly— spoken vs. written discourse mode, but rather genre and related register, growing out of communicative goals and context.” (Tannen, 1982b: 1-18)

Per tant, D. Tannen pensa que l’aparició de característiques considerades prototípiques del llenguatge oral en l’escrit i viceversa, pot explicar-se pels objectius concrets dels parlants quan elaboren els seus discursos seguint les convencions existents sobre cada gènere.

Encara en aquesta línia que destaca altres paràmetres de la variació dels textos, F.N. Akinnaso (1985)<sup>44</sup> no en proposa un de concret, sinó una perspectiva diversa amb múltiples factors influent alhora:

“(…) the conventional method of pairing spoken and written data for comparative analysis cannot account for all dimensions of relationship between spoken and written discourse types (...). Rather, the points of intersection must be located here and there: the meeting place is sometimes in method of acquisition, sometimes in discourse form or genre, sometimes in degree of formality, sometimes in functional load, sometimes in communicative goal, sometimes in the manipulation of particular aspects of language structure.” (Akinnaso, 1985: 352)<sup>45</sup>

---

<sup>44</sup> F.N. Akinnaso juga amb els títols de dos articles que va publicar amb tres anys de diferència en la mateixa revista, *Language and Speech*. El primer, de 1982, es titula “On the Differences between Spoken and Written Language” i el segon, de 1985, “On the Similarities between Spoken and Written Language”. El primer se centra en el contrast entre oralitat i escriptura i el segon, en canvi, destaca alguns aspectes que, segons l’autor, han resultat ser menys diferents del que semblava.

<sup>45</sup> F.N. Akinnaso (1985) assenyala alguns aspectes inadvertits de les societats no alfabetitzades que poden representar una crítica o, almenys, una matisació a algunes de les afirmacions de W. Ong i J. Goody presentades en el punt 2.1.1. Aquest autor —que escriu sobre aquests temes des d’una posició privilegiada, ja que firma el primer dels articles des de la *University of California*, Berkeley, i el segon des de la *Lagos State University*, Nígeria— destaca unes similituds entre els dos tipus de societat i acaba insistint en el paper relatiu de la diferenciació de mode en lingüística:

“As indicated earlier, one reason for the belief in differences between spoken and written language is the assumption that literacy must have unparalleled linguistic consequences. But by taking the written «academic language» as normative (...) scholars have not only ignored the various forms and functions of writing, they have also virtually neglected the nature of the oral and, more importantly, the complex oral resources used in nonliterate societies to perform the functions associated with literacy and written language in literate societies. (...) Yet, it is clear (...) that the «elevated» variety in both types of society must have similar consequences for their users, in addition to sharing similarities of form and function. (...) It should be pointed out at this juncture that many of the observations made about oral ritual communication in nonliterate societies also apply to literacy-motivated oral ritual communication in literate societies. For, in many respects,



Com a darrera aportació per destacar en aquest punt, cal citar la de D. Biber (1988) que fa una anàlisi de gran abast, amb un nombre important de textos i de trets lingüístics i amb la utilització de suport informàtic i de tècniques quantitatives sofisticades (vg. § 3.4.1). Aquest investigador, de manera coincident amb altres ja citats, conclou que l'oposició oral / escrit no és l'únic factor, potser ni tan sols el factor fonamental, en la variació discursiva, ja que les relacions entre els textos no poden ser definides unidimensionalment i considera també que cada gènere pot tenir un espai propi en el *continuum* oral / escrit:

“The present study is based on the assumption that there are few, if any, absolute differences between speech and writing, and that there is no single parameter of linguistic variation that distinguishes among spoken and written genres. (...) For each genre we need to know the particular ways in which it is oral and the ways in which it is literate. Thus the present study attempts to identify the linguistic parameters along which genres vary, so that any individual genre can be located within an ‘oral’ and ‘literate’ space. (...) This analysis shows that there is no single, absolute difference between speech and writing in English; rather there are several dimensions of variation, and particular types of speech and writing are more or less similar with respect to each dimension.” (Biber, 1988: 55-199)

D. Biber troba, a partir d'agrupacions estadístiques de trets lèxics i gramaticals, sis dimensions que, segons els resultats de la seva recerca, constitueixen els eixos bàsics de la variació dels discursos en anglès. Aquestes dimensions, que comentaré detingudament més endavant, són les següents:

- “1- Involved versus Informational Production
- 2- Narrative versus Non-Narrative Concerns
- 3- Explicit versus Situation-Dependent Reference
- 4- Overt Expression of Persuasion
- 5- Abstract versus Non-Abstract Information

---

formal lectures, conference presentations, job interviews, Ph. D. orals, etc., are ritual encounters, requiring specialized (para)linguistic skills as well as formal training. What this implies is that formality, function, audience, and communicative goal may be more crucial to the determination of linguistic structures than modality.” (Akinaso, 1985: 351-352)

Cadascuna d’aquestes dimensions es relaciona amb uns grups de trets lingüístics i, alhora, amb uns gèneres i amb uns trets contextuals<sup>47</sup>. Les sis dimensions de D. Biber (1988) són els paràmetres que ell considera explicatius de la variació lingüística, per davant de l’oposició oral / escrit del mode del registre<sup>48</sup>.

---

<sup>46</sup> D. Biber (1995a i 1995b) —i, poc abans, D. Biber i M. Hared (1994) i Y.-J. Kim i D. Biber (1994)— compara diverses llengües, en un estudi multidimensional similar al de 1988 sobre l’anglès. Les llengües triades són de famílies lingüístiques ben distants entre si: coreà, somalí i *nukulaelae tuvaluan* (Oceania). D. Biber troba una notable similitud en els models de variació lingüística d’aquestes tres llengües entre si i de les tres amb l’anglès, especialment en els registres expositius. A partir d’aquests resultats, D. Biber (1995a) planteja, si bé d’una manera prudent, la possibilitat de l’existència d’uns universals de la variació lingüística:

“Given the strength of these patterns, it is reasonable to hold out the possibility of cross-linguistic universals governing the patterns of discourse variation across registers and text types. However, considerable additional research is required to determine which of the cross-linguistic generalizations identified in the present study are valid as universals, and to formulate more precise statements of those relationships.” (Biber, 1995a: 359)

<sup>47</sup> A partir d’aquesta investigació, D. Biber (1989) elabora una tipologia de textos de l’anglès contemporani. Aquesta tipologia, que he comentat i comparat amb les d’altres autors a J.M. Castellà (1995 i 1996b), és la següent:

- 1- Intimate interpersonal interaction
- 2- Informational interaction
- 3- ‘Scientific’ exposition
- 4- Learned exposition
- 5- Imaginative narrative
- 6- General narrative exposition
- 7- Situated reportage
- 8- Involved persuasion” (Biber, 1989: 20-22)

La tipologia de D. Biber (1989) té l’interès d’haver estat obtinguda per procediments estadístics, a diferència de la majoria de les tipologies elaborades en l’àmbit de la lingüística del text, que han estat concebudes d’una manera teòrica i aplicades posteriorment als textos. De totes maneres, sorprèn agradablement la coincidència notable entre aquesta tipologia empírica i moltes de les tipologies teòriques elaborades amb anterioritat.

<sup>48</sup> En un estudi posterior, L.W. Smith (1994) argumenta al contrari de D. Biber i afirma que, malgrat la innegable intervenció d’altres factors, les diferències entre llengua oral i llengua escrita no queden pas totalment diluïdes:

“This study clearly indicates that when topic and participants were held constant and task-audience was structured to vary there were measurable differences between three spoken and written language domains at both the micro-structural level of analysis and at the macro-discourse level of analysis. On the one hand, although there are not simple, universal, exclusive differences between spoken and written language use, on the other hand, there are discernable differences between uses of spoken and written language ranging along a continuum based on genre and context considerations.” (Smith, 1994: 74)

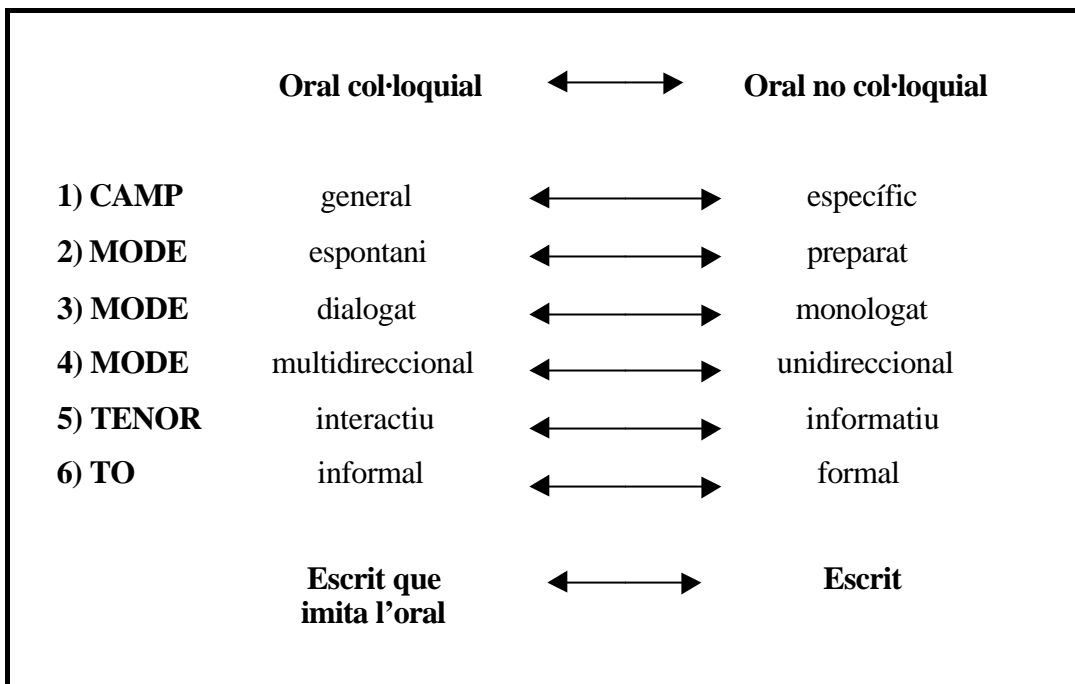
En resum, d'acord amb les diverses aportacions teòriques i experimentals analitzades al llarg d'aquest apartat 2.2, l'oposició llenguatge oral / llenguatge escrit es mostra com una relació en el mode del registre, no dicotòmica sinó gradual, que resulta de l'acció de diversos paràmetres contextuais: a) el canal; b) els altres paràmetres relatius a les condicions de producció i de recepció del discurs; c) els paràmetres corresponents a les altres dimensions del registre —camp, formalitat i to—; i d) les convencions en matèria de gènere i tipus de text.

La combinatòria dels paràmetres esmentats dóna lloc a una multitud de gèneres intermedis entre l'oral prototípic (conversa informal) i l'escrit prototípic (prosa expositiva). Així, hi ha discursos fònics que incorporen trets típics de la llengua escrita i discursos sobre suport gràfic que presenten característiques típiques de la llengua oral. No sembla que hi hagi un únic tret que pugui donar compte globalment de la distinció oral / escrit —ni, en general, de la variació funcional en el llenguatge. Més aviat hi ha un conjunt de factors que actuen al mateix temps i en direccions diverses, i que plegats configuren el que anomenem *mode oral* i *mode escrit*. Una representació gràfica d'aquesta combinació podria ser la que s'indica en el quadre adjunt (vg. pàgina següent).

D'altra banda, l'argumentació que he desenvolupat sobre la multiplicitat de factors contextuais que afecten la relació entre oralitat i escriptura, i sobre la importància dels gèneres com a conjunts de convencions que condicionen la forma final dels textos, confirma la necessitat, que també es posarà de relleu en el capítol metodològic (vg. § 3.4.3), de tractar els textos de la mostra preferentment com a exponents dels gèneres discursius dels quals formen part i no tant com a representants del conjunt de textos orals i escrits de la llengua —només a partir de la comparació entre la classe magistral i la prosa acadèmica es podran cercar possibles tendències generals en el contrast acústic / gràfic.

---

Aquesta qüestió de l'especificitat de la distinció oral / escrit en el conjunt dels altres factors del registre la podré tractar a partir de la comparació entre els gèneres classe magistral i prosa acadèmica gràcies al control de la resta de variables (vg. § 3.1.3, objectiu específic núm. 6).



(Castellà, 1996a: 225)

## **2.3- LA COMPLEXITAT LINGÜÍSTICA EN LA LENGUA ORAL I EN LA LENGUA ESCRITA**

### **2.3.1- LA COMPLEXITAT ESTRUCTURAL DE LA LENGUA ESCRITA (I): W. CHAFE**

**“I have tried to show that spoken and written language differ with regard to two sets of features. The features of one set can be assigned to an opposition of fragmentation and integration, which I suggested is a consequence of differences in the use of time in speaking and writing. The features of the other set reflect an involvement vs. detachment dichotomy, attributable to the different relations of a speaker or writer to the audience.”**

W. Chafe, “Integration and Involvement in Speaking, Writing, and Oral Literature”, 1982, pàg. 52.

En la majoria d'estudis que s'han realitzat sobre les diferències gramaticals entre la llengua oral i l'escripta es parteix d'una assumpció, a vegades verificada empíricament i a vegades simplement acceptada sense discussió, sobre el grau de complexitat lingüística d'aquests dos modes. L'assumpció consisteix a considerar que la llengua oral és més simple gramaticalment i que l'escripta és més complexa. Com que la simplicitat gramatical se sol associar amb la parataxi i la complexitat amb la hipotaxi, és habitual trobar ressenyat, tant en les obres de tipus general com en la bibliografia especialitzada, que la llengua oral conté més juxtaposició i més coordinació i que l'escripta conté més subordinació.

La simplicitat sintàctica i el caràcter fragmentari de la llengua oral són els conceptes centrals del model de W. Chafe (1982), un dels més citats i considerat un clàssic en

aquest camp de recerca. W. Chafe duu a terme un treball empíric a partir de discursos orals informals i de discursos escrits formals<sup>49</sup>, i agrupa els trets gramaticals que es mostren més prominents en cada mena de discurs. Els resultats li permeten d'associar a cada mode unes propietats generals: per a la llengua oral, la *fragmentació* i la *implicació* amb la situació, per a la llengua escrita, la *integració* i la *no-implicació* amb la situació. L'esquema de la pàgina següent representa la relació que mantenen entre si aquestes propietats i el quadre que hi ha a continuació especifica els trets lingüístics en què, segons W. Chafe, es concreta cadascuna<sup>50</sup>.

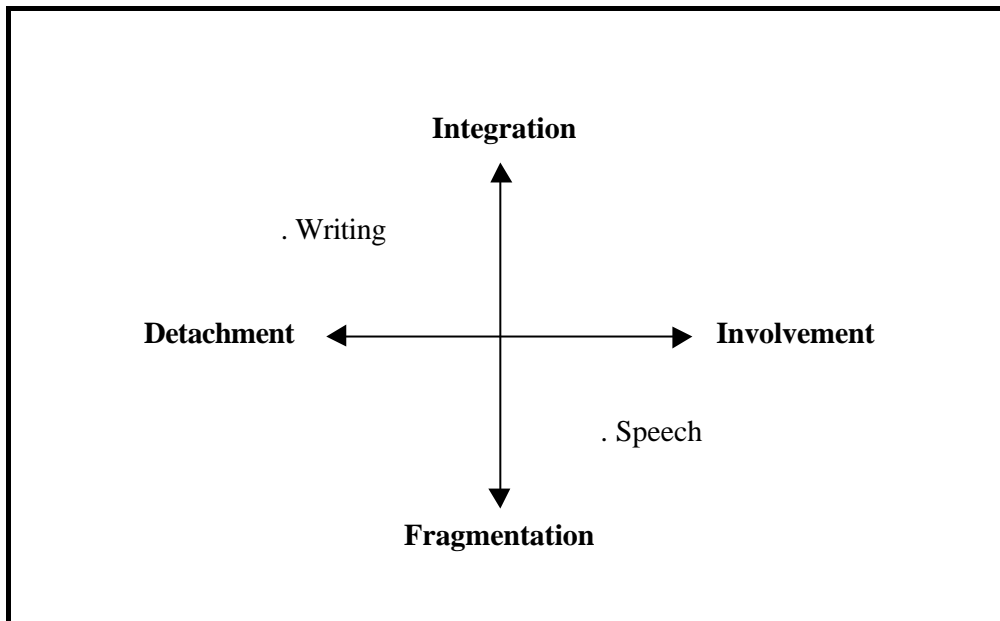
W. Chafe considera que l'oposició *fragmentació* / *integració* és conseqüència de les diferències en el temps de què disposen emissor i receptor en cada canal: la producció de la llengua oral és molt ràpida i un parlant no té temps d'integrar les idees d'una manera tan compacta com l'escriptor en un escrit. Alhora, l'oposició *implicació* / *no-implicació* és atribuïda a les diferències en la relació entre l'emissor i l'auditori: el parlant està implicat amb els receptors i hi comparteix un mateix context situacional, mentre que l'escriptor hi manté una relació més —o totalment— distanciada.

D'acord amb això, W. Chafe troba més conjuncions coordinades en el discurs oral i més nominalitzacions, més adjectius, més sintagmes preposicionals i més clàusules subordinades (completives i de relatiu) en l'escrit. Pel que fa al caràcter gradual de l'oposició oral / escrit, W. Chafe suggereix, com es pot veure en la il·lustració anterior, que les mostres d'escrit prototípic quedarien situades en el quadrant superior esquerre i les d'oral prototípic en l'inferior dret, mentre que els casos no prototípics se situarien en els punts intermedis, en algun dels altres dos quadrants.

---

<sup>49</sup> W. Chafe (1982: 36) treballa amb discursos procedents de 14 subjectes, amb 9911 mots de llenguatge oral informal (converses de sobretaula) i 12368 mots de llenguatge escrit formal (escrits acadèmics).

<sup>50</sup> Cito els quadres publicats per K. Beaman (1984: 48-49) perquè són un excel·lent resum del treball de Chafe (1982).



<b>Integration</b>	<b>Involvement</b>
<i>Nominalizations</i> <i>Past / present participles</i> <i>Attributive adjectives</i> <i>Conjoined phrases</i> <i>Series (of noun / verb phrases)</i> <i>Sequences of prepositional phrases</i> <i>Complement clauses</i> <i>Relative clauses</i> <i>Introductory phrases (e.g. 'In view of this')</i> <i>Adverbial conjunctions</i>	<i>First person references</i> <i>Speakers' mental processes (e.g. 'I think')</i> <i>Monitoring information flow</i> <i>Emphatic particles (e.g. 'just', 'really')</i> <i>Fuzziness, vagueness, hedges</i> <i>Direct quotes</i>
<b>Detachment</b>	<b>Fragmentation</b>
<i>Passives</i> <i>Nonactive subjects</i> <i>Precision</i> <i>Indirect quotations</i>	<i>Coordinating conjunctions</i> <b>Afterthoughts</b> <i>Repetitions</i> <i>False starts</i>

(Beaman, 1984: 48-49)

## **2.3.2- LA COMPLEXITAT ESTRUCTURAL DE LA LLENGUA ESCRITA (II): ALTRES AUTORS**

**“Prosody in spoken discourse establishes cohesion in much the same way as complex syntax and connectives do in written discourse.”**

J.J. Gumperz, H. Kaltman i M.C. O'Connor, 1984, “Cohesion in Spoken and Written Discourse: Ethnic Style and the Transition to Literacy”, pàg. 10.

Com que el fet de considerar la llengua escrita més complexa sintàcticament és l'assumpció majoritària, es poden trobar molts autors, a més de W. Chafe, que mantenen aquest punt de vista. Destaco, en primer lloc, tres dels autors citats en punts anteriors:

**1– R. Horowitz i S. Samuels (1987)**, en la seva llista general de característiques, assumeixen que l'estil oral es basa en “*simple linear structures*”, mentre que l'escrit conté “*complex hierarchical structures*” (vg. § 2.2.1).

**2– E.O. Keenan / E. Ochs (1977 / 1979)**, en l'argumentació a favor del tret planificat / no planificat, indiquen que els discursos planificats contenen “*complex syntactic structures*” (vg. § 2.2.3).

**3– D. Tannen (1982b)**, en la seva anàlisi de discursos narratius orals i escrits, assenyala que “*the written story combines syntactic complexity expected in writing with features which create involvement expected in speaking*” (vg. § 2.2.3).

En cap d'aquests tres casos la finalitat de la recerca no era estudiar la complexitat sintàctica i, per tant, la idea que expressen sobre aquest tema té el caràcter d'assumpció indiscutida.



A continuació presento, per ordre cronològic de publicació, les argumentacions d'altres autors del meu recull bibliogràfic que, des de camps de recerca diferents, mantenen el mateix punt de vista. Es tracta dels treballs d'A.M. Badia i Margarit (1952), B. Bernstein (1962a, 1962b i 1965), Real Academia Española de la Lengua (1973), M. Seco (1974), A.R. Luria (1975), B. Kroll (1977), F.N. Akinnaso (1982), G. Brown i G. Yule (1983a i 1983b), K. Perera (1984 i 1986), J.J. Gumperz, H. Kaltman i M.C. O'Connor (1984) i H. Calsamiglia i A. Tusón (1999):

**4– A.M. Badia i Margarit (1952)**, des d'una perspectiva diacrònica, relaciona la parataxi amb la llengua parlada, rural i inculta, i amb la fase inicial del català i, en general, de les llengües romàniques. L'argumentació s'enriqueix amb una citació del prestigiós W. von Wartburg:

“És general, en efecte, el predomini de la parataxi (o juxtaposició gramatical) sobre la hipotaxi (o subordinació gramatical) en la fase primitiva de les llengües romàniques, en la modalitat parlada d'abans i d'ara, i en qualsevol manifestació lingüística dins d'un ambient popular. Tots hem pogut observar, ben sovint, que el llenguatge popular —dels medis rurals o de gent de poca cultura— es caracteritza per la manca o escassetesa de domini gramatical d'oració a oració, i, d'altra banda, per la juxtaposició, a voltes monòtonament reiterativa, de les frases successives. Veus ací què ha escrit el romanista Walther von Wartburg sobre aquest tipus de construcció sintàctica, al·ludint concretament a la llengua francesa: «La parataxi és, en primer lloc, característica de totes les maneres de parlar populars, lliures, espontànies. La preferència de la llengua popular per una seriació deslligada i poc consistent, correspon a la minvada disposició que el poble pugui mostrar per les concatenacions lògiques i per a l'engranatge de llargues sèries de pensaments (...).»” (Badia i Margarit, 1952: 15)<sup>51</sup>

**5– B. Bernstein (1962a, 1962b i 1965)** fa la seva famosa definició dels codis *elaborat* i *restringit* de la llengua i els relaciona amb les classes socials. B. Bernstein aplica

---

<sup>51</sup> La referència bibliogràfica de W. von Wartburg que dona el mateix A.M. Badia i Margarit és la versió espanyola de D. Alonso i E. Lorenzo, de 1951, els quals, malauradament, no esmenten en la traducció ni l'any d'edició ni el títol original alemany de l'obra de W. von Wartburg.

ambdues nocions a mostres de llengua oral, però els seus resultats són fàcilment extrapolables al contrast oral / escrit<sup>52</sup>:

“Two general types of code can be distinguished: *elaborated* and *restricted*. They can be defined, on a linguistic level, in terms of the probability of predicting for any one speaker which syntactic elements will be used to organize meaning. In the case of an elaborated code, the speaker will select from a relatively extensive range of alternatives and therefore the probability of predicting the pattern of organizing elements is considerably reduced. In the case of a restricted code the number of these alternatives is often severely limited and the probability of predicting the pattern is greatly increased.” (Bernstein, 1962a: 76-77)

“A restricted code can arise *at any point* in society where its conditions may be fulfilled but a special case of this code will be that in which the speaker is *limited* to this code. This is the situation of members of the lower working class, including rural groups. An elaborated code is part of the life chance of members of the middle class; a middle-class individual has access to the two codes, a lower working-class individual access to one. It follows from this formulation that orientation towards the use of these codes is independent of measured intelligence and is a function of the form social relationships take. (...) Middle-class groups used a high proportion of the following: Subordinations; Complex verbal stems; Passive voice; Total adjectives; Uncommon adjectives; Uncommon adverbs; Uncommon conjunctions; Egocentric sequences; ‘of’ as a proportion of the sum of the prepositions ‘of’, ‘in’ and ‘into’ (...)” (Bernstein, 1962b: 109-115)<sup>53</sup>

**6– El Esbozo de una nueva gramática de la lengua española (Real Academia Española de la Lengua, 1973)**, dins del punt 3.17.4, titulat “Diferencias entre la coordinación y la subordinación”, fa unes afirmacions en la línia de les d’A.M. Badia i Margarit:

“Desde hace tiempo, la Lingüística ha demostrado en firme que la unión asindética, la paratáctica y la hipotáctica, son fases distintas de un mismo proceso histórico. Ya hemos dicho antes que en la época prerrománica desaparecen casi todas las conjunciones latinas, porque no eran necesarias para la expresión en aquellos siglos de baja cultura; y las

---

<sup>52</sup> L. Cortés (1986a: 186-204) arriba a unes conclusions similars a les de B. Bernstein: en el seu treball sobre la llengua castellana troba més subordinació en els parlants del grup de més nivell sociocultural.

<sup>53</sup> Aquesta llista de trets gramaticals és altament coincident amb les que obtenen anys després W. Chafe i molts altres investigadors en l’estudi del contrast llengua oral / llengua escrita. Una part d’aquests trets els analitzaré amb detall en els capítols 4 i 5.

lenguas romances van creando un nuevo sistema conjuntivo, a medida que adquieren seguridad en sí mismas y necesitan enriquecer sus medios de enlace oracional. No se trata tanto de un proceso diacrónico en la evolución del lenguaje humano, como un hecho observable en la sincronía de todas las lenguas de cultura, donde conviven los tres procedimientos de enlace en proporciones variables según el grado de instrucción literaria de las diferentes clases sociales: los niños y las hablas vulgar y rústica usan muy pocas conjunciones en comparación con la riqueza expresiva del habla culta y literaria.” (Real Academia Española de la Lengua, 1973: 502)<sup>54</sup>

7– **M. Seco (1974)**, en un fragment molt citat dins de l'àmbit hispanòfon, parla de la “*simplicidad*” i de la falta d'elements de connexió en la llengua oral:

“Al mismo impulso a que obedece la oración sincopada se debe la frecuente simplicidad en el encadenamiento de oraciones, donde la falta de elementos de conexión (tan alejada, en su sentido, del asíndeton literario) acentúa el relieve de los enunciados parciales que se suceden. (...) Los elementos de la frase tienden a flotar separados unos de otros, ajenos a

---

<sup>54</sup> La idea de “pobresa lingüística”, que ja hem vist a A.M. Badia i Margarit (1952) associada a unes èpoques, a unes classes socials i a unes modalitats lingüístiques determinades (les orals, les dialectals, etc.) ha condicionat tota una visió del llenguatge que ha imperat durant segles i que encara avui dia és ben present en la ideologia social majoritària. Aquesta concepció ha dominat també entre els filòlegs, i alguns dels autors que investiguen l'oralitat avui dia encara n'estan impregnats. Vegeu, per exemple, les paraules següents d'A.M. Vigara (1980):

“La permanencia en la capital confiere a los hablantes un espíritu lingüístico no siempre coincidente. Los madrileños tienen fama de pecar contra el idioma. No creemos que los madrileños pequen más que los murcianos, pongamos por caso; (...). El tan criticado «o sea», que ya en los años sesenta señalaba Zamora Vicente echando raíces en Madrid, puede escucharse hoy como tic lingüístico en hablantes españoles de cualquier procedencia; a veces, incluso reforzado en un largo «o sea, bueno, vamos a ver...». (...) No podemos ignorar, aunque lo eludamos, que, generalmente, la pobreza de palabras conjuntivas es mayor en las gentes de escasa instrucción (y en los niños); (...) que el lenguaje bajo tiende a rebajar o degradar la expresión (...).” (Vigara, 1980: 33)

A la vista d'aquestes afirmacions, no ha d'estranyar que A.M. Vigara (1980) utilitzi exclusivament fragments de diàlegs literaris en una obra que titula *Aspectos del español hablado. Aportaciones al estudio del español coloquial*. Un altre especialista en la llengua oral que s'expressa d'aquesta manera és L. Cortés (1990b), que parla d'«anomalías» i d'«error» en referir-se als usos orals del relatiu *que* per comptes de *cuyo*:

“Al hablar de anomalías, por tanto, sólo haremos referencia a aquellos casos en que *cuyo* fue sustituido por *que*; las siete formas señaladas en el grafico nº 4 responden a dicho error.” (Cortés, 1990b: 188, publicat de nou sense cap canvi el 1992.)

Curiosament, aquest enfocament no invalida moltes de les observacions concretes, de gran interès, que fan aquests autors en les seves obres, sinó que dóna la mesura de la concepció del llenguatge oral que a vegades tenen fins i tot els qui es dediquen a estudiar-lo.

una estructura orgánica, liberados de un centro magnético que los engarce en una oración unitaria. (...) Es frecuente que la estructura del mensaje sea ceñido ropaje de los latidos del pensamiento (o del pensamiento-sentimiento), brotando las frases en chorros cortados, desiguales y que rebasan una y otra vez los estrechos cauces sintácticos regulares.” (Seco, 1974: 128)

**8- A.R. Luria (1975: 86-94)**, des de la perspectiva de la psicología evolutiva, estableix quatre tipus de llenguatge, el més elaborat i complet gramaticalment dels quals és el llenguatge escrit. Aquesta tipologia, que presento resumida a continuació, té l’interès de contemplar el *llenguatge monològic parlat* com a forma intermèdia entre l’oral i l’escrit:

**1- Lenguatge afectiu oral:** consta només d’exclamacions, interjeccions i tòpics discursius, elements que “*sólo convencionalmente pueden llamarse lenguaje.*” (Luria, 1975: 86).

**2- Lenguatge dialogal oral:** és fragmentari, abreviat, el·líptic i contextual. La plenitud gramatical hi és optativa. “*Puede ser incompleto, abreviado (...), pero sin embargo no deja de ser comprensible. (...) Dicho lenguaje transmite en lo esencial una idea (o sea, un significado individual correspondiente a la situación) que a menudo es incomprendible fuera del contexto.*” (Luria, 1985: 87-88).

**3- Lenguatge monològic parlat:** té una formulació discursiva més completa, tot i que també admet incompletitud. “*No presupone que el interlocutor al que va dirigido conozca forzosamente la situación, y por eso ha de contener en sí una formulación discursiva bastante completa de las informaciones que transmite. (...) En su lenguaje monológico hablado, el hombre continúa empleando los gestos y la mímica; (...) Esto da la posibilidad de conservar también en el lenguaje monológico hablado cierta insuficiencia gramatical, hace admisibles algunas elipsis (omisión de ciertos componentes gramaticales superfluos), permite utilizar construcciones gramaticales insuficientemente circunstanciadas, cuya reducción se compensa con la entonación y los gestos de que se acompañan.*” (Luria, 1975: 89-90)

**4- Lenguatge monològic escrit:** No recolza en mitjans extradiscursius, és gramaticalment complet i és l’instrument bàsic del pensament racional. “*Una*

*peculiaridad esencial del lenguaje monológico escrito es que carece de la posibilidad de apoyarse en cualesquiera medios extradiscursivos (...). El lenguaje escrito se configura como forma especial independiente de la actividad discursiva, que requiere una preparación máxima y el más pleno uso de los códigos lógico-gramaticales del idioma. (...) El lenguaje circunstanciado no es sólo un medio de comunicación sino también un instrumento del pensamiento.” (Luria, 1975: 92-93)*

A.R. Luria defensa, doncs, com també veurem que fan altres autors, que els gestos, la mímica, la veu, la prosòdia, etc. justifiquen i compensen la incompletitud gramatical de la llengua oral.

**9– B. Kroll (1977)**, en un estudi sobre la coordinació i la subordinació en el discurs oral i escrit, considera que els comunicadors “no sofisticats” prescindeixen d’aquests dos recursos connectius i, tot referint-se a E.O. Keenan / E. Ochs, considera que la codificació dels enunciats sota la pressió del temps és el factor responsable de la menor presència de subordinades i coordinades en els textos no planificats:

“Following from what Keenan has said about planned and unplanned discourse (this volume), we would expect that communication which is planned and allows time for encoding information in more «difficult» structures will exhibit a greater degree of combined ideas than communication which is spontaneous and encoded under pressure of time, which does not allow the communicator to use those combining strategies which require major manipulations of word order and sentence structure.” (Kroll, 1977: 105-106)

**10– F.N. Akinnaso (1982)** fa una revisió bibliogràfica de l’estudi del contrast entre oralitat i escriptura i elabora una llista especialment completa dels trets lingüístics de la llengua escrita —i, per defecte, de la llengua oral. F.N. Akinnaso comença afirmant-hi la superior elaboració sintàctica i semàntica de la llengua escrita<sup>55</sup>:

---

<sup>55</sup> Aquesta llista té a favor seu el mèrit d’haver estat adoptada íntegrament —i exclusivament, sense citar-ne cap altra— per l’antropòleg J. Goody (1987) a l’hora de caracteritzar sintàcticament i pragmàticament la relació entre llengua oral i llengua escrita:

“In addition there are a number of syntactic and pragmatic features which differentiate the two registers and I replace my earlier list (...) by the much more extensive inventory given by Akinnaso (1982: 104).” (Goody, 1987: 263)

“Though there are notable contradictions in the findings, it is generally argued that, unlike speech, written language promotes the following:

1. Preferential usage of elaborate syntactic and semantic structures, especially complex nominal constructions (noun groups, noun phrases, nominalizations, relative clauses, etc.) and complex verb structures (...).
2. Preference for subordinate rather than coordinate constructions (...).
3. Preferential usage of subject-predicate constructions instead of reference-proposition (...).
- 4- Preferential usage of declaratives and subjunctives rather than imperatives, interrogatives, and exclamations (...).
- 5- Preferential usage of passive rather than active verb voice (...).
- 6- Preferential usage of definite articles rather than demonstrative modifiers and deictic terms (...).
- 7- Higher frequency of certain grammatical features, e.g., gerunds, participles, attributive adjectives, modal and perfective auxiliaries, etc. (...).
- 8- The need to produce complete information or idea units and make all assumptions explicit (...).
- 9- Reliance on a more deliberate method of organizing ideas, using such expository concepts as «thesis», «topic sentence», and «supporting evidence» (...).
- 10- Preferential elimination of false starts, repetitions, digressions, and other redundancies which characterize informal spontaneous speech (...).” (Akinnaso, 1982: 104)

F.N. Akinnaso forneix, doncs, una llista de trets ben completa, alguns dels quals —preferència de l’escrit per la construcció nominal, per la subordinació, pels gerundis, pels adjectius atributius, etc.— tractaré amb detall en aquesta recerca.

**11– G. Brown i G. Yule**, des de l’àmbit de l’anàlisi del discurs (1983a) i de la didàctica de l’ensenyament de llengües estrangeres (1983b), observen també una simplicitat sintàctica i un nombre menor d’oracions subordinades en la llengua oral:

“El efecto general es la producción de una lengua hablada con una organización menos rica<sup>56</sup> que la de la lengua escrita, y que contiene menor densidad informativa, pero mayor cantidad de «marcadores» de interacción y «articuladores discursivos». (...) La sintaxis de

---

<sup>56</sup> És destacable la insistència en aquesta valoració pejorativa —global, no només sintàctica o lèxica— de la llengua oral com a “*menos rica*”, en dos analistes del discurs del prestigi i la subtilitat de G. Brown i G. Yule. En canvi, just a continuació, fan notar que la llengua oral té més marcadors i articuladors discursius.

la lengua hablada suele estar mucho menos estructurada que la de la lengua escrita: i. La lengua hablada contiene muchas oraciones incompletas, a menudo simplemente secuencias de sintagmas, ii. La lengua hablada suele contener menos subordinación.” (Brown i Yule, 1983a: 34-35)

“Most spoken language consists of paratactic (unsubordinated) phrases which are marked as related to each other, not so much by the syntax as by the way the speaker says them. The speaker uses the resources of pausing and rhythm and, to a lesser extent, intonation, to mark out for the listener which parts of his speech need to be co-interpreted. Of course, what syntax there is will contribute to this structuring, but is frequently the case that the syntax is rather simple. (...) It seems to be the case that rather limited syntax is required for adequate performance in producing the spoken language. As we have suggested, simple noun phrases and a very few subordinate declarative structures, together with an interrogative structure to ask questions with, appear to characterise typical spoken language produced by native speakers. Similarly, a great deal of the vocabulary which is produced is of a very general, non-specific sort: *chap, guy, individual, one, other* (one), *place, thing, be, have, got, do, fine, good, bad.*” (Brown i Yule, 1983b: 4-9)<sup>57</sup>

**12– K. Perera (1984 i 1986)**, com W. Chafe, A.R. Luria i B. Kroll, atribueix a la producció immediata en temps real (*on-line*) de la llengua oral, el fet que sigui menys compacta i menys acabada sintàcticament. El discurs oral és produït per l'emissor al mateix temps que és rebut pel receptor, de manera que l'emissor no té temps d'elaborar-lo perquè tingui més densitat lèxica i més organització sintàctica i, en conseqüència, és un discurs més redundat. Alhora, el receptor també ha d'interpretar el discurs en temps real, a mesura que l'emissor el va produint, i per això, si tingués més densitat, tampoc no tindria temps de processar-lo. És a dir, es dona una coincidència en les necessitats de l'emissor i del receptor, derivades de la naturalesa física del canal auditiu: a ambdós els cal un grau de redundància superior al de la llengua escrita.

El discurs escrit, al seu torn, pel fet de poder ser redactat amb menys restriccions de temps, permet més elaboració, amb més densitat lèxica i amb una estructuració sintàctica més acabada i integrada, basada en els procediments jeràrquics de la subordinació:

---

<sup>57</sup> Al final d'aquest fragment G. Brown i G. Yule destaquen un aspecte que també estudiaré en aquesta recerca: es tracta dels mots de freqüència alta i de significat molt general com ara *cosa, lloc, persona*, etc., que aporten menys pes a la densitat informativa que els mots menys freqüents i de significat més específic.

“It is generally the case that information units of speech are most often linked by devices of coordination, while clauses in writing are frequently joined by the hierarchical processes of subordination, which gives a more tightly integrated texture to the language.” (Perera, 1984: 183)

D’acord amb aquesta idea general, K. Perera (1984) presenta una llista dels trets lingüístics de l’oral i de l’escrit, organitzats seguint el fil de la seva argumentació:

### **1. Differences related to physical production**

- 1.1. *High redundancy in speech*: false starts, incomplete utterances, hesitations, repetitions, rephrasing, fillers such as *well* and *you know*, voiced pauses (usually written *er* or *umm*), coordination.
- 1.2. *Low redundancy in writing*: non-finite subordinate clauses, ellipsis of common elements elsewhere in the discourse, nominalizations, attributive adjectives, subordination.

### **2. Differences related to situational factors**

- 2.1. *Context-dependence in speech*: pronouns, context-dependent ellipsis, *yes / no* questions, tags, *wh*-questions.
- 2.2. *Autonomy in writing*: long subject noun phrases, anonymous style.

### **3. Differences related to functional contrasts**

- 3.1. *Informality in speech*: intimacy signals (*you know, you see, well*), contractions, *this* and *these* for indefinite reference, clause completers (like *and that*).
- 3.2. *Formality in writing*: formal relatives, impersonal use of the passive.

### **4. Differences related to absence of prosodic features in writing**

- 4.1. *End-focus* (emphasizing elements that occur clause-finally): passive, cleft sentences.
- 4.2. *Emphasis*: interrupting constructions (delaying the completion of the basic clause pattern).
- 4.3. *Thematic prominence*: change of word order.

(Adaptació a partir de Perera, 1984: 182-199)<sup>58</sup>

---

<sup>58</sup> La llista de trets de K. Perera (1984) és, amb la de F.N. Akinnaso (1982), de les més extenses que he trobat. La que ofereixo és, però, una versió més reduïda en la qual he omès diverses característiques pròpies de la llengua anglesa que, en una comparació ràpida, m’ha semblat que no es donen en català. La meua adaptació, però, necessitaria un acarament més detallat entre les dues llengües.



**13– J.J. Gumperz, H. Kaltman i M.C. O’Connor (1984)**, en la mateixa línia que els autors anteriors, consideren que la llengua escrita ha de suplir amb complexitat sintàctica allò que la llengua oral resol amb la prosòdia (vegeu la citació que encapçala aquest punt) o amb el context. A més, en la producció del discurs escrit es disposa de temps per construir una sintaxi elaborada, cosa que no permet la producció *on-line* del discurs oral:

“The syntactic constructions associated with written language are those that supply information usually available from context in speaking, and those that, perhaps for reasons of cognitive complexity, tend not to be used in the more time-pressured oral speech situation: subject + predicate constructions instead of referent + proposition, subordinate rather than coordinate constructions, passive verb choice, definite articles and nominal modifiers instead of demonstratives used in a contextually referring way, etc. (...) The conversation was made up of single-proposition sentences. The thematic connections that linked these into a successive argument were signaled through prosody and lexical markers of intent. The same cohesive structure is signalled in the written passage through (a) more complex syntactic structure, and (b) transitional words.” (Gumperz, Kaltman i O’Connor, 1984: 8-14)

**14– H. Calsamiglia i A. Tusón (1999)**, en una obra ben recent, resumeixen en poques paraules el punt de vista majoritari sobre la complexitat sintàctica en relació amb la variació funcional:

“En general, puesto que se tiende a facilitar la comprensión por parte de la audiencia, en las formas más comunes y habituales de discurso oral se tiende al uso abundante de la yuxtaposición y la coordinación para relacionar oraciones y a un menor uso de nexos de subordinación. (...) La selección sintáctica también sirve como marcador de la variedad funcional o registro que se utiliza. En general, se puede decir que a menor complejidad sintáctica se corresponde un registro más coloquial y a mayor complejidad sintáctica, un registro más formal o culto.” (Calsamiglia i Tusón, 1999: 59)

Defensen punts de vista similars als que s’acaben de veure les obres d’altres autors que em limitaré a esmentar, sense citar-ne cap fragment, per no fer inacabable la lectura d’aquest punt. Es tracta de J. Vendryes (1923)<sup>59</sup>, P. Fraisse i M. Breyton (1959: 70), E.A. Nida (1967: 156), P. Kay (1977: 21-24), A.M. Vigara (1980: 21), L. Cortés (1986a: 176-

---

<sup>59</sup> Fragment citat per L. Payrató (1988: 104).

204), G. Herrero (1988: 193), J. Lerot (1991: 148), M.T. Serafini (1992: 174), S. Greenbaum i G. Nelson (1995: 16-17), J. Miller (1995: 132), J.J. Bustos (1997: 10-12) i S. Luque i S. Alcoba (1999: 34-36). La sèrie podria ser més llarga ja que, com s'ha dit, la simplicitat de la llengua oral i la poca presència que hi té la subordinació són llocs comuns que es poden rastrejar al llarg de tota la tradició lingüística. En tot cas, les dues llistes que he donat ja són de prou envergadura i el prestigi de molts dels noms que hi figuren és també prou indiscutible.

El criteri més generalitzat és, doncs, que la llengua oral presenta una simplicitat sintàctica i lèxica, manifestada en el predomini de les conjuncions de coordinació i en la redundància de vocabulari, a causa de la producció i la recepció en temps real del discurs que no dóna prou temps al parlant per elaborar-lo d'una manera més compacta. La llengua escrita, al seu torn, es mostra complexa sintàcticament i lèxicament, a través de la major presència de conjuncions de subordinació i d'una densitat lèxica superior. La idea complementària a aquesta, també molt defensada, és que la prosòdia, el context immediat i els llenguatges no verbals associats amb la llengua oral constitueixen una riquesa comunicativa que compensa sobradament la falta d'elaboració sintàctica. El llenguatge escrit, en canvi, ha de suplir la manca d'aquests elements amb la puntuació, la tipografia i l'estructuració sintàctica. L'esquema següent pot ser la síntesi de les idees exposades en aquest punt i en l'anterior:

<b>MODE ORAL</b>	<b>MODE ESCRIT</b>
<b>Simplicitat lèxica i gramatical</b>	<b>Complexitat lèxica i gramatical</b>
<b>Complexitat prosòdica, de context immediat i de llenguatges no verbals</b>	<b>Simplicitat prosòdica, de context immediat i de llenguatges no verbals</b>

### 2.3.3- LA COMPLEXITAT ESTRUCTURAL DE LA LLENGUA ORAL (I): M.A.K. HALLIDAY

**“Contrary to what many people think, spoken language is on the whole more complex than written language in its grammar; and informal spontaneous conversation, especially sustained and rather rapid conversation, is the most grammatically complex of all.”**

M.A.K. Halliday, “Differences between Spoken and Written Language: Some Implications for Literacy Teaching”, 1979, pàg. 47.

Als punts anteriors s’han vist diverses mostres de l’ampli conjunt d’estudis que han postulat la simplicitat sintàctica i/o lingüística de la llengua oral. Els autors que han defensat la posició contrària són pocs i han tingut com a figura més destacada M.A.K. Halliday. Aquest lingüista, en estudis de tipus gramatical i en treballs sobre el mode del registre (1979, 1985a, 1985b i 1987), ha afirmat que la llengua oral és gramaticalment més complexa que l’escripta. M.A.K. Halliday vol combatre l’assumpció generalitzada, tant en el món de la lingüística com en la societat en general, que la llengua oral és simple, elemental i, fins i tot, caòtica, al costat de la llengua escrita que és considerada més complexa, elaborada i purificada<sup>60</sup>. La llengua oral i la llengua escrita no són més simples o més complexes l’una que l’altra, sinó que ambdues presenten menes diferents de complexitat:

“(…) speech is not, in any general sense, ‘simpler’ than writing; if anything, it is more complex. (...) Looked at from the point of view of the sentence structure, it is the spoken text that appears more complex than the written one.” (Halliday, 1987: 64-65)

“(…) with adults the sentence structure of spontaneous conversation can reach a remarkable degree of complexity, such as is rarely attained in writing and indeed is difficult to follow when written down. At first sight this seems surprising, since we are

---

<sup>60</sup> Com que la llengua escrita passa per un procés de codificació gramatical, s’ha fet la filigrana conceptual de considerar-la “el model” o “l’essència” de l’idioma, i la llengua oral una “degeneració” de la llengua escrita, de manera que moltes persones pensen sincerament que hi ha individus de determinats dialectes o grups socials que quan parlen “fan malbé” o “embruten” la llengua nacional.

accustomed to thinking of written language as having the more complex syntax of two. But it is not really surprising when one takes into account the nature of the two media. Writing is a deliberate and, even with modern technology, a relative slow process; the text is created as an object and is perceived by the reader as an object —it exists. Spoken text does not exist; it happens. The text is created, and is perceived by the listener, as a process. (...) So, while speech and writing can both be very complex, the complexities tend to be of different kinds. The complexity of speech is choreographic —an intricacy of movement. That of writing is crystalline —a denseness of matter.” (Halliday, 1979: 49)

M.A.K. Halliday utilitza la metàfora de l'absència o la presència de moviment per afirmar que la llengua escrita té l'estructura *crystal·lina* del diamant mentre que la llengua oral la té *coreogràfica*, com el ball. L'oralitat es basa, com un riu o com una dansa, en el moviment i en la fluïdesa. La descripció gramatical d'aquest moviment no és una tasca fàcil i per això no disposem de gramàtiques de la llengua oral. La tendència a veure l'oralitat com a simple és, en realitat, producte de la falta d'instruments gramaticals adequats i d'una tradició cultural que tendeix a valorar l'escriptura i a menystenir la parla. La llengua oral, segons M.A.K. Halliday, no només no és simple, sinó que és la forma de llenguatge on el sistema lingüístic es desenvolupa al màxim i caldrien gramàtiques coreogràfiques per poder-la abastar:

“It is in spontaneous, operational speech that the grammatical system of a language is most fully exploited (...)” (Halliday, 1985a: xxiv)

“Since the study of grammar grew out of writing —it is when language comes to be written down that it becomes an object of study, not before— our grammars are grammars of the written language. We have not yet learnt to write choreographic grammars; so we look at spoken language through the lens of a grammar designed for writing. Spoken discourse thus appears as a distorted variant of written discourse, and not unnaturally it is found wanting.” (Halliday, 1987: 67)

Per a M.A.K. Halliday la complexitat de la llengua oral no s'ha d'oposar a la de la llengua escrita, sinó que ambdues són complementàries i afecten aspectes diferents del llenguatge. La complexitat de cada mode es manifesta lingüísticament d'una manera pròpia: la de l'oral a través de la gramàtica i la de l'escrit a través del lèxic. Vegem-ho a continuació en les seves paraules:

“Spoken and written language, then, tend to display different KINDS of complexity; each of them is more complex in its own way. Written language tends to be lexically dense, but grammatically simple; spoken language tends to be grammatically intricate, but lexically sparse. But these but should be really ands, because the paired properties are complementary, not counterexpectative.” (Halliday, 1987: 66)

“In linguistic terms, spoken language is characterized by complex sentence structures with low lexical density (more clauses, but fewer high content words per clause); written language by simple sentence structures with high lexical density (more high content words per clause, but fewer clauses). We could express this even more briefly, though at the cost of distorting it somewhat, by saying that speech has complex sentences with simple words, while writing has complex words in simple sentences.” (Halliday, 1979: 49)

Així doncs, els trets bàsics que diferencien la llengua oral de l'escrita són la *densitat lèxica* i la *complexitat gramatical*. Això s'explica, com hem vist abans, per la naturalesa mateixa de cada mode, oral i escrit, i per la manera específica que té cadascun de representar la realitat que comunica. La llengua oral, que és procés i moviment, representa la realitat per mitjà de verbs i de clàusules, i propicia, per tant, un estil verbal de construcció dels enunciats. La llengua escrita, que és producte i objecte estàtic, ho fa per mitjà de noms i prefereix un estil més nominal:

**“The spoken language is, in fact, no less structured and highly organised than the written.** It could not be otherwise, since both are manifestations of the same system. (...) Speech and writing are two rather different ways of representing our experience. The basis of the distinction is this. Written language represents phenomena as **products**. Spoken language represents phenomena as **processes**. Each code represents reality as being like itself. A piece of writing is an object; so what is represented by written language is also given the form of an object. (...) But when you talk, you are doing; so when you represent by talking you say that something happened or something was done. (...) To refer to an object takes a noun; but to say that something is happening takes more than a verb —it takes a clause. The immediate reason for this lies in the grammar: verbs do not occur by themselves.” (Halliday, 1985b: 79-82)

L'argumentació de M.A.K. Halliday és, a més de sorprenent, impecable en el fil lògic i absolutament coherent amb la resta de la seva obra. I la contradicció entre aquesta opinió i les que s'han vist en els dos punts anteriors és tan gran que demana —i exigeix—

reflexió teòrica i recerca empírica immediates per tal de resoldre el problema científic que es planteja<sup>61</sup>.

Pel que fa al darrer punt que he destacat, és a dir, el de la influència de la naturalesa de cada mode en la forma lingüística amb què aquest mode representa el món, m'ha sobtat la coincidència amb un fragment de la *Retòrica* d'Aristòtil, que he trobat com a citació d'autoritat encapçalant algun article dedicat a aquest tema. Vegem-lo a continuació:

“(…) no ha d'oblidar-se que cada gènere s'harmonitza amb un estil diferent: l'estil dels escrits, en efecte, i l'estil dels debats no és pas el mateix. (...) L'estil dels escrits, nogensmenys, és sens dubte més precís, mentre que l'estil dels debats és més adequat a l'acció. (...) Ahora, tot comparant-los, els discursos escrits apareixen en els debats com a lacònics, mentre que els discursos dels oradors, tot essent ben pronunciats, semblen vulgars i simples quan es tenen a les mans. (...) allí on domina sobretot l'acció, l'exactitud hi ocupa molt poc lloc. Això s'esdevé també allí on tot depèn de la veu i sobretot allí on tot depèn de la veu forta i enèrgica.” (Aristòtil, c.335aC: 282-285)

---

<sup>61</sup> La difusió d'aquestes idees, potser pel contrast que tenen amb la concepció dominant, ha estat escassa i lenta. En el món anglosaxó no hi he trobat cap ressò fins a K. Beaman (1984) i D. Biber (1988). A casa nostra, la primera referència que conec és la de V. Salvador (1989), que en fa un resum precís i complet a partir de M.A.K. Halliday (1985b). Vegem a continuació una part d'aquest resum:

“Siga com siga, el que Halliday assenyala com a característica diferenciadora essencial —que ja es palesa en algunes estratègies de focalització informativa— és la tendència a la nominalització en el discurs escrit, per contrast amb una proclivitat a la utilització més freqüent de verbs en l'oralitat. Això implica una oposició entre la densitat lèxica com a estratègia de l'escriptura («tota la carn del missatge —diu— es troba en els nominals») i el dinamisme sintàctic (*intricacy of movement*, que remet a la idea de condició abrupta, moguda, sorpresiva) del discurs oral. Per explicar-ho recorre al contrast metafòric entre el diamant i el riu. (...) La conclusió principal de Halliday és que el llenguatge oral i l'escrit corresponen a dues estratègies diferents de representar la realitat, d'activar el *meaning potential* d'una llengua: en un cas es representa els fenòmens com a *processos*, en l'altre com a *productes*.” (Salvador, 1989: 17-18)

Un any després, H. Calsamiglia (1990) també es refereix, d'una manera més sumària, a la posició de M.A.K. Halliday —malgrat que després no s'integri en el manual publicat el 1999 amb A. Tusón que he citat anteriorment (vg. § 2.3.2):

“La gramàtica [de la llengua oral] té particularitats. (...) La majoria dels autors criden l'atenció sobre el predomini de les relacions paratàctiques entre els enunciats (o les *unitats informatives* de Chafe). Simplicitat, doncs, de les relacions, juxtaposades o coordinades. Darrerament, Halliday (1985) ha destacat, però, la relativa complicació sintàctica que comporta la llengua oral a causa de la poca densitat del lèxic, les paràfrasis, les repeticions i redundàncies.” (Calsamiglia, 1990: 42)

Sobta i complau de veure idees tan similars en autors tan allunyats en el temps. Així, Aristòtil considera l'estil dels debats *‘més adequat a l'acció’*, i el dels escrits *‘més precís’*, com M.A.K. Halliday, que caracteritza la llengua oral com un *procés* en moviment i la llengua escrita com un *producte* estàtic. Aristòtil també s'adona de la trampa que representa judicar els discursos orals a partir de la transcripció, que llavors semblen *‘vulgars i simples’*, així com M.A.K. Halliday deia que quan la parla és transcrita i *‘reduced to writing’* sembla *‘like a dog’s dinner’* (vg. fragment ja citat, § 2.1.2). Aristòtil, per fi, podria apartar-se de la posició de M.A.K. Halliday i donar la raó als autors del punt anterior quan afirma que *“allí on tot depèn de la veu”* —és a dir, de la prosòdia—, *“l’exactitud hi ocupa molt poc lloc”*<sup>62</sup>.

Al llarg d’aquest punt s’ha fet palesa, doncs, una contradicció notable entre la visió de W. Chafe i de tota la tradició lingüística, i les afirmacions de M.A.K. Halliday que postulen la posició contrària. D’acord amb aquest autor, l’esquema que havia donat al final del punt anterior seria substituït per un altre que, tot manejant conceptes similars, els oposa de manera diferent:

<b>MODE ORAL</b>	<b>MODE ESCRIT</b>
<b>Simplicitat lèxica (redundància)</b>	<b>Complexitat lèxica (densitat)</b>
<b>Complexitat gramatical</b>	<b>Simplicitat gramatical</b>

---

<sup>62</sup> Malgrat tot, sembla difícil de discernir quina seria la interpretació exacta d’aquest passatge. L. Payrató (comunicació personal) troba que en el mot *precís* —que, de fet, no és d’Aristòtil, sinó un mot català del traductor— hi ha també un judici de valor, que pot dur a la típica infravaloració posterior de la llengua oral com a imprecisa o menys rica.

En aquest esquema, els dos aspectes destacats de la complexitat lingüística —densitat lèxica i elaboració sintàctica— queden oposats en una relació inversament proporcional. Així, el fet de ser *simple* o *complex* lingüísticament no és propietat d'un dels dos modes, sinó que a cadascun se li atorga un aspecte o una cara de la complexitat. La prosòdia, el context immediat i els llenguatges no verbals, no presents en aquest quadre, serien elements afegits a la complexitat del mode oral.



### 2.3.4- LA COMPLEXITAT ESTRUCTURAL DE LA LLENGUA ORAL (II): ALTRES AUTORS

**“The subordination features (...) seem to be associated with the expression of information under real-time production constraints, when there is little opportunity to elaborate through precise lexical choice.”**

D. Biber, *Variation across Speech and Writing*, 1988, pàg. 107.

En les pàgines següents, com he fet abans per als que mantenen la posició contrària, consignaré diversos autors que, amb connexió directa o sense, amb M.A.K. Halliday, han postulat la complexitat sintàctica de la llengua oral o han arribat a conclusions similars amb més o menys matisos i observacions. Els tractadistes que citaré —també, per ordre cronològic de publicació— són M.E. Poole i T.W. Field (1976), L.V. Aracil (1978), G. Brown (1978), L. Polanyi i R.J.H. Scha (1983), K. Beaman (1984), A. Narbona (1986, 1988 i 1995), D. Biber (1988), J.R. Martin (1989), M.I. Macaulay (1990), D. Allerton (1991), C. Goodwin (1994), C. Blanche-Benveniste (1995a) i A. Briz (1998):

**1– M.E. Poole i T.W. Field (1976)** troben la sintaxi de la llengua oral més complexa que la de la llengua escrita i justifiquen aquest fet per les diferències en el temps de producció d’ambdós modes. D’aquesta manera, l’argumentació de molts dels autors del punt 2.3.2 es capgira i la producció de la llengua oral en temps real (*on-line*), que allà era motiu de simplicitat, aquí és causa precisament de la complexitat sintàctica:

“The findings relative to syntax suggest that oral language is more structurally complex than written. (...) At first this seems something of a paradox. In writing, however, there is temporal advantage which provides a greater opportunity for structural planning, a factor which may account for both increased sentence length and structural simplicity. In writing, the focus is more on stylistics. (...) Pleas for simplicity, conciseness, precision, and clarity contained in style manuals could therefore mitigate against the use of syntactically complex and elaborate structures.” (Poole i Field, 1976: 309)

**2– L.V. Aracil (1978)**, pioner a casa nostra a tractar aquest tema, aboga clarament per la complexitat sintàctica i discursiva de la llengua col·loquial i explica, amb el seu habitual to reivindicatiu, les raons per què no ha estat estudiada:

“Hi ha aspectes molt interessants des d’aquest punt de vista —per exemple: les diferències (meta)sintàctiques entre col·loquial i no-col·loquial. La Gramàtica tradicional les ha negligides maquinalment —segurament perquè pressuposava que parlar col·loquial és «parlar malament». Com que el judici de valor és arbitrari, no el refutaré pas. Ara: la imputació indemostrada de «simplicitat» i «pobresa» és demostrablement falsa. L’estructura més profunda (a nivell de discurs) del col·loquial és, si de cas, prou més complexa que la de qualsevol discurs formal. La veritat és que el col·loquial té qualitats ben «positives» —com ara una agilitat desconcertant i una fecunditat exuberant— que desafien l’anàlisi superficial i posen a prova l’enginy inventiu del lingüista més expert. La (meta)sintaxi del col·loquial ha restat inexplorada perquè, per alguna raó, la Lingüística estricta moderna (com la vella Gramàtica) han preferit d’eludir-la. Una de les mil proves d’això és la confusa i intrigant categoria de les «interjeccions» —fins i tot les «conjuncions» (a excepció possiblement de «i») són embarassoses. Hom diria que un detallisme miop, justificable en una etapa inicial, ha descartat com a «marginal» tot allò que transcendia l’estructura de l’oració simple regular.” (Aracil, 1978: 238)

L.V. Aracil no només troba inacceptable parlar de “pobresa” en referir-se a la llengua oral, sinó que coincideix amb M.A.K. Halliday a considerar que té “*una agilitat desconcertant i una fecunditat exuberant*”, expressió que recorda molt la “*intricacy of movement*” citada anteriorment i que també destacava V. Salvador (1989) (vg. § 2.3.3).

**3– G. Brown (1978)**, sense parlar específicament de complexitat, explica d’una manera que sembla ben satisfactòria la successió sintàctica en la llengua oral, a partir de la noció psicològica de *memòria a curt termini*:

“The speaker’s control of syntax over a series of clauses is obviously limited by his short term memory. If he is embarked on his third dependent clause it may well be in a relation to the main clause at which he would shudder if he saw it written down. But of course in speech he would probably only be conscious of the detailed syntax of the previous clause (as would his hearer, for the same reasons). So one frequently finds, in transcripts of spontaneous speech, that in a sequence of clauses ABC the sequence AB is fine and the sequence BC is fine but the overall relationship is not.” (Brown, 1978: 274)

Aquest desenvolupament amb enllaç sintàctic entre clàusules contigües i pèrdua de l'estructuració sintàctica —no pas de la informativa— a llarga distància és típic de la llengua oral i pot desembocar en un nombre més alt de connectors, ja que els encadenaments ABC... arriben a ser bastant llargs i en cada seqüència bimembre AB, BC, CD, etc. hi pot aparèixer un nexa oracional o un connector textual.

**4– L. Polanyi i R.J.H. Scha (1983: 264)**, des de l'àmbit de l'anàlisi del discurs, afirmen que una conversa qualsevol “*may exhibit a complex interactively constructed syntactic structure*” i que aquesta estructura es relaciona amb mecanismes semàntics, lèxics, entonatius i pragmàtics també complexos.

**5– K. Beaman (1984)** parteix dels treballs de W. Chafe (1982) i de M.A.K. Halliday (1979), n'assenyala la contradicció i fa una anàlisi de la problemàtica en narratives orals i escrites, en les quals cerca l'aparició de clàusules coordinades i subordinades. Els seus resultats mostren que hi ha tanta subordinació en les narratives orals com en les escrites i que les diferències entre els dos modes radiquen en el tipus de subordinació (p. ex., en la llengua oral hi ha més clàusules subordinades nominals amb el verb en forma personal i en la llengua escrita n'hi ha més amb el verb en forma no personal). Les conclusions d'aquesta autora són les següents:

- 1- “Directly contrary to many previous assumptions, subordinate clauses are frequent in spoken as well as in written narrative.”
- 2- “The increased lexical density of writing is balanced by a relatively greater number of dependent clauses in speech.”
- 3- “(...) Previous studies have often confounded the spoken / written distinction with differing degrees of formality and purpose of the discourse.” (Beaman, 1984: 78-80)

En aquesta tercera conclusió, K. Beaman es planteja una pregunta clau que reprendré en l'apartat 3.4: “*Per què s’han produït resultats tan discrepants en les recerques de diferents autors sobre aquest tema?*” La resposta és de gran interès metodològic: atès que l’oposició oral / escrit no és dicotòmica i que les convencions de gènere hi poden influir en gran manera, els discursos comparats han de diferir, si és possible, únicament en el canal i han d’estar controlades totes les altres característiques de formalitat,

especialització, gènere, tipus de text, etc. Altrament, les diferències gramaticals observades entre els discursos comparats poden haver estat causades pel contrast oral / escrit sinó per d'altres factors contextuals. Però, com es veurà, aquesta tasca de control comporta, al seu torn, problemes importants de validesa de la mostra.

**6– A. Narbona (1986, 1988 i 1995)** —a parer meu, l'autor de l'àmbit hispanòfon que ha tractat amb més encert la qüestió de la sintaxi del registre col·loquial— parla dels tòpics, a vegades no comprovats, que existeixen sobre la llengua oral i ho fa des d'una visió oberta, no limitada a la gramàtica oracional (posar a prova algun d'aquests tòpics és una de les tasques de la present recerca):

“(…) Es casi un tópico, por ejemplo, hablar del predominio de la yuxtaposición en el lenguaje coloquial, frente al empleo abundante de la subordinación en la lengua escrita, lo que ha sido puesto en relación frecuentemente con la andadura sintáctica de la lengua primitiva y se ha considerado como reflejo de pobreza o escasez de recursos de trabazón sintáctica. No nos parece que sea ésta la vía más indicada para entender la peculiar arquitectura de la frase coloquial. (...) En primer lugar, conviene decir que la afirmación de que la frase coloquial es breve y de que en ella no abundan los términos de subordinación, aun en el caso de que estuviera empíricamente comprobada, sería escasamente reveladora. (...) La segunda observación se refiere a esa supuesta falta de elaboración de la sintaxis coloquial. Es preferible arrancar de la hipótesis de que el tipo básico de organización sintáctica es distinto del dominante en la lengua culta. Hay que dejar de ver en la disposición secuencial de la frase coloquial una mera «des-ordenación» del orden tenido como lineal, normal o lógico, y contemplarla como una ordenación que responde a factores prioritarios diferentes, factores que, insistimos, tienen que ver con el valor informativo que el hablante asigna e impone a los distintos constituyentes, con independencia de las relaciones estrictamente funcionales.” (Narbona, 1986: 164-167)

“A falta de término mejor, he calificado en alguna ocasión este discurrir propio del coloquio conversacional de *parcelado*, término que no debe entenderse como sinónimo de *fragmentario*. Lo de menos es que las frases sean más breves y escaseen las consideradas *complejas*, particularmente las *subordinadas*; la inclinación a organizar los discursos en numerosas *parcelas*, cada una de las cuales dispone de su propia configuración melódica, lejos de suponer una vertebración *centrífuga* y desmembrada, potencia en la mayoría de

los casos una articulación del todo resultante de un modo mucho más eficaz y pertinente.”  
(Narbona, 1995: 41)<sup>63</sup>

Efectivament, una de les idees que més combat A. Narbona és la pretesa “simplicitat” o “desordre” de la sintaxi oral, alhora que insisteix en la falta de comprovació empírica de la suposada poca abundància de subordinades en la llengua oral. A més, en la segona citació, canvia el terme *fragmentari* per *parcel·lat* en referir-se a la naturalesa de la sintaxi col·loquial<sup>64</sup>. Les afirmacions d’aquest autor sobre la subordinació i els marcadors discursius les reprendré més endavant en els apartats corresponents (vg. 5.2 i 5.5).

**7– D. Biber (1988: 102-103)**, en un treball que esmentaré sovint, arriba a uns resultats que confirmen, si bé només parcialment, la posició de M.A.K. Halliday: troba algunes menes de subordinació com a típiques de determinats textos orals. Es tracta de les construccions *causative subordination*, *conditional subordination*, *sentence relatives* i *wh clauses*, que es relacionen amb la producció implicada en la situació (“*Factor 1: Involved production*”), i de les construccions *that clauses as verb complements*, *that relative clauses on object positions*, *that clauses as adjective complements* i *wh relative clauses on object positions*, que es relacionen amb la producció en temps real (“*Factor 6: On-line Informational Elaboration*”). D. Biber interpreta aquests resultats de la manera següent:

“The distribution of Factor 1 disagrees with the expectations of O’Donnell (1974), Kay (1977), and others —that subordination marks greater elaboration and thus should be characteristic of informational discourse— but it agrees with the findings of Poole and Field (1976) and Halliday (1979) that subordination is associated with the production constraints characteristic of speech. (...) The subordination features grouped on Factor 6 apparently mark informational elaboration that is produced under strict real-time constraints, resulting in a fragmented presentation of information accomplished by tacking

---

<sup>63</sup> Amb els mots “*vertebración centrífuga*”, A. Narbona es refereix a l’expressió encunyada per M. Seco segons la qual la sintaxi oral presenta una “*tendencia centrífuga*”.

<sup>64</sup> Una de les tasques inicials en l’estudi de la llengua oral ha de ser la de canviar els termes pejoratius tradicionals per uns altres que, tot reflectint l’especificitat del fenomen referenciat, no ho facin d’una manera que impliqui una valoració negativa.

on additional dependent clauses, rather than integrated presentation that packs information into fewer constructions containing more high-content words and phrases (as on Factor 1). (Biber, 1988: 107-113).

Per tant, D. Biber, després de plantejar les contradiccions en els resultats de les recerques, corrobora la posició de M.A.K. Halliday respecte de la complexitat sintàctica i de la relació de la subordinació amb la mobilitat constructiva de la llengua oral, si bé considera que aquesta relació es dóna especialment en els gèneres orals de funció informativa —que, com hem vist anteriorment, no és la típica de la llengua oral, però tampoc no li és estranya (vg. les paraules de G. Brown i G. Yule, 1983b: 16-23, en el punt 2.2.3)<sup>65</sup>. A causa de la producció dels discursos simultàniament a la recepció, en la llengua oral es distribueix la informació en més clàusules, principals i dependents, mentre que en la llengua escrita es comprimeix en un nombre menor d'estructures que contenen una major densitat de paraules i sintagmes. Una altra vegada, com en M.A.K. Halliday (1979) i M.E. Poole i T.W. Field (1976), la producció espontània, en temps real, s'ha convertit en argument per a la complexitat i no per a la simplicitat de la llengua oral.

**8– J.R. Martin (1989)**, del cercle australià de M.A.K. Halliday, repeteix la idea que la llengua oral és més complexa gramaticalment, però matisa també que això es produeix quan el parlant produeix un torn de parla llarg. D'aquesta manera, s'aparta de M.A.K. Halliday i coincideix amb G. Brown i G. Yule (1983b) i D. Biber (1988) en la influència que tenen els torns de paraula llargs i d'intenció informativa en la complexitat sintàctica:

---

<sup>65</sup> Aquesta especificitat dels gèneres orals informatius resulta de gran valor per a la meua recerca, en la qual estudio uns discursos, les classes magistrals, que s'inclouen plenament en aquesta categoria i crec que és una qüestió clau a l'hora de situar correctament les afirmacions de M.A.K. Halliday. Com ja comentaré en analitzar els resultats, les classes magistrals són discursos orals informatius i monologats que es produeixen amb la intenció explícita, regulada socialment, que “s'assemblin a la llengua escrita”, encara que mai no aconseguixin d'aproximar-s'hi absolutament. D'altra banda, potser això és el que explicaria que D. Biber (1988: 37) hagi declarat que la conversa és “*structurally simple*” i “*fragmented*” (vg. la citació del subapartat 3.2.5), afirmació que sembla en contradicció amb la línia argumentativa principal del seu treball. Sigui com sigui, D. Biber i M.A.K. Halliday difereixen en el fet que el primer atribueix la complexitat sintàctica als discursos orals informatius i el segon, en canvi, considera que com més col·loquial i espontània és la llengua oral més complexitat gramatical presenta.

“Many people have the mistaken idea that speaking is simpler than writing, and that this is one reason writing is hard to learn<sup>66</sup>. In fact both speech and writing make use of complex linguistic patterns, although the complexity tends to be of different kinds. The complexity of speaking is grammatical —very long sentences may well be produced once speakers depart from the rapid-fire repartee of texts like Text 2.3 (see before) and take a longer turn, with clause added in a complex array.” (Martin, 1989: 78)

**9– M.I. Macaulay (1990)** planteja l’oposició entre els punts de vista de W. Chafe i M.A.K. Halliday i busca el nombre d’oracions simples i complexes que apareixen en un corpus de textos orals i escrits. Tot i que trobi més oracions complexes en la llengua escrita, M.I. Macaulay afirma que les posicions de tots dos autors són correctes —és a dir, que cadascuna té la seva part d’encert:

“Taking all sentences from all texts used in this study, I found marked differentiation only in the percentages for complex sentences; 20 per cent of all oral sentences are complex, while the number is 35 per cent for written sentences. This finding would seem to confirm Chafe’s view that written speech is «integrated», supporting more subordination than oral speech, and disconfirm Halliday’s characterization of written speech as «crystalline», supporting primarily simple sentence structures. However, simple sentences appear with almost comparable frequency in all oral and written texts (49 % oral; 45 % written), so that Halliday’s argument for the simplicity of written syntax is equally validated. Both theorists are, in fact, correct in their respective analyses of written speech. What the data show is that both oral and written speech favour simple sentences structures, while written speech alone seems to favour complex structures.” (Macaulay, 1990: 162-163)<sup>67</sup>

---

<sup>66</sup> La dificultat de l’aprenentatge de la llengua escrita en contrast amb la facilitat —la *naturalitat*, en sentit literal— amb què els humans aprenen a parlar ha fet pensar, efectivament, que la llengua escrita és més complexa. Tot i això, crec que cal fer la reflexió en el sentit contrari. La llengua oral la desenvolupem com a part del nostre creixement com a persones i és per això que “no costa”, perquè estem programats genèticament per fer-ho, mentre que la llengua escrita és un afegit, de gran importància cognitiva per desenvolupar el pensament de tipus lògic i científic, constitutiu de la condició humana com a potencialitat, però no com a component obligatori (vg. K. Perera, 1984, fragment citat en el punt 2.1.3). De fet, si fos al revés, que primer escrivíssim i després parléssim, l’aprenentatge de la llengua oral, amb la sintaxi més lliure, la fonètica, l’entonació, la fraseologia, els girs, els torns de parla, les regles de la conversa, les expressions facials, gestuals i corporals, etc., seria enormement més difícil que no ho és ara el de l’escrita.

<sup>67</sup> L’argumentació de M.I. Macaulay (1990) en aquest fragment sembla estranya. Malgrat que el nombre de *simple sentences* sigui similar en els dos modes, el de *complex sentences* no ho és pas i, per tant, el conjunt d’ambdós resultats sembla donar molt més la raó a W. Chafe que no pas a M.A.K. Halliday. La seva conclusió, però, que considera correcta en algun aspecte cadascuna de les dues opinions enfrontades, coincideix amb la que serà la del present treball, encara que jo em basaré en arguments i en dades diferents. D’altra banda, cal destacar que, com ja he assenyalat en el punt 3.4.1, els textos orals de M.I. Macaulay no són converses sinó entrevistes formals d’intenció predominantment informativa, cosa que per

**10– D. Allerton (1991)**, sense fer referència estrictament a la sintaxi, defensa que la llengua oral no és més simple que l'escrita i que, a més, en alguns aspectes conté distincions més subtils:

“I intend to show that in some respects the rules for spoken English are stricter than those that seem to apply in written, and that spoken English can make more subtle distinctions. Each of my examples will involve a rule or set of rules of English which only apply consistently in spoken English. Written English is less strict or less systematic in these areas. Each example can be linked to a different basic characteristic of spoken language, the first to rhythm, the second to intonation, the third to situation-boundness, and the last to redundancy.” (Allerton, 1991: 470)<sup>68</sup>

**11– C. Goodwin (1994)**, en la línia de L. Polanyi i R.J.H. Scha (1983) i sense parlar explícitament de complexitat, afirma que la conversa és una de les formes d'organització social més intricades, que s'hi poden trobar *“highly structured products”* i que les oracions de la conversa es construeixen en un procés interactiu amb la col·laboració dels diversos interlocutors. La sintaxi resultant presenta un dinamisme molt pròxim a la *“intricacy of movement”* de M.A.K. Halliday:

“Conversation is one of the most pervasive, as well as intricate, forms of human social organization. (...) Further, the vocal and nonverbal activities of the participants within the turn produce highly structured products (...). For example, in traditional linguistics it has been assumed that the analysis of sentences can be performed upon examples isolated from the processes of interaction within they naturally emerge. (...) The analysis presented here has argued, to the contrary, that the sentence actually produced in a particular turn of talk is determined by a process of interaction between speaker and hearer. Their collaborative work in constructing the turn systematically modifies the emerging structure of the sentence, adding to it, deleting from it [sic], and changing its meaning. Insofar as this is the case the procedures utilized to construct sentences are, at least in part,

---

força ha de condicionar els resultats que obté i que, a més, els relaciona amb el que han dit D. Biber (1988) i J.R. Martin (1989) sobre l'oral informatiu monologat.

<sup>68</sup> D. Allerton (1991) analitza alguns aspectes com la posició dels mots en les interrogatives en anglès, les contruccions focalitzades del tipus *Ha estat en Joan qui s'ha menjat el pastís* i l'ordre dels constituents oracionals (subjecte - verb - complement directe - complements preposicionals), que creu que és més fix en la llengua oral que en l'escrita.



interactive procedures. (...) Both the length and the meaning of the sentence eventually constructed within a turn of talk might thus emerge as the products of a dynamic process of interaction between speaker and hearer.” (Goodwin, 1994: 199-217)

**12– C. Blanche-Benveniste (1995a: 25)** troba més subordinació en la llengua oral que en l’escrita (“*il est frappant de voir que les subordonnées sont plus nombreuses par oral que par écrit, et qu’elles sont utilisées de façons différentes.*”), tot i que, com comentaré en el punt 3.4.1, ho calcula amb una mostra de textos molt reduïda.

**13– A. Briz (1998)** atribueix al registre col·loquial de la conversa una ‘*sintaxis no convencional*’ i una sèrie de trets organitzatius que situen l’opinió d’aquest autor en una posició intermèdia entre les dues que tenim a debat. A. Briz elabora una llista de característiques sintàctiques, contextuals, fòniques i lèxiques del registre col·loquial que inclou les aportacions de diversos autors i que haurà de ser tinguda en compte com una bona guia en aquest camp de recerca. A continuació ofereixo alguns dels punts d’aquesta llista:

“La ausencia de planificación o, más exactamente, la obligada planificación rápida, y, así pues, el escaso control de la producción del mensaje, el tono informal, determinan una *sintaxis no convencional* y una estructura gramatical específica, que se concreta en los rasgos textuales, estructurales y de organización y formulación de los mensajes que siguen:

#### 1.1. LA SINTAXIS CONCATENADA

Uno de los rasgos más notables de la construcción coloquial es la concatenación y, a veces, acumulación de enunciados, no necesariamente independientes (Narbona, 1989: 166 y 180): la *sintaxis concatenada*, frente a la *sintaxis incrustada* del modo escrito y del registro formal (...).

#### 1.2. LA PARCELACIÓN

Como resultado de la concatenación y acumulación de enunciados, el modo de glosar se presenta parcelado, si bien no por ello menos trabado (Narbona, 1989: 163 y 192-94) (...).

#### 1.3. EL RODEO EXPLICATIVO

Todo ello favorece la paráfrasis, los continuos rodeos, consecuencia de lo cual la información avanza de forma lenta, a diferencia del modo rápido en que el mensaje se enuncia (...).

#### 1.4. LA REDUNDANCIA

La conversación coloquial presenta un alto grado de redundancia en las intervenciones (Narbona, 1989: 181-83). Las frecuentes repeticiones y reelaboraciones son mecanismos de formulación cuya función discursiva va más allá del mero papel retardatorio que algunos autores les han asignado (...).

#### 1.5. LA UNIÓN ABIERTA

La unión entre los enunciados es abierta, sin fuertes ataduras sintácticas (Narbona, 1989: 167 y 180), pero no por ello menos cohesionados. Esto permite ir y venir, volver sobre lo dicho, reformular, reelaborar, introducir digresiones, incrustar comentarios, etc. (...).

#### 1.6. LA CONEXIÓN A TRAVÉS DE CONECTORES PRAGMÁTICOS Y ENTONACIÓN

La trabazón y conexión entre los enunciados y, así pues, la cohesión textual y discursiva, se logra a través de lo que S. Gili Gaya (1976) denominó *enlaces extraoracionales* (...).

#### 1.7. EL ORDEN PRAGMÁTICO

La comunicación al instante, característica de la conversación coloquial, explicaría asimismo que el orden de palabras responda a la función pragmática de la topicalización y al realce informativo, sólo a veces meramente subjetivo, de los elementos. Las palabras se adelantan a menudo como preludios de la expresión organizada, a modo de titulares que presentan y anuncian la idea que se pretende desarrollar.” (Briz, 1998: 68-82)

Com es pot veure, en la llista d’A. Briz hi ha arguments que serien propis dels autors del punt 2.3.2, defensors de la simplicitat sintàctica de la llengua oral (concatenació i no incrustació, parcel·lació, etc.), però l’enfocament general no és que la llengua oral sigui pobre o simple, sinó que s’hi empren uns recursos especials (acumulació, reformulació, connexió extraoracional) que permeten que no sigui pas menys travada que l’escripta. A més, A. Briz cita contínuament A. Narbona, autor que ja he destacat pel tractament que fa de la sintaxi col·loquial.

Aquest conjunt de valoracions sobre el grau i el tipus de complexitat sintàctica i lingüística en general, de la llengua oral i de la llengua escrita, ofereix una panoràmica àmplia sobre la qüestió. Les dues conclusions principals que em sembla que se’n poden extreure són, en primer lloc, que el grau de complexitat sintàctica dels textos deu situar-se en algun punt intermedi entre la posició de W. Chafe i la de M.A.K. Halliday, ja que s’han aportat molts arguments i moltes dades en tots dos sentits, i sembla difícil de

pensar que tants investigadors, d'una banda o de l'altra, puguin equivocar-se completament —possibilitat que, malgrat tot, també ha de ser comprovada. I, en segon lloc, que aquesta posició intermèdia s'haurà d'acomodar al tipus de construcció sintàctica que s'analitzi en cada cas i podrà ser diferent segons si es tracta d'oracions subordinades o coordinades, i, encara, segons si és en el terreny de la subordinació substantiva, adjectiva o adverbial, etc. El present treball, per tant, haurà de tenir en compte els matisos de la complexitat —és a dir, la complexitat de la complexitat.

### 2.3.5- SOBRE LA NOCIÓ MATEIXA DE COMPLEXITAT LINGÜÍSTICA

**“The basic assumption of many linguists in the past, i.e. that subordination implies complexity, is false. The evaluation of syntactic complexity is simply more complex than that.”**

K. Beaman, “Coordination and Subordination Revisited: Syntactic Complexity in Spoken and Written Narrative”, 1984, pàg. 79-80.

Fins aquí he esmentat els punts de vista d'un bon nombre d'estudiosos sobre la complexitat lingüística en la llengua oral i en la llengua escrita. Per acabar amb el plantejament del tema, crec que cal qüestionar-se sobre el concepte mateix de complexitat lingüística.

En la citació que encapçala aquest punt, K. Beaman (1984) nega el postulat tradicional segons el qual és vàlida l'equació “subordinació = complexitat”. Aquesta és la mateixa qüestió que planteja A. Narbona (1995), que desfà la relació teòrica o conceptual — caldria fer-ho, també, empíricament— entre subordinació i complexitat o entre parataxi i simplicitat:

“La noción de *simplicidad*, además, se vincula generalmente a una presunta *falta de elaboración* o de *trabazón sintáctica*, y, en la práctica, a la utilización preferente de oraciones *simples* y, dentro de las llamadas *compuestas* o *complejas*, de las consideradas de menor grado de vertebración, esto es, las que se sirven de mecanismos de conexión *paratáticos*. Ha acabado por convertirse casi en un lugar común sostener que el escaso

empleo de subordinadas es una de las principales características del habla popular (...), que, por lo mismo, acaba siendo interpretada a menudo como *inferior* (...) o *deficitaria*. (...) Los conceptos de *parataxis* y de *hipotaxis* —junto con los de *hipertaxis* y *antitaxis*— han de entenderse como propiedades generales del lenguaje (...). En consecuencia, no cabe establecer una correspondencia estricta entre ellos y las nociones de *simplicidad* y *complejidad* sintácticas. Por lo mismo, el grado de *madurez* de una modalidad de uso no puede medirse, sin más, por el empleo que haga de estructuras *subordinadas*. El hecho de que la sintaxis del latín arcaico y la de los primitivos romances haya sido caracterizada igualmente como *suelta* o poco *trabada* (Wartburg, 1969)<sup>69</sup> revela que los usos orales requieren ser contemplados en cuanto tales, previamente a cualquier confrontación con las variedades escritas. (...) Descubrir su especificidad requiere, entre otras cosas, romper con la perspectiva lineal y secuencial que domina en la gramática oracional y adoptar un punto de vista discursivo global.” (Narbona, 1995: 40-41)

A. Narbona (1995), que es refereix a W. von Wartburg com A.M. Badia i Margarit, però per capgirar el camí argumentatiu que havíem vist en aquest darrer, acaba proposant de tractar la complexitat de la llengua oral amb una visió oberta en què la gramàtica incorpori la dimensió discursiva.

Aquesta ampliació al discurs com a sortida per al problema de la complexitat també és propugnada per M.J. Schleppegrell (1992). Aquesta autora assenyala que hi ha conjuncions considerades de subordinació en el discurs escrit que, en el discurs oral, poden funcionar com a elements coordinants i, fins i tot, com a marcadors discursius o omplidors<sup>70</sup>. M.J. Schleppegrell (1992) posa l'exemple de *because*, que en anglès col·loquial presenta usos no subordinants molt diversos. A la vista de la plurifuncionalitat d'alguns mots gramaticals ben importants en la construcció dels discursos, conclou que la noció de complexitat lingüística ha de ser revisada i que cal plantejar-la en el nivell supraoracional:

---

<sup>69</sup> A. Narbona treballa, curiosament, sobre la traducció francesa de l'obra de W. von Wartburg i no sobre la versió espanyola, traduïda i adaptada per D. Alonso i E. Lorenzo, que ja havíem vist citada per A.M. Badia i Margarit (1952) (vg. § 2.3.2). D'altra banda, atès l'any d'aparició de la traducció espanyola, el 1951, no sembla probable que el 1969 hagi de ser la data de la primera publicació en francès, sinó, versemblantment, d'alguna edició posterior.

<sup>70</sup> La denominació *omplidor* l'he adoptada per al que se sol anomenar *mot-crossa* —*muletilla* en castellà i *filler* en anglès. En el punt 5.6.1 pot veure's una presentació d'aquesta unitat i uns comentaris sobre les denominacions que té.

“Linguistic skill and linguistic complexity are constructs that are not objective facts, but rather require interpretation and judgements. (...) Linguistic complexity, although used as an indicator of language skill or of higher levels of linguistic development, is a construct which is not yet well defined. Sentence-level indicators, such as use of subordinate clauses, are not adequate as measures of linguistic complexity. A conjunction such as *because* contributes not only to the structure of sentences, but also to the construction of discourse. (...) Characteristics of linguistic complexity as a measure of language skill should not be focused at the sentence level, but should account for complexity at the discourse level as well, also considering cultural and situational contexts of language use.” (Schleppegrell, 1992: 126-129)

Per tant, una primera conclusió que cal incorporar és que l'estudi de la complexitat lingüística s'ha d'abordar també des de la perspectiva del discurs. D'altra banda, un altre aspecte fonamental de la complexitat, ja apuntat per M.A.K. Halliday i per altres autors citats anteriorment és que es pot manifestar de maneres diferents i en dimensions gramaticals i discursives diferents. Aquesta és la tesi fonamental de l'article de D. Biber (1992) titulat precisament “On the Complexity of Discourse Complexity: A Multidimensional Analysis”. Vegem-ne uns fragments ben representatius:

“The study shows that discourse complexity is a multidimensional construct, that different types of structural elaboration reflect different discourse functions, and that different kinds of texts are complex in different ways (in addition to being more or less complex). (...) Discourse complexity is not a single, unified construct (...). Given this multidimensional nature of discourse complexity, a complete description of a text (or register) cannot focus exclusively on the degree of complexity; rather (...) texts are complex in different ways as well as to differing extents.” (Biber, 1992: 133-149)

D. Biber (1992) parteix per al seu experiment d'unes llistes dels trets associats a la simplicitat o a la complexitat lingüística. Els criteris per a la classificació d'aquests trets no hi són definits sinó que es donen com a sobreentesos i sembla que dependrien bastant dels “tòpics” que denuncia A. Narbona (1995). Per tant, la revisió que demanava M.J. Schleppegrell (1992) de la noció de complexitat és ben necessària. Encara que negar radicalment la relació entre subordinació i complexitat —com fa K. Beaman (1984) en la citació que encapçala aquest punt— sigui potser exagerat, la noció de complexitat lingüística mereixeria una definició acurada que no sembla que estiguem en condicions

d'ofrir en l'estat actual de la recerca. Per fer-ho, caldria saber quins serien els trets lingüístics constituents d'aquesta complexitat, més enllà de la subordinació que se cita en tots els treballs. Caldria veure si es tractaria d'una complexitat sintàctica o també —i més aviat— d'una complexitat discursiva, cognitiva, pragmàtica, prosòdica, etc. Caldria, a més, veure el paper que podrien tenir en l'establiment del concepte els estudis de tipus psicolingüístic, que avaluarien la facilitat o la dificultat amb què els usuaris produeixen i interpreten unes determinades construccions oracionals o discursives<sup>71</sup>. En resum, caldria recerca en moltes direccions.

Pel que fa a aquesta tesi i a la qüestió central que la motiva, és a dir, la contradicció flagrant entre els diversos autors en tractar sobre la complexitat lingüística i, especialment, sintàctica, hi ha d'haver diverses explicacions que actuïn alhora. Algunes d'aquestes explicacions poden ser les que exposo a continuació:

1- Bona part dels resultats contradictoris s'ha d'haver produït per les diferències entre els textos i els corpus analitzats en cada estudi, de manera que les especificitats de gènere, de grau de formalitat, de funció, etc., poden haver provocat un biaix que hagi allunyat els resultats dels diversos investigadors.

2- Sembla que una altra part del problema pot raure en les múltiples direccions de la complexitat i en el fet que es manifesti de manera diferent segons cada tret lingüístic (construccions nominals i verbals; sintagmes preposicionals precedits d'una preposició o d'una altra; tipus de coordinades, de subordinades i d'interordinades, etc.).

---

<sup>71</sup> K. Perera (1986), en parlar de les diferències oral / escrit, afirma que aquest darrer conté més expressions indicatives de "maduresa gramatical":

"From a grammatical point of view, writing can be differentiated from spontaneous speech in three ways. First, there are some structures that regularly occur in speech which are not normally used by competent writers (...). Secondly, writing contains a higher proportion than speech of constructions widely held to be indicative of grammatical maturity." (Perera, 1986: 498)

Caldria veure si la noció de maduresa gramatical, que només es pot definir des d'un punt de vista psicolingüístic, coincidiria amb la de complexitat, és a dir, si a un adult li costa més de produir — tant en temps de processament com en esforç cognitiu— allò que ha après més tard en la seva evolució lingüística. D'altra banda, caldrien més estudis de psicologia evolutiva centrats no només en construccions sintàctiques isolades sinó en les habilitats de construcció del discurs.

3- En molts casos, i també com a enfocament general, pot ser més rellevant determinar la mena de complexitat que no pas veure'n el grau —de fet, arribaria a ser absurd d'anar buscant quin mode o quin gènere és *més complex*. Les preguntes clau hauran de ser, doncs, *De quina manera és complex?* o *Com s'hi manifesta la complexitat lingüística?*.

4- Algunes de les afirmacions fetes pels tractadistes podrien adquirir un altre caire si s'amplia el marc de referència al text o discurs.

5- Si, d'acord amb els punts anteriors, s'analitzen els diversos trets lingüístics en cadascun dels apartats i subapartats, el més probable és que, com ja he dit, de les dues posicions enfrontades cadascuna tingui una part de raó —seria estrany que tants lingüistes prestigiosos, demostradament experts en el seu camp, anessin errats en totes les seves observacions.

6- Hi pot haver, encara, una darrera raó, que no he trobat ressenyada enlloc, però que pot tenir-hi un paper destacat. En l'afirmació de M.A.K. Halliday, que tant sorprèn, en el sentit que la llengua oral és més complexa gramaticalment que la llengua escrita, és clau preguntar-se què entén aquest autor per *gramàtica*. Efectivament, la gramàtica sistemicofuncional, de la qual M.A.K. Halliday és un dels principals teòrics, postula l'existència de tres nivells d'explicació gramatical del llenguatge, que donen lloc a tres estructures superposades per a cada enunciat: el nivell ideatiu (que inclou nocions semàntiques com *actor*, *meta*, *sensor*, *fenomen*, *portador*, *atribut*, etc.), el nivell interpersonal (on es manegen unitats més pròpiament sintàctiques com *subjecte*, *mode*, *predicat*, *complements*, *adjunts*, etc.) i el nivell textual (que distingeix entre *tema* i *rema*). Moltes de les unitats d'aquests tres nivells són, per tant, elements de tipus semàntic i pragmàtic que no solen incloure's en gramàtiques elaborades des d'altres punts de vista. Així, quan M.A.K. Halliday parla de «complexitat gramatical» o d'«estructura complexa de l'oració» o «de la clàusula» s'està referint a un conjunt de trets molt més ampli del que generalment s'entén per sintaxi o per gramàtica<sup>72</sup>.

---

<sup>72</sup> La gramàtica sistemicofuncional, poc difosa entre nosaltres, es pot trobar desenvolupada a M.A.K. Halliday (1973 i 1985a). [També pot veure-se'n una brevíssima síntesi a J.M. Castellà (1992: 209-216).]

7- Caldrà mantenir una reserva teòrica global sobre la idea de complexitat lingüística ja que, com he argumentat anteriorment, aquest concepte demanaria una definició i unes concrecions de les quals no podem disposar en l'estat actual de la recerca en els diversos camps que hi estan relacionats.

Aquests set punts, que serveixen de resum final per al marc teòric, han de ser la guia d'actuació al llarg de la recerca, en la qual intentaré d'aclarir alguns aspectes de la complexitat lingüística en tres gèneres orals i escrits en llengua catalana.



### 3. METODOLOGIA

“For speech is so much more than words and sentences.”

John Steinbeck, *Travels with Charley in Search of America*, 1962<sup>73</sup>.

#### 3.1- OBJECTIUS I JUSTIFICACIÓ DE LA RECERCA

##### 3.1.1- ÀMBIT DE LA RECERCA

Aquesta tesi s’inscriu en l’*estudi de l’ús del llenguatge* i, dins d’aquest, en l’anomenada *lingüística social* o *microsociolingüística*<sup>74</sup>, que s’ocupa de la descripció de la *variació funcional*:

“La variació funcional —que també ha estat denominada *contextual* o *estilística*— té a veure amb les diferències que presenta una llengua segons els diversos àmbits d’ús en què es fa servir, o sigui, segons els diferents propòsits o funcions, interlocutors, canals i temes propis dels actes comunicatius en què apareix el llenguatge verbal humà. Qualsevol llengua varia en funció de factors com els esmentats, fins i tot aquelles que no tenen forma escrita: en aquest cas la variació pot ser menor, però mai no serà menyspreable. Ara bé, en les (mal) anomenades «llengües de cultura», la variació és més visible donat que aquestes

---

<sup>73</sup> Fragment citat per K. Beaman (1984: 45).

<sup>74</sup> La *microsociolingüística* té en compte, segons H. Calsamiglia i A. Tusón (1991: 22-23), “(...) tots els elements lingüístics i extralingüístics (tant els cognitius com els situacionals) que contribueixen a *crear significació*”, en contrast amb la *macrosociolingüística*, en la qual “la llengua es considera com una variable entre altres presents en la vida social i, com a tal, subjecta a normes d’ús, explícites o implícites. S’analitza l’ús de la llengua per àmbits, les funcions que les llengües compleixen i també les actituds dels parlants envers les llengües, des dels criteris de prestigi, poder i identitat”. Es tracta de la diferència, en altres paraules, que hi ha entre la *lingüística social* i la *sociologia de la llengua* (Castellà, 1992: 116-117).

llengües es manifesten a través d'un conjunt de varietats funcionals —els registres— més ampli i complex.” (Payrató, 1998a: 5)

Més en concret, la present recerca forma part dels estudis dedicats a la caracterització del *mode del registre*, en la línia que s'inicià amb M.A.K. Halliday i que més recentment s'ha concretat en els treballs de D. Biber i altres autors anglòfons esmentats en el capítol anterior. És una investigació de tipus *microscòpic*, en la mesura que s'hi analitzen només tres gèneres i un grup limitat de trets lingüístics que, encara que sigui prou nombrós, no té la pretensió d'abraçar el conjunt de les relacions lèxiques, sintàctiques i discursives. Una opció diferent és la recerca de tipus *macroscòpic*<sup>75</sup>. La definició que fa D. Biber (1988) de l'una i de l'altra és la següent:

“Within the broad framework of investigation into the psychological and sociological underpinnings of linguistic variation, researchers have investigated textual variation through macroscopic and microscopic analyses. Macroscopic analysis attempts to define the overall dimensions of variation in a language. Microscopic analysis, on the other hand, provides a detailed description of the communicative functions of particular linguistic features. (...) Microscopic text analysis is necessary to pinpoint the exact communicative functions of individual linguistic features. It complements macroscopic analysis in two ways: (1) it identifies the potentially important linguistic features and genre distinctions to be included in a macro-analysis, and (2) it provides detailed interpretation of the textual dimension in functional terms.” (Biber, 1988: 61-62)

En confluència amb la perspectiva bàsicament variacionista, utilitzaré, lògicament, conceptes i enfocaments provinents de diverses disciplines lingüístiques, com la *lexicologia*, la *sintaxi*, la *pragmàtica*, la *lingüística del text*, l'*anàlisi del discurs*<sup>76</sup> i, fins i tot, la *lingüística cognitiva*.

---

<sup>75</sup> Els treballs de tipus macroscòpic solen exigir la tasca coordinada d'un equip d'investigació o d'un institut de recerca i un tractament informatitzat de grans volums de dades a partir de corpus preferiblement preconstituïts.

<sup>76</sup> La distinció entre aquestes tres darreres disciplines és confusa i a vegades depèn de les maneres d'autodenominar-se que tenen els diversos autors i tendències. Per exemple, partint d'una definició àmplia d'*anàlisi del discurs*, aquesta tesi també podria incloure-s'hi.

### 3.1.2- NECESSITAT I ORIGINALITAT DE LA RECERCA

Diversos autors han expressat la conveniència de l'estudi de la variació funcional i del registre col·loquial en llengua catalana, com per exemple F. Vallverdú (1973), I. Marí (1983 i 1994), V. Salvador (1984), L. Payrató (1988 i 1992), A. Viana (1990) i B. Montoya (1992). Un dels primers a fer-ho va ser L.V. Aracil (1978), en els termes següents:

“Hom diria que, per alguna raó, ha prevalgut pertot una preferència immemorial i inconfessada pel discurs escrit formal —el més allunyat de fet, del «no-marcad». Cal concloure que els prejudicis anticol·loquials han estat dominants pertot. I en el cas de la Gramàtica Prescriptiva catalana, obsedida i fascinada per la «Llengua literària», hi ha hagut virtualment una compacta conspiració anti-col·loquial. (...) En aquestes circumstàncies, vindicar i dignificar el col·loquial, prendre'l seriosament i fer-li justícia, fins i tot el fet mateix de contrastar-lo i reconèixer-lo objectivament, és un acte revolucionari —i un bell punt de partida de la recerca estilística en aquest racó de món.” (Aracil, 1978: 239)

Vint anys després, malgrat alguns esforços ben meritoris però poc nombrosos, L. Payrató (1998a) descriu una situació no gaire diferent, encara molt deficitària en estudis sobre la variació funcional:

“La llengua catalana, que ha estat estudiada d'una manera força aprofundida pel que fa a la variació històrica i geogràfica, no ha estat objecte en canvi de gaires estudis de caire funcional. Per comprovar-ho, només cal repassar les gramàtiques històriques i els atles i monografies dialectals amb què comptem, i comparar-los —quantitativament— amb els estudis sobre els registres catalans. I si el que demanéssim, més enllà de monografies sobre un registre concret, fossin obres de contrast entre varietats funcionals, el resultat s'acostaria molt al buit absolut. (...) De fet, les «caselles buides» de la sociolingüística catalana es fan ben paleses en el vessant de les varietats funcionals, així com en el de les socials (o socioculturals).” (Payrató, 1998a: 5-6)

El meu treball s'emmarca en la tasca d'omplir una d'aquestes *caselles buides* de la sociolingüística catalana. Es tracta, doncs, d'una recerca de *caràcter bàsic*, d'aquelles de què una llengua ha de disposar per aspirar a tenir una descripció sociolingüística mínimament completa. En concret, una recerca de les característiques d'aquesta sobre el mode del registre és, per la informació de què dispo, la primera vegada que es fa en

català<sup>77</sup>. De fet, en la bibliografia consultada, un estudi quantitatiu del contrast oral / escrit en un conjunt variat de trets i relativament ampli de textos no l'he trobat tampoc sobre la llengua castellana, francesa ni italiana. En anglès, en canvi, hi ha hagut, com ja he mostrat en el capítol precedent i com es veurà en els posteriors, una línia intensa i extensa de treball en aquest camp<sup>78</sup>.

Més enllà de la novetat que representa a casa nostra, des de la perspectiva internacional la meua contribució haurà de ser necessàriament més discreta, atès el nombre important de publicacions existents sobre el tema en anglès que acabo d'assenyalar. Així, respecte de l'estudi macroscòpic de D. Biber (1988), que és per la seva amplitud el principal punt de referència, les meves aportacions —modestes, però no per això menys interessants— són les següents:

a- Definir els gèneres a estudiar d'una manera més acurada (vg. les observacions als gèneres del corpus de D. Biber en el punt 3.4.1).

b- Comparar de manera especial, amb més control de variables i amb els mateixos individus, la classe magistral i la prosa acadèmica, sense deixar mai de treballar amb textos produïts en condicions reals i no de laboratori (vg. els comentaris a recerques anteriors en el punt 3.4.1).

---

<sup>77</sup> Cal fer justícia, però, als treballs inclosos en el volum col·lectiu *Oralment*, a cura de L. Payrató (1998a) —aparegut durant el procés d'elaboració d'aquesta tesi—, els quals, tot i ser d'un abast més restringit, tant per la mostra emprada com pel nombre de trets estudiats, constitueixen inequívocament una incursió de pes en aquest àmbit.

<sup>78</sup> En castellà hi ha els treballs quantitatius de J. de Kock, C. Gómez Molina i R. Verdonk (1992), que es refereixen només a la llengua escrita i, com a molt, a l'oral teatral, i els de L. Cortés (1986), que analitza exclusivament la llengua oral. Hi ha, també, el de S. Pons (1998), de caràcter quantitatiu i qualitatiu sobre els connectors textuais en espanyol, que es basa exclusivament en el corpus de conversa del grup *Val.Es.Co* (vg. Briz, 1995). Altres autors, com A. Briz, A. Narbona o A.M. Vígara, fan estudis centrats en l'oral, però no quantitatius. De la mateixa manera, C. Blanche-Benveniste ha desenvolupat una destacada línia d'estudis sobre el francès col·loquial, que també ressenyo, i que tampoc no aborden en cap cas la perspectiva quantitativa.

c- Incorporar a la llista trets de tipus discursiu —en concret, els connectors textuais (marcadors o connectors parentètics i interjeccions)—, no tractats per la majoria d'autors d'aquest camp.

d- Intentar de donar una resposta, des de la llengua catalana i amb les limitacions d'una recerca individual, a la divergència entre els diversos autors sobre la complexitat lingüística (lèxica i sintàctica) en la llengua oral i en la llengua escrita.

### **3.1.3- OBJECTIUS I PREGUNTA DE RECERCA**

D'acord amb els aspectes teòrics presentats en el capítol anterior, amb la necessitat palesada de disposar en català d'una recerca descriptiva mínimament extensa en aquest camp, i amb l'interès indiscutible que té, per a la lingüística de la variació en general, de cercar respostes explicatives a la contradicció generada al voltant de la complexitat lingüística en el mode del registre, he establert per a aquesta tesi la pregunta i els objectius de recerca que especifico a continuació.

#### **A- Pregunta de recerca**

Com es diferencien entre si i quines menes de complexitat lingüística manifesten els gèneres conversa informal, classe magistral i prosa acadèmica pel que fa als trets lingüístics associats amb la densitat lèxica, la composició oracional i la connexió textual?

#### **B- Objectiu general**

Efectuar una anàlisi comparativa de tres gèneres de discurs en llengua catalana (conversa informal, classe magistral i prosa acadèmica) en tres aspectes de l'estructuració i de la complexitat lingüística: densitat lèxica, composició oracional i connexió textual.

### **C- Objectius específics**

1- Comparar i caracteritzar els tres gèneres estudiats pel que fa als trets lingüístics relacionats amb la densitat lèxica.

2- Comparar i caracteritzar els tres gèneres estudiats pel que fa als trets lingüístics relacionats amb els estils nominal i verbal de construcció dels enunciats.

3- Analitzar i determinar el grau i el tipus d'intervenció de les diverses formes de composició oracional en cada gènere. Especialment, resoldre, per a la llengua catalana i en els gèneres estudiats, la contradicció entre els analistes que atorguen a la llengua escrita un grau superior de complexitat sintàctica —i, sobretot, una tendència més gran a la subordinació com a recurs per a la interrelació dels enunciats—, i els que consideren que és en la llengua oral on es desenvolupen al màxim les potencialitats gramaticals de la llengua.

4- Analitzar i determinar el grau i el tipus d'intervenció de les diverses formes de connexió textual en l'estructuració discursiva a cada gènere.

5- Comparar, d'una manera especial, la conversa informal i la prosa acadèmica, com a gèneres polars de l'estudi i com a representants més o menys arquetípics de la llengua oral i de la llengua escrita, respectivament.

6- Comparar, també especialment, la classe magistral i la prosa acadèmica, com a gèneres amb més control de variables en la present investigació, cercant-hi possibles diferències més afinades entre la llengua oral i la llengua escrita.

### **D- Objectius secundaris o consecutius**

1- Experimentar i resoldre els problemes metodològics associats a una recerca que treballa sobre un corpus de llengua oral i escrita, i que combina el tractament informàtic amb la codificació manual de les dades.

- 2- Contribuir a la caracterització lingüística del gènere conversa informal.
- 3- Contribuir a la caracterització del gènere classe magistral i, en general, del discurs oral formal expositiu.
- 4- Contribuir a la caracterització del gènere prosa acadèmica.
- 5- Aportar dades per a la descripció del discurs acadèmic oral i escrit.
- 6- Aportar dades per a la descripció dels tipus textuais conversacional i expositiu.
- 7- Col·laborar en l'estudi d'algunes unitats lingüístiques, especialment de l'àmbit discursiu o textual, que resulten de difícil definició i que estan en una fase incipient de delimitació (connectors textuais, interjeccions, omplidors, etc.).
- 8- Contribuir a l'anàlisi i a la descripció del registre col·loquial en llengua catalana.

### **3.1.4- HIPÒTESIS DE RECERCA**

Per elaborar les hipòtesis de recerca, formulades amb anterioritat al gruix del treball quantitatiu de la present recerca, he partit de tres fonts. En primer lloc, hi ha les que són congruents amb alguns dels principis establerts i/o discutits en el marc teòric. En segon lloc, les que es desprenen d'estudis anteriors sobre l'anglès, especialment aquelles en què hi ha unanimitat —o quasiunanimitat— entre els autors. I, en tercer lloc, les que es deriven dels resultats del meu treball de doctorat (Castellà, 1996c), que ja estava orientat a permetre l'establiment d'hipòtesis per a un estudi posterior. De totes maneres, en el cas d'aquesta darrera font, he hagut de reelaborar la llista perquè hi contrastava només dos gèneres, la classe magistral i la prosa acadèmica, i aquí hi he incorporat la conversa informal.

He agrupat les hipòtesis en dos tipus: les més *generals*, que es refereixen a les tendències globals de la complexitat lingüística en els tres gèneres, i les *particulars*, que es formulen sobre el comportament de trets lingüístics concrets o de grups de trets.

### **A- Hipòtesis generals**

1- Els gèneres conversa informal, classe magistral i prosa acadèmica presentaran importants diferències entre si pel que fa a la densitat lèxica, la composició oracional i la connexió textual.

2- Els gèneres conversa informal i prosa acadèmica, com a representants arquetípics dels modes oral i escrit se situaran en les posicions extremes o polars de la mostra.

3- El gènere classe magistral, atès que combina característiques contextuals del mode oral i de l'escrit, se situarà en una posició intermèdia entre els altres dos gèneres.

4- La complexitat lingüística es manifestarà en direccions diferents a tots tres gèneres, de manera que cap no resultarà globalment més complex que un altre.

5- La prosa acadèmica serà més densa lèxicament que la classe magistral i aquesta ho serà més que la conversa informal.

6- La conversa informal tindrà un estil més verbal de construcció dels enunciats, la prosa acadèmica tindrà un estil més nominal i la classe magistral ocuparà una posició intermèdia.

7- Els tres gèneres presentaran diferències en les formes de composició oracional i en la preferència a l'hora de triar cada construcció sintàctica. Aquestes preferències es produiran en direccions diverses, de manera que caldrà detallar el comportament de cada gènere respecte de cada construcció.

8- Els dos gèneres orals presentaran més ocurrencies dels connectors textuais. La prosa acadèmica, tot i tenir-ne menys en nombre, serà més variada en el repertori.



## **B- Hipòtesis particulars**

1- La prosa acadèmica tindrà un nombre major de noms i un nombre menor de verbs respecte a la classe magistral, i aquesta respecte a la conversa informal.

2- Els verbs copulatius seran més habituals en els dos gèneres orals.

3- Els noms, adjectius i verbs de freqüència alta apareixeran més en el mode oral que en l'escrit.

4- El nombre total d'adjectius serà superior en la prosa acadèmica, tot i que aquesta tendirà a tenir més adjectius atributius i la conversa informal més adjectius predicatius.

5- Els adverbis en *-ment* seran més habituals en la classe magistral i en la prosa acadèmica.

6- La prosa acadèmica serà més variada lèxicament, mentre que la classe magistral i, sobretot, la conversa informal, ho seran menys.

7- La longitud de mot serà superior en la prosa acadèmica.

8- La prosa acadèmica presentarà una quantitat més elevada de sintagmes preposicionals complements del nom, especialment els introduïts per la preposició *de*.

9- Les nominalitzacions seran més habituals en la prosa acadèmica que en la classe magistral, i més habituals en la classe magistral que en la conversa informal.

10- Els dos gèneres orals presentaran un nombre més elevat de clàusules subordinades substantives que la prosa acadèmica. Aquesta, per contra, tindrà més clàusules subordinades adjectives. En xifres relatives al total de clàusules, els resultats seran similars en els tres gèneres.

11- En els dos gèners orals hi haurà més coordinació clausal i en l'escrit més coordinació sintagmàtica. Tant en els uns com en l'altre, la coordinació clausal serà un recurs quantitativament molt inferior a la subordinació.

12- La major part dels nexes d'interordinació apareixeran en un nombre lleugerament superior en els gèneres orals que en l'escrit.

13- Els connectors textuais de tota mena (conjuncions, marcadors parentètics i interjeccions) presentaran més ocurrences en els dos gèneres orals que en l'escrit.

14- Els gèneres orals presentaran una tendència general a unir les clàusules entre si per mitjans explícits de connexió, mentre que la prosa acadèmica farà relativament menys ús dels elements de connexió (nexes de connexió oracional i connectors textuais).

## 3.2- LA VARIABLE DEPENDENT

### 3.2.1- LES VARIABLES DE LA RECERCA

En qualsevol estudi que impliqui, en tot o en part, procediments de tipus quantitatiu i, en general, sempre que es vulgui assegurar la validesa interna d'una recerca, és important de determinar amb exactitud les variables que es manejaran. D'acord amb E. Hatch i H. Farhady (1982) i M.T. Turell (1995)<sup>79</sup>, les definicions de les variables que intervenen en el present treball són les següents:

**A/ Variable independent:** és aquella que exerceix la seva influència sobre les variables que són l'objecte d'estudi:

*“The independent variable is the major variable which you hope to investigate. It is the variable which is selected, manipulated, and measured by the researcher.”* (Hatch i Farhady, 1982: 15)

“Les variables independents són els factors i contextos de diversos tipus amb els quals es correlaciona la variable dependent per poder establir el marc favorable o desfavorable d'aplicació d'una determinada regla variable, o de realització d'una determinada variant.” (Turell, 1995: 25)

**B/ Variable dependent:** és la variable que evoluciona segons els canvis que s'operen en les variables independents:

*“The dependent variable, on the other hand, is the variable which you observe and measure to determine the effect of the independent variable.”* (Hatch i Farhady, 1982: 15)

---

<sup>79</sup> Més enllà de les publicacions, les explicacions de M. Teresa Turell a l'assignatura “Metodologia de la recerca en lingüística”, dins del programa de doctorat *Lingüística Aplicada (lèxic i discurs)*, i la seva col·laboració personal amb una lectura crítica de l'esborrany, han resultat essencials per a la redacció del present capítol.

“La variable dependent és la dada lingüística per excel·lència; és l’objecte d’estudi de la lingüística de la variació, ja que és en l’ús on s’observa una variació prou important per ser analitzada. És la variable que es correlaciona amb els factors i contextos que poden afavorir la seva realització en una, dues, tres o més formes; formes que, d’altra banda, seran anomenades *variants* de la variable.” (Turell, 1995: 24)

**C/ Variable de control:** és la que es manté inalterable per assegurar la validesa de l’experiment:

“A control variable is a variable which is held constant in order to neutralize the potential effect it might have on behaviour. (...) Notice that you also limit the generalizability of your research by using control variables.” (Hatch i Farhady, 1982: 16)

En aquest estudi, la variable dependent és el *gènere discursiu*, amb tres possibles realitzacions (*conversa informal*, *classe magistral* i *prosa acadèmica*) que definiré en els punts següents. Les variables independents, que són la llista de trets lingüístics a mesurar, i les variables de control, que són diverses i demanen una discussió específica, les presentaré en apartats posteriors.

La recerca que desenvolupo es caracteritza per un nombre alt de trets lingüístics i un nombre mínim de gèneres, amb tota una sèrie de paràmetres contextuals que han quedat situats com a variables de control. La consideració dels trets lingüístics com a elements caracteritzadors dels gèneres, que són la variable central que es vol estudiar, dóna a aquesta recerca un caràcter *horitzontal*: és d’un tipus diferent de les més habituals en lingüística de la variació, que estudien pocs o només un tret lingüístic situat en la condició de variable dependent, amb diversos elements contextuals o independents que el condicionen (són exemples d’aquestes darreres M.T. Turell, 1988; N. Alturo i M.T. Turell, 1990 i N. Alturo, 1998). És, també, una recerca que practica l’anàlisi quantitativa —que inclou, necessàriament, una part qualitativa— en trets lingüístics de tipus lèxic, sintàctic i discursiu<sup>80</sup>.

---

<sup>80</sup> D. Schiffrin (1985b) defensa l’anàlisi quantitativa aplicada a trets sintàctics i discursius:

“Quantitative methods of analysis are often used by sociolinguists to answer parallel questions about constraints on phonological or morphological variation (...). However, such methods are not

### 3.2.2- EL GÈNERE DISCURSIU

El mot *gènere*, procedent dels estudis literaris, adoptat i cada vegada més utilitzat en sociolingüística i anàlisi del discurs, ha estat central en la present recerca<sup>81</sup>. Aquest concepte, que, lògicament, ha ampliat l'abast en incorporar-se a les disciplines de l'ús lingüístic, és de difícil definició perquè se superposa amb altres nocions que s'hi relacionen —o que hi competeixen—, com *registre*, *tipus de text* i *estil*, tal com assenyala D. Biber (1994):

“The term *register* has been used in many different ways, as have the related terms *genre*, *text type*, and *style*. Most researchers agree in using *register* to refer to situationally defined varieties, as opposed to *dialect*, which refers to varieties associated with different speaker groups. For example, Ure (1982: 5) states that the “register range of a language comprises the range of social situations recognized and controlled by its speakers —situations for which appropriate patterns are available.” Beyond this general association of register with situational variation, however, there is less consensus.” (Biber, 1994: 51)<sup>82</sup>

---

easily transferred to syntactic or discourse domains (...). In this paper, I demonstrate that quantitative analyses are nevertheless useful procedures through which to isolate constraints from different levels of discourse, and through which to tease apart the ways in which structure, meaning, and social action differently influence the production of discourse —even though they are interrelated in the process of so doing.” (Schiffrin, 1985b: 282)

Crec que el present treball és una demostració i una posada en pràctica de l'afirmació de D. Schiffrin en aquest paràgraf.

<sup>81</sup> L'anàlisi i el contrast de diversos gèneres és propugnat per C. Blanche-Benveniste (1997a) amb els arguments següents:

“J'ai parlé des variations syntaxiques en les ramenant à deux grands types, variants utilisées pour classer les locuteurs et variants de genre, qu'on peut s'attendre à trouver chez tous les locuteurs. (...) Les analyses sur des corpus diversifiés montrent que les locuteurs peuvent manier une grande dose de variation, interne à leurs productions, selon les genres dans lesquels ils se situent. Selon qu'ils parlent en conversation privée ou en situation professionnelle, ils utiliseront ou non des nominalisations ou des types de prépositions. Cette dimension n'est pas prise en compte par des corpus fondées sur une seule situation d'entretien.” (Blanche-Benveniste, 1997a: 27-28)

<sup>82</sup> De fet, en els orígens del concepte, el terme *registre* abraça qualsevol aspecte de la forma dels textos condicionat pel context i, per tant, inclou les nocions de *gènere* i de *tipus de text* (vg. § 2.2.1). Podem veure-ho a M.A.K. Halliday (1978) i a M.A.K. Halliday, A. McIntosh i P. Strevens (1964), segons els quals el gènere forma part del mode del registre, i aquest, al seu torn, es relaciona íntimament amb l'idiolecte:

Efectivament, l'ús que es fa d'aquests termes depèn molt de cada disciplina, de cada autor i, fins i tot, de diferents èpoques en un mateix autor. Així, per exemple, el mateix D. Biber usa generalment el mot *gènere* en els treballs anteriors a 1993, per designar els mateixos objectes (p. ex., les subdivisions del *LOB Corpus* i del *London-Lund Corpus* dels seus treballs de 1986 i 1988) que a partir d'aquesta data anomena *registres* (vg. D. Biber, 1993, 1995a i 1995b). C.A. Ferguson (1994), en un intent de clarificació del panorama, proposa les definicions següents:

**A. [Dialect]** *A group that operates regularly in a society as a functional element (e.g., in terms of physical location, marriage patterns, or economic, religious, or other interactional behavior) will tend to develop identifying markers of language structure and language use, different from the language of other social groups.*

**B. [Register]** *A communication situation that recurs regularly in a society (in terms of participants, setting, communicative functions, and so forth) will tend over time to develop identifying markers of language structure and language use, different from the language of other communication situations.*

**C. [Genre]** *A message type that recurs regularly in a community (in terms of semantic content, participants, occasions of use, and so on) will tend over time to develop an identifying internal structure, differentiated from other message types in the repertoire of the community.”* (Ferguson, 1994: 18-21)

Per tant, el registre es caracteritza més per ser un conjunt de trets lingüístics (“*identifying markers of language structure and language use*”) i el gènere, en canvi, per ser un tipus de missatge recurrent i tipificat. Al costat d'això, la noció de *tipus de text*, que C.A. Ferguson no tracta, és de caràcter més general, referida a operacions cognitives i a estructures abstractes (narració, exposició, argumentació, etc.), que s'apliquen en gèneres i en textos

---

“El concepto de género analizado con anterioridad es un aspecto de lo que nosotros llamamos «modo». Los diversos géneros del discurso, incluso los géneros literarios, son las funciones semióticas específicas del texto que poseen valor social en la cultura.” (Halliday, 1978: 189)

“Individual style, however, is linked to register. It is the writer's idiolect, especially the grammar and lexis of the idiolect, in a given register. Insofar as 'style' implies literary style, register here means literary, including poetic, genre and medium. Style is thus linguistic form in interrelation with literary form.” (Halliday, McIntosh i Strevens, 1964: 158)

concrets. La definició que en fa D. Biber —amb més convenciment, com es pot veure, per als tipus de text que per als gèneres— es basa en la distinció *intern / extern*:

“The term *genre* is loosely defined as «text categorizations made by the basis of external criteria relating to author / speaker purpose» (Biber 1988: 68) and «the text categories readily distinguished by mature speakers of a language; for example... novels, newspaper articles, editorials, academic articles, public speeches, radio broadcasts, and everyday conversations. These categories are defined primarily on the basis of external format» (Biber 1989: 5-6). (...) In contrast, the term *text type* has been used in multidimensional analyses to refer to text categories defined in strictly linguistic terms.” (Biber, 1994: 52)<sup>83</sup>

Finalment, *registre*, *gènere* i *tipus de text* són delimitats entre si per L. Payrató (1998a) en una proposta més global, que té l’originalitat de jugar amb tots tres conceptes, tot redefinint-los i adaptant-los:

“(...) el registre és una modalitat lingüística (és a dir, de *la llengua*) enquadrada en unes coordenades que es refereixen als *usos*: és la llengua *situada*, en els seus contextos *situacionals*. Els *tipus de text*, en canvi (...) poden entendre’s com a abstraccions o categoritzacions que tenen en compte exclusivament factors lingüístics, verbals, o, si es prefereixen d’altres paraules, criteris *interns*, intrínsecs al text, no dependents del seu entorn sociocultural. (...) Els *gèneres*, per últim, (...) representen abstraccions o categoritzacions ja pròpiament socioculturals: són classes de *fets de parla* (o *fets comunicatius*) típics d’una comunitat, és a dir, manifestacions comunicatives socioculturals on la parla és essencial, resulta definidora de l’activitat i està modelada per la tradició. Cada comunitat disposa d’un conjunt prototípic de gèneres, coneguts (de forma activa i/o passiva) pels seus parlants, i de fet les denominacions que se solen utilitzar per etiquetar-los són les mateixes que les emprades en el llenguatge corrent (...).” (Payrató, 1998a: 27-28)<sup>84</sup>

---

<sup>83</sup> Aquest autor no és sempre consistent amb la distinció que proposa: només cal veure la citació (D. Biber, 1986: 386, § 3.4.1), en què parla de *text type* on hauria d'utilitzar, segons el seu propi criteri, la paraula *genre*.

<sup>84</sup> Una distinció molt semblant entre aquests tres conceptes és la de J.M. Castellà (1995 i 1996b):

“. **Gènere:** forma discursiva estereotipada, definida principalment per les seves característiques externes, de caràcter social i cultural. Els gèneres són reconeguts pels parlants segons el seu format extern i segons els paràmetres situacionals en què se solen produir. Constitueixen una llista oberta i es poden classificar per àmbits d'ús. Són exemples de gèneres la classe, la conferència, el debat, la notícia, l'article d'opinió, la novel·la, el conte, l'examen, la tesi, etc.

. **Tipus de text:** forma discursiva estereotipada, definida per les seves característiques internes, de caràcter lingüístic i informacional. Els tipus de text són descrits pels lingüistes segons la seva estructura

Els conceptes de gènere i de tipus de text són presentats d'una manera similar per J.-M. Adam (1992) i C. Blanche-Benveniste en diversos dels seus treballs. J.-M. Adam (1992: 20-35), en concret, proposa la *seqüència textual dominant*, noció que dona compte del fet que els tipus de text no es troben en estat pur sinó combinats entre si, de manera que en els discursos reals hi sol haver seqüències prototípiques diverses, sovint organitzades per —o supeditades a— una seqüència que adopta una posició dominant (vg., també, Castellà, 1994).<sup>85</sup>

---

informativa i segons els trets gramaticals que els configuren. Constitueixen una llista tancada i tenen un caràcter més abstracte que els gèneres. Són exemples de tipus de text la descripció, la narració, l'argumentació, les instruccions, etc.

. **Registre:** conjunt de característiques lingüístiques que apareixen com a manifestació de la relació dels textos amb la situació comunicativa. Els trets lingüístics dels registres pertanyen al nivell fonològic, al morfològic, al sintàctic i, sobretot, al lèxic. Segons el registre, els textos s'adeqüen al tema, al grau de formalitat, al canal i a l'objectiu de cada comunicació. Cada text que es produeix constitueix una combinació concreta de trets del registre. A vegades pot ser pràctic utilitzar denominacions com *registre col·loquial*, *administratiu*, *científic*, etc., però sovint la col·locació d'aquestes etiquetes és una simplificació excessiva de les característiques del text en qüestió." (Castellà, 1995: 74)

<sup>85</sup> Una definició de gènere que ha estat molt difosa és la de J.M. Swales (1990), que segueix, però, una línia diferent, ja que el relaciona amb els *communicative events* i els *communicative purposes* i, per tant, no fa més que substituir *registre* per *gènere*:

"A genre comprises a class of communicative events, the members of which share some set of communicative purposes. These purposes are recognized by the expert members of the parent discourse community, and thereby constitute the rationale for the genre. This rationale shapes the schematic structure of the discourse and influences and constrains choice of content and style." (Swales, 1990: 58)

J. Pujolar (1999), per la seva banda, repassa diverses definicions de gènere i en planteja una visió relativista i, sobretot, no taxonòmica, que comparteixo especialment:

"Crec que és possible que l'estudi dels gèneres es mostri més productiu des del punt de vista de la investigació dels processos socials de comunicació que no pas amb la simple elaboració de tipologies o inventaris de gèneres. Les formes de treballar arquetípiques de l'estructuralisme han afavorit tradicionalment el treball taxonòmic, fins al punt que alguns autors semblen confondre una taxonomia per una teoria. És de preveure, però, que l'elaboració de tipologies generals de gèneres verbals ha de ser una tasca excessivament àrdua donat el caràcter fluid d'aquests mateixos en la forma en què apareixen en la vida ordinària. Això no vol dir que no sigui possible l'establiment d'unes categories canòniques per a ús en àmbits pedagògics o d'un altre tipus, com de fet ja passa actualment. Una altra cosa és que aquestes categories siguin analíticament productives en d'altres àmbits. Per contra, estudiar les condicions de «fluïdesa» dels gèneres promet ser una tasca enormement suggerent i susceptible de trobar-hi multitud d'aplicacions." (Pujolar, 1999: 151-152)



### 3.2.3- LA CONVERSA INFORMAL

El primer gènere de l'estudi és la conversa informal. Es tracta de la forma més bàsica i, sens dubte, més freqüent de comunicació humana, possiblement la manifestació del llenguatge per excel·lència —i, curiosament, com ja s'ha dit, una de les menys estudiades. L'he anomenada *informal* — i podria haver dit *quotidiana, natural, espontània, col·loquial*, etc.— per diferenciar-la d'altres interaccions o formes dialogades que també podrien anomenar-se *converses*, però que impliquen uns rols i uns estatus socials que les fan menys lliures o més subjectes a regles relacionades amb la distància i la formalitat (p. ex., conversa cap-subordinat, metge-pacient, informador-informat, etc.). Les converses que analitzaré hauran de ser *informals, no preparades, entre familiars o amics, de tema o temes improvisats i amb una distribució lliure de torns de parla*.

El gènere conversa informal és, dels tres d'aquesta tesi, el que ha estat definit amb més precisió. D'entre els autors consultats (Edmonson, 1981: 6; Viana, 1986: 93-94; Tusón, 1988: 140 i 1995: 54-56; Cots, Nussbaum, Payrató i Tusón, 1990: 59-60; Bobes, 1992: 41; Vigara, 1992: 39-43 i Goodwin, 1994: 217), destacaré la definició —o, més aviat, la caracterització— esdevinguda ja clàssica, de H. Sacks, E. Schegloff i G. Jefferson (1974):

“In any conversation, we observe the following:

- (1) Speaker-change recurs, or at least occurs.
- (2) Overwhelmingly, one party talks at a time.
- (3) Occurrences of more than one speaker at a time are common, but brief.
- (4) Transitions (from one turn to a next) with no gap and no overlap are common. Together with transitions characterized by slight gap or slight overlap, they make up the vast majority of transitions.
- (5) Turn order is not fixed, but varies.
- (6) Turn size is not fixed, but varies.
- (7) Length of conversation is not specified in advance.
- (8) What parties say is not specified in advance.
- (9) Relative distribution of turns is not specified in advance.
- (10) Number of parties can vary.
- (11) Talk can be continuous or discontinuous.

(12) Turn-allocation techniques are obviously used. A current speaker may select a next speaker (as when he addresses a question to another party); or parties may self-select in starting to talk.

(13) Various ‘turn-constructural units’ are employed; e.g., turns can be projectedly ‘one word long’, or they can be sentential in length.

(14) Repair mechanisms exist for dealing with turn-taking errors and violations; e.g., if two parties find themselves talking at the same time, one of them will stop prematurely, thus repairing the trouble.” (Sacks, Schegloff i Jefferson, 1974: 700-701)

Per tant, els torns de parla, les convencions que en regulen la successió i, bàsicament, la llibertat amb què s’organitzen en la conversa quotidiana són les característiques fonamentals d’aquest gènere discursiu. A. Tusón (1995), en un comentari de la llista de trets que acabem de veure, destaca el fet que, malgrat aquesta gran variabilitat, les converses no resulten pas caòtiques sinó ben organitzades:

“Veiem, doncs, que una de les característiques fonamentals de la conversa espontània és la seva «indeterminació». Ni el tema, ni el nombre de participants, ni la distribució dels torns de paraula, ni la duració de cada torn o de la mateixa conversa en conjunt s’especifiquen prèviament. (...) Tanmateix, els mateixos autors apunten també, en aquesta llista de trets, l’aspecte organitzat i sistemàtic de les converses espontànies, perquè el que resulta altament interessant és que, a partir d’unes coordenades indefinides i canviant, que podrien produir simplement *caos*, existeixen tècniques per a la distribució de torns, a penes es produeixen solapaments i, quan apareix un problema, els parlants despleguen mecanismes per reparar-lo.” (Tusón, 1995: 55-56)

Des de la perspectiva de les modalitats de llengua, la conversa informal està associada al *registre col·loquial*, que L. Payrató tracta en diverses de les seves obres (1988, 1992 i 1994) i que defineix de la manera següent:

“El registre col·loquial català és la modalitat de la llengua catalana que utilitzen els parlants en l’àmbit d’ús col·loquial o quotidià (...). Malgrat que sigui irreductible en últim terme a la categorització, la col·loquialitat ideal tendiria a la tria següent:

- 1) CAMP:    quotidianitat
  - 2) MODE:    oral espontani
  - 3) TENOR:   interactiu
  - 4) TO:       informal”
- (Payrató, 1988: 49-50)

El fet que les converses informals puguin tenir un nombre variable d'interlocutors i que aquests puguin allargar més o menys el torn, pot tenir conseqüències importants per a l'arquitectura sintàctica i discursiva resultant. Hi ha converses —o fragments de conversa— que tenen els torns de parla més curts, amb més mobilitat i més facilitat per als trencaments i les superposicions (són així, principalment, les converses 2 i 3 de la mostra). N'hi ha d'altres, en canvi, en què un dels interlocutors pren la iniciativa de parlar d'un tema determinat, explicar una història, etc. i els altres es limiten a anar-hi intercalant preguntes, assentiments o petits comentaris (hi ha fragments així a les converses 4 i 5, i sobretot a la conversa 1, la primera meitat de la qual és protagonitzada per un dels parlants que s'explica llargament sobre un tema). Si els torns de parla es fessin més i més extensos, ens podríem trobar ja fora de l'àmbit de la conversa espontània, en la tertúlia organitzada, la taula rodona o la conferència. La diferència entre els torns de parla segons la duració és apuntada per G. Brown i G. Yule (1983b) en distingir entre *short turns* i *long turns*, aspecte que es veurà en el punt següent.

### 3.2.4- LA CLASSE MAGISTRAL

Aquest és un gènere menys tractat que la conversa i de definició més difícil. De fet, la classe magistral comparteix moltes característiques amb la conferència, en la mesura que tots dos són discursos *orals formals, de tema més o menys especialitzat i amb un grau important de planificació o preparació prèvia*. Vegem, a continuació, la coneguda definició de *lecture* d'E. Goffman (1981)<sup>86</sup>:

“A lecture is an institutionalized extended holding of the floor in which one speaker imparts his views on a subject, these thoughts comprising what can be called his «text». The style is typically serious and slightly impersonal, the controlling intent being to generate calmly considered understanding, not mere entertainment, emotional impact, or immediate action. Constituent statements presumably take their warrant from their role in

---

<sup>86</sup> De fet, com encertadament em fa notar M.J. Cuenca (comunicació personal), el mot *lecture* es refereix a la conferència però també a la classe magistral, especialment formal, que es dona dins de cursos universitaris (un professor titular d'universitat és un *lecturer*). Aquesta ambivalència del mot anglès va encara més a favor, doncs, de la interrelació entre conferència i classe magistral.

attesting to the truth, truth appearing as something to be cultivated and developed from a distance, coolly, as an end in itself.” (Goffman, 1981: 165)

Ja des d’una perspectiva d’anàlisi del discurs pedagògic, A. Cros (1995 i 1999) elabora una definició de classe magistral que remet i, alhora, es distancia de la conferència i hi afegeix els aspectes institucional i didàctic:

“A partir de les consideracions dels estudis pedagògics i de les aportacions de Goffman, entendrem la *classe magistral* com un gènere discursiu que es produeix en el marc de la institució acadèmica, especialment la universitària, la qual atorga una autoritat a l’enunciador, considerat expert, que el situa en un estatus superior al del destinatari, permet que gestioni el discurs, que imposi unes normes als estudiants i que orienti les seves interpretacions en una direcció determinada. Malgrat que es tracti d’un discurs monogestionat, a diferència d’altres gèneres, com ara la conferència, els estudiants poden participar de maneres diferents durant la classe i poden interrompre les explicacions dels ensenyants per formular preguntes o fer comentaris. Igualment, els ensenyants reorienten el seu discurs en funció del *feed-back* que reben dels estudiants.” (Cros 1999: 66)

Per tant, la classe magistral comporta una situació d’enunciació que combina la presa de possessió del torn de parla i l’autoritat pròpies de la conferència amb la funció didàctica adreçada a uns estudiants i, sobretot, amb la ubicació dins d’una institució acadèmica (amb uns programes, unes taxes, unes titulacions, etc.). Aquest gènere m’ha interessat, doncs, perquè constitueix una forma de llengua oral generalment monologada —o, almenys, amb seqüències monològiques bastant llargues— ben diferenciada socioculturalment de la conversa quotidiana —i cal pensar que també lingüísticament, aspecte que és més l’objecte d’aquesta tesi. El monòleg és la realització màxima del *long turn* de G. Brown i G. Yule (1983b) que he esmentat en el punt anterior, en què el parlant pren la paraula i assumeix la responsabilitat de construir tot sol el discurs:

“A short turn consists of only one or two utterances, a long turn consists of a string of utterances which may last as long as an hour’s lecture. (...) As soon as a speaker ‘takes the floor’ for a long turn, tells an anecdote, tells a joke, explains how something works, justifies a position, describes an individual, and so on, he takes responsibility for creating a structured sequence of utterances which must help the listener to create a coherent mental representation of what he is trying to say. (...) The ability to construct such long turns appears to vary with individuals, in part, no doubt, depending on the opportunity they have had to produce long

turns which other people bother to listen to. The ability to produce long transactional turns, in which clear information is transferred, is, we claim, not an ability which is automatically acquired by all native speakers of a language. It is an ability that appears to need adequate models, adequate practice and feed-back.” (Brown i Yule, 1983b: 16-19)<sup>87</sup>

Aquesta habilitat per produir torns de parla llargs de tipus transaccional té a veure també amb el coneixement i la pràctica que l'individu té de l'escriptura. La classe magistral és un gènere prototípicament monologat que queda situat entre la llengua oral col·loquial i la llengua escrita, en una posició intermèdia que la fa especialment adequada per al present estudi. La classe magistral, com la conferència i altres gèneres orals, forma part d'una franja de registres que en un altre lloc he denominat *llengua oral no col·loquial* (Castellà, 1996a). La caracterització que feia d'aquesta franja, en relació amb els dos pols del mode del registre, és la següent:

**“1- Coincidències que sembla mantenir l'oral no col·loquial amb l'escrit:**

1.1- En el pla de la coherència textual, l'adopció d'unes formes d'ordenació i d'estructuració de la informació ben distants de la conversa quotidiana (especialment en els discursos expositius de caràcter monologat).

1.2- En el pla de la cohesió textual, la utilització de connectors i organitzadors textuais que reflecteixen l'ordenació particular esmentada al punt 1.1 (*com hem dit abans, en el punt anterior, que esmentarem a continuació*, etc.) comuns amb l'escrit però molt rars en l'oral col·loquial.

1.3- En el pla sintàctic, la voluntat del parlant d'utilitzar al màxim possible una sintaxi sistematitzada a l'estil de l'escrit. (...)

---

<sup>87</sup> Una explicació semblant d'aquest fenomen, des del punt de vista de la sintaxi utilitzada, és la que fa C. Blanche-Benveniste (1995b):

“Queríamos obtener enunciados pronunciados por diferentes tipos de hablantes, con una sintaxis bastante compleja. Para ello necesitábamos turnos de habla largos, sin las interrupciones incesantes de las conversaciones. Intentamos identificar cuáles eran las situaciones que favorecían estas características. Resultó que los ejemplos de sintaxis más complejos eran proporcionados en situaciones de «explicaciones», y en particular cuando los hablantes aceptaban responder a preguntas sobre su profesión, o cuando eran llevados a hablar públicamente de su profesión. Estos turnos de habla suelen ser muy diferentes de lo que son las conversaciones de los mismos hablantes. En efecto, al habla profesional se transmite todo un saber lingüístico que estaría fuera de lugar en las charlas entre amigos.” (Blanche-Benveniste, 1995b: 155)

## 2- Coincidències que sembla mantenir l'oral no col·loquial amb el col·loquial:

2.1- En el pla de la coherència textual, la necessitat d'uns mecanismes de repetició de la informació, de conducció de l'oient, etc., que tenen a veure amb la producció i la recepció del discurs en temps real.

2.2- En el pla de la cohesió textual, l'aparició dels anomenats marcadors del discurs (Schiffrin 1987) (*bé, vull dir, ara, perquè, saps?*), en part comuns amb el col·loquial però rars en l'escrit.

2.3- En el pla sintàctic, tot i la voluntat indicada a 1.3 d'incorporar una sintaxi regular, l'existència d'anacoluts, vacil·lacions, repeticions, mots crossa, etc. que en realitat el distancien de l'escrit.” (Castellà, 1996a: 268-269)

D'acord amb aquesta anàlisi, la classe magistral i els registres orals no col·loquials en general comparteixen característiques de la llengua oral i de la llengua escrita prototípiques. Algunes d'aquestes característiques he pogut estudiar-les amb detall en la present recerca.

### 3.2.5- LA PROSA ACADÈMICA

En el cas de la prosa acadèmica, fer-ne una delimitació com a gènere és encara més complex. De fet, aquesta denominació no es correspon pròpiament amb un únic gènere, sinó que n'hi ha diversos de l'àmbit acadèmic que podrien englobar-s'hi: articles i capítols de llibre científics i de divulgació científica, llibres de text, projectes acadèmics, treballs de recerca, tesis, etc.<sup>88</sup> Aquesta multiplicitat o ventall d'opcions queda reflectida en la definició que A. Cros (1999) fa del discurs acadèmic i, també, en la que L. Trimble (1985) fa de l'*EST (English for Science and Technology)*:

“Els termes *discurs acadèmic* o *discurs de vulgarització científica* són utilitzats per diferents disciplines per designar el conjunt de discursos orals i escrits produïts en àmbits relacionats amb l'ensenyament, amb un objectiu que, d'una manera general, es podria denominar *didàctic* i on la relació que s'estableix entre els participants pot ser d'expert a expert o de més

---

<sup>88</sup> També es poden incloure dins de la prosa acadèmica els textos produïts pels alumnes —apunts, exàmens, treballs, etc. (vg., per exemple, el tractament que en fan P. Battaner, E. Atienza, C. López i M. Pujol, 1997: 11-12)—, entre els quals, de fet, caldria classificar la tesi, per més que sigui el treball més llarg i més especialitzat que un alumne pugui presentar.

a menys expert. Es tracta de *discursos segons*, en el sentit que són el resultat de l'elaboració i la transformació d'altres discursos produïts anteriorment.” (Cros, 1999: 61)

“(…) it (EST) covers that area of written English that extends from the ‘peer’ writing of scientists and technically oriented professionals to the writing aimed at skilled technicians. (...) EST covers the areas of English written for academic and professional purposes and of English written for occupational (and vocational) purposes, including the often informally written discourse found in trade journals and in scientific and technical materials written for the layman.” (Trimble, 1985: 5-6)

L. Trimble (1985) assenyala que hi ha una àrea de superposició entre l'anglès per a propòsits científics i tècnics i l'anglès per a finalitats acadèmiques. Aquesta relació entre ciència i pedagogia és expressada també per A. Ali Bouacha (1981):

“Si l'on considère l'université comme le lieu privilégié de la transformation des énoncés scientifiques en énoncés pédagogiques et le cours universitaire comme un «macro-acte de parole» qui prend en charge cette transformation, le discours pédagogique peut apparaître comme le simulacre sans cesse répété du discours de la découverte. Le professeur-énonciateur construit son discours de manière à apparaître comme un chercheur objectif s'effaçant devant une vérité qui se proclame d'elle-même, par la voix de la science, à l'intention d'un auditoire qui, au-delà du groupe d'étudiants présents, s'élargit à l'ensemble de la communauté scientifique.” (Ali Bouacha, 1981: 41)

Per tant, els textos que he seleccionat sota el rètol *prosa acadèmica* són relativament pròxims a les classes magistrals, tant per la intenció didàctica com pel grau d'especialització. Per a aquest segon aspecte, m'ha calgut buscar textos publicats en llibres d'un grau d'especialització mitjà, més pròxims a la divulgació periodística de la ciència que a l'article científic adreçat a especialistes que només es difon en àmbits reduïts. Així, els escrits acadèmics que analitzaré es diferencien de les classes magistrals sobretot per la dimensió oral / escrit o, per ser més exactes, per l'oposició acústic / gràfic.

En la prosa acadèmica la qüestió de la denominació dels gèneres es mostra, doncs, amb tota la problemàtica. Els textos escrits que analitzo en aquest treball no tenen una designació genèrica establerta socialment. Són *articles i capítols de llibre de tema científic, dins de l'àmbit de les ciències socials, d'un grau d'especialització mitjà, escrits amb una intenció divulgativa i/o didàctica però sense arribar a ser de to periodístic*, però, naturalment,

aquesta manera d'anomenar-los resultaria impracticable. Per aquesta raó purament pràctica he triat el rètol *prosa acadèmica* —també podia haver estat *prosa expositiva*—, tot i que sóc conscient que aquesta denominació no és del tot exacta<sup>89</sup>.

Una altra qüestió important és la mesura en què la prosa acadèmica es pot considerar representativa de la llengua escrita en general, ja que en això no hi ha un acord total entre els especialistes. Així, per exemple, W. Chafe (1992) accepta la conversa com a gènere prototípic de la llengua oral, però qüestiona que existeixi una forma prototípica de llengua escrita:

“I assume that ordinary casual conversation has a special position as the prototypical use of spoken language. Speech is of course used in other ways—in interviewing, lecturing, orating, and so on— but conversation presumably is and always has been the most common use of speaking, and hence the most common use of language overall. (...) It is interesting that written language seems not to offer anything comparable as a prototype. The special circumstances of writing allow a writer to produce languages of quite varied kinds, no one of which has any privileged status with relation to the rest. This fact has led to much confusion in comparisons of spoken and written language. Although one may feel secure in taking conversation to be the spoken basis for comparison, there is no particular version of written language with which one can compare it.” (Chafe, 1992: 19-24)

Hi ha altres autors, en canvi, que defensen la prototipicitat tant de la conversa informal com de la prosa expositiva:

“Most research on spoken versus written language has analyzed casual conversation as spoken language and expository prose, or what Olson (1977) calls ‘essayist literacy’, as written. This is not by chance. There is something typically spoken about face-to-face conversation and something typically written about expository prose.” (Tannen, 1985: 124)

“In terms of its situational characteristics, stereotypical speech is interactive, and dependent on shared space, time, and background knowledge; stereotypical writing has the opposite characteristics. In terms of its linguistic characteristics, stereotypical speech is structurally simple, fragmented, concrete, and dependent on exophoric (situation-dependent) reference;

---

<sup>89</sup> Dit d'una altra manera, hi ha moltes altres menes de textos que també són prosa acadèmica i que no s'estudien en aquesta tesi. Alhora, però, és ben veritat que els textos que s'inclouen aquí formen part indiscutible de la prosa acadèmica.



again, stereotypical writing has the opposite characteristics. (...) Both perspectives indicate face-to-face conversation as typical speech and informational exposition as typical writing.” (Biber, 1988: 37)

“Prendrem com a punts de referència la conversa quotidiana, d’un costat, i la prosa expositiva, de l’altre, com les manifestacions més habituals i típiques, respectivament, de l’oralitat i de l’escriptura.” (Tusón, 1995: 15)

“Specifically, expository registers seem to be the kind of writing that develops to maximally exploit the production potential of the written mode, apparently in response to the highly informational communicative purposes.” (Biber, 1995b: 9)

Així, en la mateixa línia que D. Tannen (1985) i A. Tusón (1995), D. Biber (1988) considera l’exposició informativa com a *typical writing* i D. Biber (1995b), encara amb més èmfasi, caracteritza els registres expositius com la manifestació màxima de la potencialitat de la llengua escrita. Per part meva, crec que cal ser molt caut en aquesta qüestió. Com comentaré en els apartats següents, penso que el present estudi no podrà ser representatiu ni extrapolable a tota la llengua catalana escrita. Però, malgrat això, en la mesura que els textos de la mostra són mostres inequívokes de prosa acadèmica expositiva, podrem estar segurs que reflecteixen alguns dels trets essencials de l’escriptura.

### **3.3- LES VARIABLES INDEPENDENTS**

#### **3.3.1- VARIABLES INDEPENDENTS DE L'ESTUDI**

La llista de trets lingüístics o variables independents, com és habitual en els treballs de tipus general sobre el contrast entre llengua oral i llengua escrita, és bastant llarga. Els he seleccionat tenint com a referent els trets manejats pels autors que han argumentat sobre la complexitat lingüística, relatius als tres grans blocs que ja s'anuncien en el títol del treball: trets lexicogramaticals o de densitat lèxica (i, també, de definició de l'estil nominal i verbal), trets sintàctics o de composició oracional i trets cohesius o de connexió textual. No hi ha dubte que podia haver-ne investigat moltíssims més, de cara a disposar d'una descripció més detallada de l'oposició oral / escrit en llengua catalana. Però, obligat com estava a limitar-ne el nombre, he triat els que, en treballs anteriors, han resultat més decisius en el debat obert sobre la complexitat lingüística, tot afegint als sintàctics els de tipus textual —i d'aquests, els més relacionats amb la composició oracional, és a dir, els connectors textuais. La llista completa definitiva ha estat la següent:

#### **1/ TRETS LEXICOGRAMATICALS (DENSITAT LÈXICA)**

##### **A/ Noms**

- 1- Noms en general
- 2- Noms de freqüència alta
- 3- Pronoms personals forts
- 4- Demostratiu neutre *això*

##### **B/ Verbs**

- 5- Verbs en general
- 6- Verbs de freqüència alta
- 7- Verbs en forma personal

- 8- Verbs en infinitiu
- 9- Verbs en gerundi
- 10- Perífrasis verbals
- 11- Verbs copulatius
- 12- Verbs predicatius
- 13- Nombre de mots per clàusula

### **C/ Adjectius**

- 14- Adjectius atributius
- 15- Adjectius predicatius
- 16- Adjectius en general
- 17- Adjectius de freqüència alta

### **D/ Adverbis**

- 18- Adverbis en *-ment*

### **E/ Proporció mots lexicals i mots gramaticals**

- 19- Mots lexicals
- 20- Mots gramaticals

### **F/ Sintagmes preposicionals**

- 21- Sintagmes preposicionals (amb les preposicions febles)
- 22- Sintagmes preposicionals complements del nom
- 23- Sintagmes preposicionals complements de l'adjectiu
- 24- Sintagmes preposicionals complements del verb
- 25- Altres usos preposicionals
- 26- Noms complementats per un sintagma preposicional
- 27- Nombre de sintagmes preposicionals per clàusula
- 28- Adjectius-pronoms demostratius i indefinits *aquest/a/s/es, un/a/s/es* i *algun/a/s/es*

### **G/ Nominalitzacions**

- 29- Nominalitzacions (noms derivats d'adjectius i noms derivats de verbs)

### **H/ Variació lèxica, repetició lèxica i longitud de mot**

- 30- Variació lèxica
- 31- Repetició lèxica
- 32- Longitud de mot

## **I/ Estil nominal i estil verbal**

33- Total estil nominal

34- Total estil verbal

## **2/ TRETOS SINTÀCTICS (COMPOSICIÓ ORACIONAL)**

### **J/ Nexes i construccions de subordinació**

35- Nexes i construccions de subordinació substantiva

36- Nexes i construccions de subordinació adjectiva

37- Nexes i construccions de subordinació adverbial

38- Nexes i construccions de comparació clausal

39- Nexes i construccions de consecució descontínua

40- Subordinades complement nominal

41- Subordinades complement verbal

### **K/ Nexes de coordinació**

42- Nexes de coordinació sintagmàtica

43- Nexes de coordinació clausal

### **L/ Nexes i construccions d'interordinació**

44- Nexes i construccions condicionals

45- Nexes i construccions concessives

46- Nexes i construccions causals

47- Nexes i construccions consecutives

48- Nexes i construccions adversatives

49- Nexes i construccions finals

## **3/ TRETOS COHESIUS (CONNEXIÓ TEXTUAL)**

### **M/ Connectors textuais**

50- Conjuncions en usos textuais

51- Marcadors o connectors parentètics en usos oracionals

52- Marcadors o connectors parentètics en usos textuais

### **N/ Interjeccions**

53- Interjeccions amb valor connectiu

54- Inicials de torn de parla

- 55- Medials de torn de parla
- 56- Finals de torn de parla
- 57- Absolutes (constituents de torn de parla)

### **O/ Omplidors o mots-crossa**

- 58- Nombre d'omplidors o mots-crossa

Resulten, doncs, un total de 58 trets agrupats en 15 apartats, els quals es distribueixen, al seu torn, en els tres grans blocs. Dins de cadascun dels trets o variables, encara es podrien enumerar les diverses realitzacions lingüístiques de cada un (p. ex., cadascuna de les cinc preposicions febles, cadascun dels nexes de subordinació, de coordinació i d'interordinació, cadascun dels nombrosos marcadors discursius, etc.). En els apartats corresponents de la tesi detallaré aquestes realitzacions i en donaré els resultats d'aparició en els tres gèneres estudiats.

### **3.3.2- TRETOS LINGÜÍSTICS NO INCLOSOS EN L'ESTUDI**

Després d'oferir la llista de trets o variables independents que he seleccionat per analitzar en els textos de la mostra, crec que també és important que especifiqui què *no* tracta el present treball.

Com és habitual en estudis com aquest —i possiblement en qualsevol tipus d'estudi individual—, m'he hagut de limitar rigorosament en l'anàlisi<sup>90</sup>. Així, hi ha un bon nombre de trets que no han pogut formar part de l'anàlisi. Per repassar-los procediré acarant la llista anterior amb la de D. Biber (1988), que té l'interès de ser la més completa que he trobat en la bibliografia sobre el mode del registre. Aquesta llista abraça un nombre (67 trets en 16 apartats) i una varietat de trets més gran que la meua, d'acord amb l'objectiu de l'autor de

---

<sup>90</sup> Com comentaré més endavant en parlar de les dimensions de la mostra, m'he vist molt condicionat per les limitacions pròpies d'una tesi: és una tasca individual, està —sortosament— delimitada en el temps i, per tant, s'ha de definir acuradament perquè no sigui infinita. Així, algunes de les qüestions que aquí he abordat només en la mesura de les meves possibilitats, serien matèria de treball més apta per a un equip integrat per diverses persones dins d'un projecte d'investigació d'objectius més ambiciosos, a l'empar d'una institució acadèmica de recerca que li donés la cobertura adequada.

descriure la variació lingüística en anglès d'una manera global. Per facilitar-ne la comparació, he assenyalat amb vermell els trets en què hi ha hagut una coincidència plena, amb verd aquells en què la coincidència ha estat parcial i he deixat sense destacar els que no he tractat en la recerca:

“Table 4.4 Features used in the analysis

A/ Tense and aspect markers

- 1- Past tense
- 2- Perfect aspect
- 3- Present tense

B/ Place and time adverbials

- 4- Place adverbials (e.g., *above, beside, outdoors*)
- 5- Time adverbials (e.g., *early, instantly, soon*)

C/ Pronouns and pro-verbs

- 6. First person pronouns
- 7. Second person pronouns
- 8- Third person personal pronouns (excluding *it*)
- 9- Pronoun *it*
- 10- Demonstrative pronouns (*that, this, these, those* as pronouns)
- 11- Indefinite pronouns (e.g., *anybody, nothing, someone*)
- 12- Pro-verb *do*

D/ Questions

- 13- Direct WH-questions

E/ Nominal forms

- 14- Nominalizations
- 15- Gerunds
- 16- Total other nouns

F/ Passives

- 17- Agentless passives
- 18- *By*-passives

G/ Stative forms

- 19- *Be* as main verb
- 20- Existential *there*

## H/ Subordination features

- 21- *That* verb complements (e.g., *I said that he went*)
- 22- *That* adjective complements (e.g., *I'm glad that you like it*)
- 23- WH clauses (e.g., *I believed what he told me*)
- 24- Infinitives
- 25- Present participial clauses (e.g., *Stuffing his mouth with cookies, Joe ran out the door*)
- 26- Past participial clauses (e.g., *Built in a single week, the house would stand for fifty years*)
- 27- Past participial WHIZ deletion relatives (e.g., *the solution produced by this process*)
- 28- Present participial WHIZ deletion relatives (e.g., *the event causing this decline is...*)
- 29- *That* relative clauses on subject position (e.g., *the dog that bit me*)
- 30- *That* relative clauses on object position (e.g., *the dog that I saw*)
- 31- WH relatives on subject position (e.g., *the man who likes popcorn*)
- 32- WH relatives on object position (e.g., *the man who Sally likes*)
- 33- Pied-piping relative clauses (e.g., *the manner in which he was told*)
- 34- Sentence relatives (e.g., *Bob likes fried mangoes, which is the most disgusting thing I've ever heard of*)
- 35- Causative adverbial subordinators (*because*)
- 36- Conditional adverbial subordinators (*although, though*)
- 37- Conditional adverbial subordinators (*if, unless*)
- 38- Other adverbial subordinators (e.g., *since, while, whereas*)

## I/ Prepositional phrases, adjectives, and adverbs

- 39- Total prepositional phrases
- 40- Attributive adjectives (e.g., *the big horse*)
- 41- Predicative adjectives (e.g., *the horse is big*)
- 42- Total adverbs

## J/ Lexical specificity

- 43- Type / token ratio
- 44- Mean word length

## K/ Lexical classes

- 45- Conjuncts (e.g., *consequently, furthermore, however*)
- 46- Downtoners (e.g., *barely, nearly, slightly*)
- 47- Hedges (e.g., *at about, something like, almost*)
- 48- Amplifiers (e.g., *absolutely, extremely, perfectly*)
- 49- Emphatics (e.g., *a lot, for sure, really*)
- 50- Discourse particles (e.g., sentence initial *well, now, anyway*)
- 51- Demonstratives

## L/ Modals

- 52- Possibility modals (*can, may, might, could*)
- 53- Necessity modals (*ought, should, must*)
- 54- Predictive modals (*will, would, shall*)

## M/ Specialized verb classes

- 55- Public verbs (e.g., *assert, declare, mention, say*)
- 56- Private verbs (e.g., *assume, believe, doubt, know*)
- 57- Suasive verbs (e.g., *command, insist, propose*)
- 58- *Seem* and *appear*

## N/ Reduced forms and dispreferred structure

- 59- Contractions
- 60- Subordinator *that* deletion (e.g., *I think [that] he went*)
- 61- Stranded prepositions (e.g., *the candidate that I was thinking of*)
- 62- Split infinitives (e.g., *he wants to convincingly prove that...*)
- 63- Split auxiliaries (e.g., *they are objectively shown to...*)

## O/ Coordination

- 64- Phrasal coordination
- 65- Independent clause coordination (clause initial *and*)

## P/ Negation

- 66- Synthetic negation (e.g., *no answer is good enough for Jones*)
- 67- Analytic negation (e.g., *that's not likely*)” (Biber, 1988: 73-75)

Com es pot veure, dels 67 trets de D. Biber (1988) he coincidit en 28 de forma plena i en 15 de forma parcial. Dels altres, n'hi ha alguns que són específics de l'anglès (p. ex. 28- *Present participial WHIZ deletion relatives*, 54- *Predictive modals*, 60- *Subordinator that deletion*, 61- *Stranded prepositions* o 66- *Synthetic negation*) i n'hi ha força que no he pres en consideració perquè no estan inclosos en els tres blocs objectiu de la tesi.

D'altra banda, hi ha tot el grup de trets que tracto en el tercer bloc (Connexió textual), que en la llista de D. Biber estan poc o gens representats i que només hi apareixen parcialment en apartats com *K- Lexical classes*. De fet, en els treballs sobre el contrast oral / escrit és estrany de trobar-hi trets de tipus textual, i és en aquest punt on crec que estic fent una aportació més singular en aquest camp de recerca<sup>91</sup>. Malgrat això, dins dels aspectes de

---

<sup>91</sup> D'entre el nombre considerable d'estudis que he manejat sobre el contrast oral / escrit, només n'he trobat dos sobre aspectes netament textuals. El primer és el de D. Biber i B. Altenberg (1992) sobre la referència oracional i textual (si bé presenta un problema metodològic important que comentaré en el punt 3.6.2). El segon, molt més ampli en el nombre de trets considerats, és el de G.B. Sherman (1993) sobre la cohesió textual, tractada en els seus diversos apartats segons la coneguda subdivisió de M.A.K. Halliday i R. Hasan (1976): referència, substitució, el·lipsi, conjunció i cohesió lèxica.



tipus textual m'he limitat als elements de connexió i he hagut d'excloure un bon nombre de trets lingüístics rellevants en la construcció del discurs. Així, no he pogut tractar aspectes tan destacats i tan atractius com:

- Els elements de referència textual (anàfora, catàfora, dixi<sup>92</sup>).
- L'elisió sintàctica per motius discursius.
- Els fenòmens de cohesió lèxica.
- La progressió temàtica (desenvolupament tema / rema).
- En la llengua oral, les vacil·lacions, les reformulacions, els anacoluts, etc.
- En la llengua oral, l'entonació<sup>93</sup>.

---

<sup>92</sup> D'aquest aspecte, només he comptabilitzat les ocurrències totals del demostratiu neutre *això*, dels pronoms personals forts i de l'adjectiu-pronom *aquest/a/s/es*, sense distingir entre els usos dítctics o no dítctics (vg. § 3.6.2, 4.2.3 i 4.7.3).

<sup>93</sup> Sóc ben conscient de l'extraordinària importància de la prosòdia en la llengua oral i, alhora, de l'enorme dificultat de tractar-la, en primer lloc per la manca d'instruments teòrics i descriptius en català sobre aquesta matèria —de fet, n'hi ha en poques llengües. Potser arribarà un dia en què serà impossible, en què serà considerat una incorrecció tractar sobre la llengua oral sense incorporar-hi un aspecte tan fonamental.

## **3.4- LES VARIABLES DE CONTROL**

### **3.4.1- VALIDESA INTERNA I EXTERNA: REPRESENTATIVITAT DE LA MOSTRA**

Les variables de control, que ja he definit anteriorment, són un factor fonamental per assegurar la validesa de l'experiment que es realitza. En el cas del contrast oral / escrit, aquest ha estat assenyalat com un punt crucial per diversos autors: els resultats de les recerques que s'han fet han estat contradictoris entre si i el motiu principal pot haver estat la disparitat de criteris en la constitució dels corpus de cada investigació. Vegem, per exemple, el comentari que en fan K. Beaman (1984) i D. Biber (1986a), el segon en un article titulat significativament "Spoken and Written Textual Dimensions in English: Resolving the Contradictory Findings":

"The differences between spoken and written language that many linguists in the past have found may be more a result of differences in the formality and purpose of the discourse than in a spoken / written distinction." (Beaman, 1984: 46)

"It is evident that the relationship between speech and writing is highly complex; however, the methodologies of most quantitative studies have been restricted, and have not provided an adequate basis for global conclusions concerning this relationship. The restrictions include the following:

(a) Assigning undue weight to particular linguistic features: Most quantitative studies have used between three and ten features, and have considered a differential distribution in any of them to be important evidence.

(b) Assigning undue weight to individual texts: Most studies have been based on few texts, so that idiosyncratic differences can have a major influence.

(c) Assigning undue weight to the texts types chosen for analysis: Most studies have compared two text types, one spoken (e.g., conversation) and one written (e.g., academic prose), and have not controlled for the different communicative tasks represented by those text types.

In spite of such restrictions, earlier studies have in fact attempted to reach global conclusions —which have thus proved contradictory.” (Biber, 1986a: 386)

Efectivament, és un defecte important de qualsevol treball d’investigació intentar d’extreure conclusions generals a partir de casos massa particulars. Cal dissenyar la recerca i oferir-ne les conclusions de manera que es respectin els principis de validesa interna i externa:

“The internal validity of a research study is the extent to which the outcome is a function of the factor you have selected rather than other factors you haven’t controlled. (...) External validity refers to the extent that the outcome of any research study would apply to other similar situations in the real world.” (Hatch i Farhady, 1982: 7-8)

En la bibliografia que he manejat s’observen diferents maneres d’actuar respecte d’aquest principi. D’entrada, hi ha un primer grup d’autors que opten pel procediment més tradicional i, de fet, menys coherent, que és l’anàlisi de la llengua oral o bé de determinades unitats discursives, a partir de textos literaris, sovint fragments de diàleg novel·lesc o teatral. Aquest és el cas de A.M. Vígara (1980), J.-M. Adam (1985 i 1992), J. Deulofeu (1988), G. Herrero (1988), H. Mederos (1988), M.C. Bobes (1992), J. de Kock, C. Gómez i R. Verdonk (1992), J. Portolés (1998) i M.A. Martín Zorraquino i J. Portolés (1999)<sup>94</sup>. Cal destacar que el treball quantitatiu de J. de Kock, C. Gómez i R. Verdonk (1992) sobre la llengua castellana, que és una obra d’especial envergadura, recolza enterament en textos teatrals per a les seves conclusions sobre la llengua oral. Hi ha, també, el cas particular de G. Herrero (1988), que utilitza exclusivament textos de còmics, sense dubte d’un caràcter més oralitzant que els diàlegs d’altres gèneres literaris, però també sens dubte diferents de la conversa quotidiana.

Un segon grup d’autors, alguns dels quals ja he esmentat entre els anteriors, fan seva la pràctica habitual en lingüística estructural i generativa de crear els exemples *ad hoc*. Entre aquests es poden esmentar J. Deulofeu (1988), J. Portolés (1988 i 1998), D.J. Allerton

---

<sup>94</sup> Per a l’estudi de determinades unitats lingüístiques com, per exemple, els connectors textuais, no crec que basar-se exclusivament en textos escrits porti a resultats erronis sinó, més aviat, incomplets, ja que la llengua oral pot fer uns usos particulars d’aquestes unitats.

(1991) i J.-M. Adam (1992). Hi ha també el cas de J. Gumperz, H. Kaltman i M.C. O'Connor (1984), que disposen de textos reals de conversa, però elaboren ells mateixos els fragments de llengua escrita que hi comparen, cosa que a vegades també fa M.A.K. Halliday (1985b). B. Montoya (1992) s'ocupa de fer la crítica d'aquestes maneres de procedir, tant de la d'aquest grup com de la de l'anterior:

“Els lingüistes han estat, normalment, uns «hòmens de lletres». Confosos amb els estudiosos de la literatura sota l'etiqueta de «filòlegs», s'han desviat molt sovint del seu objectiu final: la llengua parlada. (...) Certament, hi ha unes especialitats lingüístiques que s'ocupen directament d'observar i recollir mostres de l'oralitat quotidiana. Però parlem de dos camps d'estudi relativament marginals: la dialectologia i la sociolingüística. Aquesta marginalitat, la confirmen els tractats gramaticals a l'ús, que ignoren completament els resultats d'aquestes branques de la lingüística. Si una gramàtica és una descripció d'un sistema lingüístic que funciona en les ments de milers o milions de persones, sobta moltíssim que tots els exemples que s'hi inclouen provenen de les pàgines dels literats o de la inventiva dels gramàtics, dos grups de persones ben poc representatives de la totalitat d'una comunitat lingüística.” (Montoya, 1992: 391-392)

Una altra resposta al problema de la validesa, aquesta vegada més reflexiva i conscient, és la dels autors que fan un esforç de control de variables per arribar a comparar dos discursos que difereixin únicament en la dimensió oral / escrit o acústic / gràfic. Aquest esforç, d'una banda ben lloable, ha comportat d'altra banda l'obtenció d'uns textos sovint poc reals, produïts amb una restricció de variables que només es pot donar en condicions de laboratori. D'aquesta mena d'experiments he trobat els de G.H.J. Drieman (1962), J.W. Gibson (1966), R. O'Donnell (1974), M.E. Poole i T.W. Field (1976), B. Kroll (1977), G. Redeker (1984 i 1990), S.A. Thompson (1984), C.S. Butler (1992), L.W. Smith (1994) i Y. Bestgen (1998), que presento i comento a continuació:

**1– G.H.J. Drieman (1962)** estudia les produccions orals i escrites de vuit llicenciats en psicologia. “*From each subject a spoken as well as a written text was obtained in one sitting under conditions that were, as nearly as possible, identical for all sittings*” (1962: 39). Per a això, col·loquen cada individu davant de dues diapositives d'obres pictòriques de Renoir i de Weissenbruch, i li demanen que escrigui i parli sobre la impressió que li causen.

– **Comentari:** les condicions de producció són, efectivament, idèntiques entre si, però radicalment diferents a les de la vida fora de l'experiment.

2– **J. W. Gibson (1966)** es basa en els textos produïts per 45 estudiants que parlen i escriuen sobre el mateix tema com a part dels seus exercicis de classe. – **Comentari:** Un exercici de classe no pot ser mai com una activitat comunicativa de fora de classe.

3– **R. O'Donnell (1974)** treballa amb una entrevista oral en la qual un personatge famós respon les preguntes de tres periodistes en un programa de televisió i amb quatre articles publicats al diari escrits pel mateix individu. – **Comentari:** una entrevista televisiva no és el gènere prototípic de la llengua oral. D'altra banda, com reconeix el mateix autor (vg. pàg. 109 del seu article), l'anàlisi d'un sol individu no pot conduir a conclusions de tipus general, perquè no té prou validesa externa.

4– **M.E. Poole i T.W. Field (1976)** enregistren 80 entrevistes individuals a estudiants universitaris a partir d'un guió estructurat de preguntes sobre la seva experiència a l'ensenyament secundari i a la universitat. – **Comentari:** els discursos resultants poden ser una mostra de la llengua oral produïda sota unes condicions particulars de temps, pressió i formalitat, però no pas de la llengua oral quotidiana.

5– **B. Kroll (1977)** estudia set narracions orals obtingudes en un seminari acadèmic i set narracions escrites sobre el mateix fet, posades com a deure per fer a casa. – **Comentari:** íd. anterior.

6– **G. Redeker (1984)** vol disposar de textos orals i escrits sobre el mateix tema i amb el mateix grau d'espontaneïtat. Per això enregistra els subjectes en una habitació insonoritzada d'un laboratori de fonètica. Els subjectes parlen a un individu que els escolta però que no els pot respondre (pot fer-ho només de manera no verbal) i se'ls demana "*to tell the story or to give the explanation as if in a casual conversation*" (1984: 46). D'altra banda, en la pretensió que els escrits tinguin un grau d'espontaneïtat similar als discursos orals, s'elaboren d'una manera ben estranya per a la llengua escrita: "*The writing task was a two-page essay in the form of a letter to the experimenter. The instructions emphasized:*

«...just write it down as it comes to mind... Please do not polish your product. Do not copy your manuscript».” (1984: 46) – **Comentari:** són situacions de comunicació que resulten gairebé pintoresques i que no tenen res a veure amb els textos que es donen a la vida real.

**7– S.A. Thompson (1984)** analitza textos breus de tres tipus: *formal written*, *informal written* i *informal spoken*. El segon grup consisteix en “*letters to the editor, personal letters, and chatty pieces about people in a radio program guide*” i el tercer en “*monologues in two-party conversation*” (1984: 88). El nombre total de clàusules per a cada tipus és d’unes 300 —si cada clàusula contingués una mitjana de 10 mots, serien uns 3000 mots. – **Comentari:** la composició del segon grup de textos sembla discutible; la del tercer no inclou la llengua oral informal prototípica, que és la dialogada; no considera la formació d’un grup per a la llengua oral formal; i, en darrer lloc, el conjunt de la mostra resulta una mica breu per arribar a conclusions gaire generals.

**8– G. Redeker (1990)** compara el funcionament dels marcadors discursius en els diàlegs i els monòlegs orals de vuit individus. Els fan veure pel·lícules mudes i explicar a un altre el contingut de la pel·lícula. “*Each pair was observed under two conditions, a dialogue and a monologue situation. In both cases, they communicated only auditorily via headphones and microphones without any visual contact*” (1990: 370-371). – **Comentari:** és, com en l’anterior de la mateixa autora, una situació de comunicació extremadament forçada.

**9– C.S. Butler (1992)** analitza quatre entrevistes entre l’investigador i sengles informants, “*(...) four ‘directed interviews’ in which the interviewer guides the discussion by asking series of questions about particular topics. (...) The corpus is claimed to represent educated Spanish as spoken in Madrid.*” (1992: 3). – **Comentari:** les *directed interviews* no poden prendre’s com a mostra del parlar quotidià de les persones, per educades que siguin. D’altra banda, l’escàs “realisme” d’aquesta experiència arriba fins a l’extrem de fer les transcripcions de la llengua oral amb comes i amb punts suspensius, amb un resultat visual molt similar a la llengua escrita.

**10– L.W. Smith (1994)** estudia tres gèneres de discurs (*informal conversation*, *formal interview* i *written essay*) produïts pels mateixos individus i sobre el mateix tema. Com que les converses i les entrevistes van ser realitzades expressament per a l'experiment, es varen donar als participants les instruccions següents: “(1) *for the informal conversation, participants were asked to have a conversation with their friend concerning the particular topic as they would at home or in the student union; (2) for the formal interview, participants were asked to conduct themselves in the interview as if they were being taped for an FM radio program to be broadcast on the university’s public radio station*” (1994: 56). – **Comentari:** el fet de realitzar els enregistraments en situacions creades expressament i no en els escenaris i en les condicions en què es produeixen en la realitat comporta sempre la producció d'un text diferent del que es vol analitzar.<sup>95</sup>

**11– Y. Bestgen (1998)** cerca els usos del que anomena *segmentation markers* (signes de puntuació, pauses, connectors i expressions referencials), en textos narratius creats per 30 individus francoparlants a partir d'una llista d'enunciats fornida per l'investigador. “*It was explained to them that they had to turn it into a continuous narrative. They were permitted to change the sentences, adding anything that was necessary to obtain a discourse that would seem natural, but they had to respect the list order.*” (1998: 759). – **Comentari:** resulta simptomàtica la petició d'obtenir un discurs que “pogués semblar natural”. En aquest context, qualsevol esforç de naturalitat sembla abocat a la teatralitat.

La intenció comuna d'aquests autors és, com he dit, el desig de controlar el màxim nombre de variables per augmentar la validesa dels experiments. El resultat són uns textos sotmesos a aquest control, però que són discursos forçats, produïts en unes condicions massa diferents de les que es donen en la comunicació real. El cas extrem és el de G. Redeker (1990), en què els parlants es comuniquen sense veure's —com per telèfon, però amb una tasca que tampoc no és l'habitual en una conversa telefònica. La prova més clara d'això és que en molts casos s'hagin donat als participants instruccions per parlar “com si fos en tal o tal situació”. De les recerques ressenyades, només R. O'Donnell (1974) treballa amb textos

---

<sup>95</sup> Seria una recerca interessant de comparar textos reals i textos produïts en condicions de laboratori i determinar en quina mesura difereixen i si podrien portar a resultats erronis en l'estudi de trets lingüístics concrets.

reals, però el problema en el seu cas és que l'entrevista televisiva es pren com a representant de tota la llengua oral<sup>96</sup>.

En conclusió, es produeix un fenomen indesitjat i amb conseqüències metodològiques importants: no es poden controlar totes les variables sense falsejar la realitat mateixa. Aquest fet, que ocorre també en altres àrees científiques, resulta particularment aplicable al llenguatge, tal com assenyala D. Tannen (1989):

“Just as the scientific study of whales or elephants or chimpanzees must include painstaking observation and description of particular, individual creatures interacting with each other in their natural environments, so the scientific study of language must include the close analysis of particular instances of discourse as they naturally occur in human and linguistic context.” (Tannen, 1989: 196)

Si el control de variables és un camí per procurar la validesa dels experiments, l'altre camí és fer la mostra prou àmplia com perquè sigui representativa de tota —o de la major part— de la producció comunicativa en una llengua. Aquest ha estat el camí triat per D. Biber (1985, 1986a, 1986b, 1988, 1989, 1990, 1992, 1993, 1995a i 1995b), D. Biber i E. Finegan (1986), D. Biber i B. Altenberg (1992), D. Biber i M. Hared (1994), Y.-J. Kim i D. Biber (1994), R. Reppen (1995) i D. Biber, S. Conrad i R. Reppen (1998).

D. Biber planteja un tipus de recerca global que anomena *multi-feature / multi-dimensional approach* i que consisteix en una comparació oral / escrit a partir d'un corpus extens,

---

<sup>96</sup> També hi ha altres autors la mostra dels quals, tot i estar constituïda per textos reals, presenta problemes de tipus diversos. Així, per exemple, J. Blankenship (1962) se sol posar com a exemple de la falta quasi total de control de variables, de manera que acaba deduint que són més importants les diferències individuals que les del contrast oral / escrit. C. Blanche-Benveniste (1995a) i G.H.J. Drieman (1962) utilitzen mostres massa curtes, de només 1050 mots (aprox.) i 2655 mots, respectivament, per arribar a conclusions de tipus de general. M. Macaulay (1990) no inclou en la seva mostra cap exemplar de conversa espontània i analitza només entrevistes formals de tema científic, a partir de les quals vol caracteritzar la llengua oral en conjunt. J.-P. Bronckart (1985), en una descripció de la tipologia dels discursos en francès modern que aspira a ser completa, col·loca només 31 textos orals, la majoria dels quals són entrevistes no espontànies, al costat de 268 textos escrits. G.B. Sherman (1993) tria els textos amb un bon control de variables per a l'objectiu que persegueix (el tipus textual expositiu oral i escrit), però analitza la producció d'un sol individu. I, per acabar, E. McClure i E. Geva (1983), en un estudi sobre l'ús cohesiu de les conjuncions adversatives en el discurs, no treballen amb textos sencers, sinó amb fragments de dues oracions amb un buit entremig, que ofereixen als seus 21 informants perquè els diguin quina conjunció, *but* o *although*, “*makes the two sentences sound better*” (E. McClure i E. Geva, 1983: 414).



amb textos pertanyents a un nombre considerable de gèneres i sotmesos a tota mena de paràmetres contextuals, dels quals s'analitzen un nombre també extens de trets gramaticals. Així, D. Biber (1988) treballa amb un total aproximat de 960000 mots en anglès, distribuïts en 481 textos, 141 d'orals i 340 d'escrits, d'entre 2000 i 5000 mots i de 23 gèneres diferents (converses, entrevistes, conferències, documents oficials, cartes personals, prosa acadèmica, etc.). Sobre aquests textos aplica tècniques d'anàlisi factorial, que assegurin que els resultats no provenen de classificacions prèvies de l'investigador sinó de les agrupacions estadístiques que presenten les dades.

Les dimensions de la recerca de D. Biber han estat, doncs, espectaculars i el seu grau de representativitat ha de ser, forçosament, molt superior a tots els treballs anteriors sobre el tema. Per tant, els seus resultats, que presentaré al llarg d'aquesta tesi, han de ser molt tinguts en compte. Malgrat això, el treball de D. Biber, com és lògic, també té els seus punts febles des de la perspectiva metodològica. Per exemple, N. Oostdijk (1988) hi fa les objeccions següents<sup>97</sup>:

- 1- "Their samples are too small."
- 2- "One may question the extent to which certain samples can be regarded as representative of a particular genre."
- 3- "Biber and Finegan never once question the validity of genre distinctions they use."
- 4- "Biber and Finegan restrict themselves to linguistic features that have been identified in previous research as potentially important." (Oostdijk, 1988: 19)

D. Biber (1990) respon N. Oostdijk (1988) amb uns càlculs sobre la longitud i el nombre mínim de textos necessaris perquè una mostra o un corpus puguin considerar-se representatius (reprenc aquesta qüestió en el punt 3.5.1). El que D. Biber (1990) no rebutja és la crítica tercera sobre la validesa dels gèneres del seu estudi, i és sobre aquest aspecte que voldria fer uns comentaris<sup>98</sup>. D. Biber (1988) es basa en textos de dos corpus de la

---

<sup>97</sup> N. Oostdijk (1988) es refereix en concret al treball de D. Biber i E. Finegan (1986), però les seves crítiques poden fer-se extensives a tota la recerca efectuada dins de la línia de D. Biber. D'altra banda, les objeccions que cito apareixen escampades al llarg de dos paràgrafs de l'article de N. Oostdijk i les he unificat en una llista numerada únicament per facilitar-ne la presentació.

<sup>98</sup> Les observacions que faré sobre els gèneres dels corpus emprats per D. Biber (1988) són extensives a R. Reppen (1995) i R.N. Smith i W.J. Frawley (1983), ja que treballen sobre els mateixos corpus de base.

llengua anglesa, el *LOB Corpus* i el *London-Lund Corpus*, que es distribueixen en els 23 gèneres següents<sup>99</sup>:

“**Written genres 1-15 from the LOB corpus**: 1. Press reportage, 2. Editorials, 3. Press reviews, 4. Religion, 5. Skills and hobbies, 6. Popular lore, 7. Biographies, 8. Official documents, 9. Academic prose, 10. General fiction, 11. Mystery fiction, 12. Science fiction, 13. Adventure fiction, 14. Romantic fiction, 15. Humor, 16. Personal letters, 17. Professional letters.

**Spoken genres from the London-Lund corpus**: 18. Face-to-face conversation, 19. Telephone conversation, 20. Public conversations, debates and interviews, 21. Broadcast, 22. Spontaneous speeches, 23. Planned speeches.” (Biber, 1988: 67)

La definició d'aquests gèneres, que ja estan preestablerts en els corpus originals, no apareix en cap moment en la presentació de l'experiment, malgrat que després els utilitza com un element central per a les seves conclusions —hi dedica un capítol sencer, el núm. 8, “Extending the description: variation within genres” (1988: 170-198). D'altra banda, el nom que reben alguns dels gèneres resulta poc clar (p. ex., *Religion* i *Skills and hobbies* no són gèneres sinó temes, i *Popular lore* o *Humor* són més aviat àmbits o funcions)<sup>100</sup>. També sembla que es barrejin criteris de gènere amb criteris de registre (es dóna una informació insuficient sobre la mena de textos inclosos sota les etiquetes *Spontaneous speeches* i *Planned speeches*, i els mateixos termes utilitzats — *spontaneous/planned*— no són denominacions pròpies de gènere sinó de registre).<sup>101</sup>

---

<sup>99</sup> D. Biber treballa exclusivament amb els textos d'aquests corpus, amb l'excepció d'unes cartes personals que ell mateix explica que va incorporar-hi pel seu compte (1988: 66).

<sup>100</sup> R.N. Smith i W.J. Frawley (1983) estudien alguns elements de cohesió en textos del *Brown English Corpus* distribuïts en quatre gèneres també bastant discutibles: *fiction*, que és una suma de gèneres; *journalism*, que és una professió o àmbit; *religion* que és un tema i, potser, un àmbit; i *science*, que és un àmbit amb molts subàmbits i moltíssims temes. Tots quatre rètols poden contenir moltes i diverses menes de textos corresponents al que he definit com a *gènere*.

<sup>101</sup> Com a detalls més quantitius, també sobta de veure que, entre els discursos triats, n'hi ha un nombre diferent per a cada gènere, amb discrepàncies que arriben des de 80 textos per al gènere *academic prose* fins a només 6 textos per als gèneres *science fiction* o *personal letters*. Cal dir, a més, que s'observa una desproporció entre el nombre de discursos orals i escrits: els orals són només 141 del total de 481 que s'analitzen, encara que el nombre de mots total sigui similar, ja que els orals són fragments de 5000 mots i els escrits són fragments de només 2000 mots —desigualtat que, d'altra banda, també seria discutible. Una altra qüestió a remarcar és la notable dispersió dels resultats que D. Biber (1988) obté per a cada gènere, que, com veurem, contrastarà amb la nitidesa de la distribució en bandes dels resultats de la present recerca (vg. § 4.10.6).

En resum, i com a conclusió d'aquest punt sobre la validesa interna i externa, cal dir que el camí que triï no pot ser cap dels comentats fins aquí. No pot ser el del control d'un gran nombre de variables, perquè conduiria a uns textos irreals. No pot ser tampoc l'anàlisi multidimensional propugnada per D. Biber, per dos motius principals: el primer és que la meua recerca presenta unes limitacions inherents que ja he comentat en el punt anterior i el segon és que tampoc no disposem en llengua catalana d'uns corpus prou generals, aptes per a l'anàlisi del discurs, tant de llengua oral com de llengua escrita.

Només em resta, doncs, un camí, més humil, que és definir amb molta exactitud les coordenades dels textos que analitzaré i acceptar que els resultats tindran un abast més reduït i, per tant, una representativitat que s'haurà de relativitzar —això és, no s'haurà de magnificar, però tampoc no s'haurà de menystenir. A les pàgines següents concretaré aquestes coordenades<sup>102</sup>.

### **3.4.2- REGISTRES I SEQÜÈNCIES TEXTUALS DOMINANTS**

Al punt corresponent als gèneres discursius ja he presentat les diferències entre els conceptes de *registre* i de *tipus de text* (o *seqüència textual* dominant). Aquí pertoca de caracteritzar la mostra de textos respecte d'aquestes variables, cosa que faig en el quadre de la pàgina següent de manera resumida i amb unes etiquetes generals que a continuació matisaré. Com es pot veure, el resultat de la comparació és que la majoria de trets varien entre la conversa informal i els altres dos gèneres, de manera que només esdevenen variables de control entre la classe magistral i la prosa acadèmica i són, en canvi, trets associats a la variable dependent en el contrast entre la conversa i el conjunt dels altres dos gèneres.

---

<sup>102</sup> El treball que més s'assembla al meu dels que he consultat és el de R. Carter i M. McCarthy (1995). Aquests autors parteixen d'una mostra de quatre gèneres reals, no obtinguts en condicions de laboratori, i com jo admeten la limitació de la seva tria. Els seus resultats, però, no podran servir de base de comparació perquè no hem coincidit gaire ni en els gèneres triats ni en els trets analitzats.

<b>TAULA 1</b>	<b>CONVERSA INFORMAL</b>	<b>CLASSE MAGISTRAL</b>	<b>PROSA ACADÈMICA</b>
<b>Tema</b>	general	específic (àmbit acadèmic)	específic (àmbit acadèmic)
<b>Mode</b>	oral dialogat espontani	oral monologat espontani / preparat	escrit monologat preparat
<b>To</b>	informal	formal	formal
<b>Tenor</b>	interactiu	informatiu	informatiu
<b>Seqüència textual dominant</b>	conversacional	expositiva	expositiva

Com ja he comentat anteriorment, em sembla que l'intent de convertir tots els trets en variables de control a les tres columnes falsejaria tant la conversa que ja no seria vàlida per a l'anàlisi. En efecte, una conversa de tema específic, planificada, formal, informativa i expositiva no seria, en absolut, el que entenem per conversa quotidiana.

Els trets del quadre han de ser comentats, ja que els rètols que hi he col·locat tenen merament el caràcter de tendència general. Així, les classes magistrals i els escrits acadèmics de la mostra coincideixen a tenir un grau d'especificitat mitjà / alt i una funció informativa similar, ja que la intenció didàctica de les classes és equiparable a la intenció divulgativa dels escrits. A més, cal dir que, sense tractar exactament el mateix tema, totes les classes i tots els escrits estan dins de l'àmbit acadèmic o científic de la ciència històrica. Difereixen, en canvi, en el grau d'espontaneïtat o de preparació prèvia, perquè les classes, per més que puguin planificar-se en els continguts i que hi hagi un guió previ de l'exposició, sempre seran discursos "improvisats" en el sentit que cadascun dels enunciats és creat sobre la marxa. Aquesta relativa improvisació de la classe magistral desapareix en

la llengua escrita, que és reflexiva per naturalesa i tot hi pot ser revisat i reelaborat. A més, a les classes m'he assegurat que hagin estat impartides sense llegir en veu alta un text escrit<sup>103</sup>: han consistit bàsicament en un monòleg del professor, sense intervencions de l'alumnat<sup>104</sup> i sense recolzar en transparències o en fotocòpies distribuïdes als assistents<sup>105</sup>.

Pel que fa a la tipologia textual, sempre és difícil de classificar textos sencers sota una etiqueta o una altra. D'acord amb la formulació més acceptada (vg. § 3.2.2), més que tipus de text hi ha seqüències arquetípiques d'enunciats, que solen aparèixer combinades entre si. El text en conjunt sol tenir un tipus dominant i per això he etiquetat els meus com a conversacionals i expositius. En realitat, dins de les converses de la mostra hi ha seqüències de tots els tipus i dins de les classes se'n poden trobar sobretot d'argumentatives i d'expositives, en combinació amb la narració, que és propi que aparegui en textos històrics. Aquest caràcter combinatori fa que els tipus de text siguin una unitat incòmoda a l'hora de considerar-los com a variable, tant si és independent, dependent o de control. Pel que fa a aquesta darrera possibilitat, la variable de control, crec que no es podria forçar que una conversa fos només descriptiva o que una classe fos només argumentativa: una altra vegada s'estaria alterant la realitat.

Per posar un exemple d'això, he estat buscant en les converses fragments de més de tres o quatre línies que puguin considerar-se *discurs en situació*<sup>106</sup> gairebé pur, amb molt poca

---

<sup>103</sup> Me n'he assegurat de la manera més fefaent, és a dir, essent present entre l'alumnat en el moment de la gravació. Pel que fa a aquesta presència, atesa la llarga experiència docent dels professors enregistrats, ha d'haver influït molt poc o gens en la seva actuació. La situació "camuflada" entre els alumnes de les darreres files de l'aula i la utilització de només mitja hora —no pas la inicial— d'un total de tres o quatre hores d'enregistrament per individu, han d'haver neutralitzat l'anomenada *paradoxa de l'observador*, que es produeix per la dificultat d'observar les persones en situacions reals sense que alterin el comportament a causa del fet mateix de l'observació.

<sup>104</sup> Només a la classe de l'individu A hi havia una intervenció breu d'un alumne, amb la corresponent resposta del professor, que he suprimit del text final.

<sup>105</sup> Pel que fa a les fotocòpies, s'hi refereixen en algun moment els individus A i E, però és un fet ocasional: la major part del discurs es basa en paraules pronunciades a l'aula amb total autonomia de cap text escrit present davant dels alumnes.

<sup>106</sup> Els conceptes de *discurs en situació* i de *discurs històric* van ser creats per É. Benveniste (1966) i es poden definir com segueix: "El discurs en situació es caracteritza perquè conté elements díctics que l'insereixen en la situació comunicativa circumdant, mentre que el discurs històric és presentat com una cosa exterior al «jo, ara, aquí» de l'enunciació." (Castellà, 1992: 187-188)

combinació amb altres tipus de discurs, i només n’he trobat un, que apareix a la conversa núm. 3 i que abraça aproximadament des de la línia 3853 (pàg. 82, *Annex: Textos analitzats*) fins a la lín. 3924 (pàg. 83, íd.). A continuació en dono una petita mostra, en la qual assenyalo els díctics<sup>107</sup> amb color vermell<sup>108</sup>:

“RIA: no sé quina era la meva de màquina\  
 SEP: ah\  
 no ho sé\  
 RIA: viam\  
 SEP: la curt- --  
 la llarga\  
 RIA: aquesta\  
 VIE: ara s'hauria de- --  
 JON: no no\  
 aquesta no\  
 l'altra\  
 MAG: ah\  
 vale\  
 RIA: no/  
 JON: aquesta aquesta\  
 SEP: viam encén-la\  
 JON: sí\  
 aquesta\  
 sí\  
 MAG: sí però gaire o\_  
 JON: aquesta aquesta és\  
 aquesta\  
 RIA: tu coneixes més el que l'encén/  
 aquesta no\  
 aquesta era la meva\  
 ALL: bueno\  
 voleu dir que hi perdereu tant/  
 JON: e\_

<sup>107</sup> Sobre l’ús dels mots *dixi* i *díctic/a/s/es*, adoptats com a formes normatives en llengua catalana, voldria assenyalar que en publicacions anteriors havia emprat *deixi* i *deíctic/a/s/es*, més similars al grec original i a les llengües modernes que ens són més pròximes. A J.M. Castellà (1992: 160) ho argumentava i em referia a L. Payrató (1988: 33), que havia defensat el mateix criteri amb anterioritat. També M.J. Cuenca emprava en els seus escrits les formes amb *e* —pot veure’s així, encara, en els fragments de M.J. Marín i M.J. Cuenca (1998) que cito en el punt 5.5.1—, tot i que a M.J. Cuenca (1996) ja hi aparegui *dixi-díctic*. I, encara, amb força anterioritat, G. Rigau (1981) havia utilitzat també les formes clàssiques. Sembla, finalment, que l’opció *deixi-deíctic* la devíem fer pocs autors, i al llarg del darrer decenni la forma oficial s’ha anat imposant en els àmbits —que tampoc no són gaires— on s’usa el mot. Sembla, doncs, que deu haver arribat l’hora d’oblidar el nostre agosarat i, alhora, modestíssim —i, potser, fins i tot indecorós— flirteig amb la dissidència i de retornar al plàcid camí comú de la norma establerta.

<sup>108</sup> Els criteris de transcripció de les converses que aniran apareixent al llarg de l’explicació es poden trobar explicats en el punt 3.6.1. El text complet de les converses és a *Annex: textos analitzats* (CD-Rom adjunt).

jo res perquè **me** la varen donar”

(Conversa 3, pàg. 82, lín. 3861-3889)

En aquest fragment, la situació en què dos dels parlants han intercanviat distretament els seus encenedors i no reconeixen quin és el de cadascú dóna lloc a un fragment llarg de referències díctiques als objectes que tenen a les mans. Malgrat això, a la darrera línia ja hi ha un enunciat a mig camí del tipus històric o narratiu, en què es combina la referència díctica representada pel pronom feble de primera persona *me* amb el temps de passat en el verb (perfet perifràstic) que ja remet a un fet ocorregut amb anterioritat. Per tant, com he dit, en un examen detallat la immensa majoria de textos són tipològicament mixtos i, en conseqüència, la variable tipus de text no és de les més aptes per ser controlades.

### 3.4.3- TEXTOS, INDIVIDUS I SITUACIONS DE PRODUCCIÓ

La mostra es compon de cinc fragments de converses informals, cinc de classes magistrals i cinc d'escrits acadèmics, d'una longitud de 5000 mots i escaig cadascun, amb un total d'uns 80000 mots (vg. a l'apartat 3.5, la justificació de la llargària dels textos i de les dimensions de la mostra).

Les converses informals formen part del *Corpus de Català Contemporani de la Universitat de Barcelona (CUB)*, en particular de l'anomenat *Corpus Oral de Conversa Col·loquial (COC)*<sup>109</sup>. Les classes magistrals i els escrits acadèmics han estat recopilats per mi mateix. A continuació comento per a cada gènere les condicions de producció, les d'enregistrament en el cas dels discursos orals i les característiques dels individus que hi participen<sup>110</sup>:

**a/ Conversa informal:** tots els textos han estat obtinguts en una situació real de comunicació, sense que els participants sabessin que eren enregistrats per a cap experiment

---

<sup>109</sup> Una descripció acurada d'aquest corpus, dels textos que el componen i de les convencions de transcripció que s'hi han aplicat es pot trobar a A. Oller i altres (2000).

<sup>110</sup> Una presentació més detallada d'aquestes dades, especificada individu per individu, es pot trobar a *Annex: Textos analitzats*, pàg. 3, 45, 73, 107, 137, 171, 187, 201, 215, 229, 245, 255, 265, 275 i 285.

—encara que posteriorment, com és lògic, se'ls demanés autorització per utilitzar la gravació, tot cobrint els aspectes ètic i legal. Són converses disteses, realitzades al menjador familiar entre parents o amb amics, abans o després d'un àpat. El nombre total de participants ha estat de 29, 11 homes i 18 dones, amb unes edats compreses entre els 18 i els 85 anys, i amb una mitjana de 39 anys (sense comptar-hi una criatura de 3 anys que hi té només tres intervencions més que breus). D'aquestes 29 persones n'hi ha 18 que tenen o estan cursant estudis universitaris. Tots els individus (excepte un de la conversa 1) parlen el català central, tenen el català per llengua de provinença familiar i per llengua habitual, i viuen a l'àrea de Barcelona.

**b/ Classe magistral:** els cinc fragments de classe magistral s'han enregistrat en aules universitàries d'una facultat de lletres. Les classes han estat impartides íntegrament en català, han estat monologades i s'han fet sense llegir ni dictar els apunts a l'alumnat. Com he dit anteriorment, el professor ha anat desenvolupant el fil sintàctic i cohesiu del discurs de manera improvisada, si bé el contingut de la classe s'ha basat en un guió previ, present al damunt de la taula en tots els casos. Aquesta "improvisació", a més, pot estar condicionada pel fet que, com és habitual en l'àmbit universitari, el professor hagi llegit, hagi escrit i hagi parlat sobre el mateix tema diverses o moltes vegades amb anterioritat<sup>111</sup>. Els cinc individus triats són quatre homes i una dona<sup>112</sup>, d'entre 39 i 56 anys, catedràtics o professors titulars d'universitat, amb una experiència i un prestigi considerables dins i fora del seu camp. L'àrea de coneixement triada ha estat *Història (història econòmica, història contemporània i història de l'extrem orient, segons el cas)* perquè és un dels camps que

---

<sup>111</sup> El grau d'espontaneïtat és molt difícil de controlar, perquè l'elaboració prèvia que cada parlant fa d'un text sol ser un procés mental, intern a l'individu, difícilment observable i mesurable.

<sup>112</sup> La proporció d'homes i de dones, tant en els individus de les classes i dels escrits com en els de les converses, no permetrà d'arribar a cap conclusió sobre la variació de gènere en els trets estudiats. Una recerca sobre les diferències en la complexitat lingüística entre dones i homes demanaria de treballar amb un nombre molt més extens d'individus i, potser, amb fragments de text més curts per fer viable quantitativament la recerca. Tot i això, sembla que les diferències de sexe, en català i en les llengües del nostre entorn cultural, s'han de manifestar més en el nivell pragmàtic que en l'estrictament lingüístic o cohesiu. Aquesta és, per exemple, la conclusió de L. Cortés (1986a):

"Ni los datos aportados por el sexo, la edad o el contexto tuvieron implicaciones de correlación. (...) Con estas últimas palabras, y ciñéndonos a las variables extralingüísticas, queremos afirmar categóricamente la superior incidencia que tuvo el nivel sociocultural de los informantes con respecto a los otros dos, edad y sexo." (Cortés, 1986a: 203-204)



funcionen amb més normalitat en català<sup>113</sup>. Els individus seleccionats són persones que fan les seves classes i publiquen una part important dels seus treballs en català i que tenen aquesta llengua com a principal, de provinença familiar i d'ús més habitual. D'altra banda, també m'he assegurat que no han fet una estada llarga (de mesos o d'anys) a l'estranger en els darrers cinc anys<sup>114</sup>.

**c/ Prosa acadèmica:** com és lògic, els cinc escrits demanen menys comentari que els textos dels gèneres anteriors. Són dels individus que ja he descrit a la classe magistral i m'han estat facilitats per ells mateixos<sup>115</sup>. Són en tots els casos articles o capítols de llibre publicats i escrits originalment en català<sup>116</sup>.

Cal fer, encara dues reflexions més sobre la constitució de la mostra. En primer lloc, hi hauria hagut la possibilitat de provocar una conversa en la qual intervinguessin els mateixos individus que en les classes i en els escrits. Vaig valorar-ho acuradament i, un cop més, vaig arribar a la conclusió que seria una situació forçada de la qual no resultaria una conversa quotidiana (caldría dir-los el motiu de la trobada, probablement mantindrien la formalitat associada a la seva posició professional, les intervencions podrien ser quasi monologades i, a més, els individus seleccionats són persones molt relacionades amb la llengua escrita i una mostra de converses seves podria ser poc representativa del català col·loquial de la major part de la població). He preferit treballar, doncs, amb converses quotidianes inqüestionables en la majoria dels seus paràmetres contextuals i conformar-me amb un nombre menor de variables de control.

---

<sup>113</sup> En altres àrees hi són més dominants altres llengües. Per exemple, a les facultats de Dret del nostre país es treballa principalment en castellà, tant pel que fa a les classes com a la bibliografia, i a algunes facultats d'Econòmiques s'utilitza tant l'anglès que arriba a ser la llengua vehicular d'un percentatge important de classes, de seminaris i, fins i tot, d'algunes reunions de treball en els departaments.

<sup>114</sup> He trobat alguns professors, que he descartat per a la mostra, que han passat tant de temps a l'estranger i en època tan recent que ells mateixos consideren que parlen en català una mena de llengua, si no híbrida, sí amb moltes influències de l'idioma del país on han residit.

<sup>115</sup> Ometo els títols dels escrits per salvaguardar l'anonimat dels autors.

<sup>116</sup> No es pot saber mai, i especialment en català, el grau d'intervenció que hi hagin pogut tenir els correctors de normativa i d'estil. No compto, però, que puguin ser responsables de canvis gaire transcendents en els aspectes lèxics i sintàctics que analitzo.

Un altre aspecte que podria afectar la validesa de la mostra és el fet que els individus de les classes i dels escrits siguin només cinc, cosa que podria comportar un biaix idiolectal. No hi ha dubte que hauria estat més vàlid treballar amb més individus, però en aquest aspecte han prevalgut les exigències pel que fa al nombre mínim de mots requerits perquè cada fragment sigui representatiu i això ha limitat el nombre d'individus estudiats —c al tenir en compte que he treballat amb una mostra per sobre dels 80000 mots, cosa que, amb les codificacions successives dels diversos trets lèxics i sintàctics, ha resultat en una quantitat de feina gairebé desproporcionada. De totes maneres, com veurem al llarg dels capítols següents, les bandes de dispersió dels resultats han estat molt coherents i els cinc individus s'han comportat en general com una sola persona, sense gaires superposicions ni barreges a les bandes —en presento un resum en el punt 4.10.6. Aquesta consistència en els resultats crec que certifica la validesa de la mostra i l'allunya del perill del biaix idiolectal.

En resum, tal com he dit en el punt anterior sobre el registre i el tipus de text, s'estableixen dos nivells de variables de control:

**1- Variables de control comunes a tots els textos de la mostra:** català central; llengua de provenença familiar: català; llengua d'ús habitual: català; nivell sociocultural: mitjà-alt.

**2- Variables de control comunes només a la classe magistral i la prosa acadèmica:** tema específic, mode monologat, to formal, tenor informatiu, tipus de text predominantment expositiu i coincidència en els individus productors del text<sup>117</sup>.

Aquests dos nivells permetran d'arribar a resultats de validesa diferent segons la base de comparació. Així, en els resultats conjunts que abracin els tres gèneres, la validesa de les meves afirmacions s'haurà de limitar a la caracterització dels gèneres mateixos i no seran aplicables de manera absoluta a l'oposició oral / escrit, ja que hi intervenen altres factors del registre associats a cada gènere. Només en la mesura que la conversa es considera la

---

<sup>117</sup> Per tant, la classe magistral i la prosa acadèmica coincideixen en tots els trets excepte en el canal (acústic / gràfic) i en el grau de planificació (espontani / preparat). Aquest darrer tret es pot suposar sempre més alt en la llengua escrita, per la naturalesa mateixa del mitjà, malgrat que es tracti d'una característica que, com ja he assenyalat (vg. § 2.2.3), és difícil de determinar amb exactitud a causa del caràcter altament subjectiu que té.

forma prototípica de la llengua oral i que la prosa expositiva acadèmica pugui ser una bona representant de la llengua escrita, es podria parlar de manera general sobre aquests dos modes, en el benentès que les conclusions tindrien més el caràcter de tendències que de característiques absolutes. En canvi, quan la comparació sigui únicament entre la classe magistral i la prosa acadèmica, les diferències que s'hi trobin, en haver-hi més variables de control, seran més atribuïbles a l'oposició oral / escrit o, més ben dit, acústic / gràfic. Així, el contrast entre la classe magistral i la prosa acadèmica pot informar sobre algunes característiques específiques dels modes oral i escrit motivades únicament pel paràmetre de canal —i pel tret *grau de planificació* que hi està associat.

## 3.5- LES DIMENSIONS DE LA MOSTRA

### 3.5.1- EXTENSIÓ IDÒNIA DELS TEXTOS

Per al treball quantitatiu amb textos reals és important de determinar quina extensió han de tenir els textos perquè siguin representatius. N. Oostdijk (1988), en l'article que ja he esmentat anteriorment (vg. § 3.4.1), qüestiona els treballs de D. Biber per diferents motius, un dels quals és que els fragments dels corpus emprats —5000 mots per als orals i 2000 per als escrits— són, a parer de N. Oostdijk, massa breus. Motivats per aquesta crítica, D. Biber (1990) presenta els resultats d'una recerca quantitativa encaminada a determinar, entre altres qüestions, el tema que es planteja. Aquest autor compara parelles de fragments, de 1000 mots cadascun, extrets del mateix text, sobre un total de 481 textos orals i escrits, i en aquests fragments hi cerca deu trets lingüístics<sup>118</sup>. D. Biber fa el raonament següent:

“If large differences are found between the two 1,000 word samples, then we can conclude that 1,000-word samples do not adequately represent the overall linguistic characteristics of a text and that possibly 2,000-word and 5,000-word samples are also inadequate. If, on the other hand, the two text samples are similar linguistically, indicating that even a 1,000-word sample reliably measures the linguistic characteristics of a text, we can be fairly confident using the 2,000-word and 5,000-word samples of the LOB and London-Lund corpora.” (Biber, 1990: 258)

Els resultats de l'experiment confirmen, segons D. Biber, que els fragments de 1000 mots són suficients per a la recerca quantitativa en llengua<sup>119</sup>:

---

<sup>118</sup> Els deu trets lingüístics són els següents: *nouns, prepositions, present tense, past tense, third person pronouns, first person pronouns, contractions, passives, wh relative clauses, conditional subordination*.

<sup>119</sup> D. Biber (1990: 259) matisa, però, que els coeficients de fiabilitat són menors amb trets com les oracions condicionals i les oracions de relatiu.

“Overall, the results presented indicate a high level of stability for these linguistic feature counts across 1,000-word sub-samples of texts. This stability holds generally across linguistic features and across text categories. Given this stability between 1,000-word samples, it seems safe to conclude that the 2,000-word and 5,000-word texts in the standard corpora are reliable representatives of their respective text categories for analyses of this type.” (Biber, 1990: 261)

Per la meua part, també he volgut experimentar sobre aquesta qüestió i comprovar amb textos de la mostra, amb alguns dels trets lingüístics que investigaré i en llengua catalana, els mínims necessaris (Castellà, 1996c). L’experiment el vaig fer amb els fragments corresponents a la classe i a l’escrit de l’individu A. Per fer la prova vaig seleccionar les quatre categories lingüístiques amb més ocurrencies de la llista (noms, adjectius, verbs i preposicions), tot considerant que per als altres trets, amb un nombre inferior d’aparicions, caldria aplicar criteris més estrictes que els que resultessin per a aquests més abundants.

Els fragments van ser de 5000 mots, de 2500 mots —resultat de la partició en dos del fragment de 5000— i de 1300 mots —resultat aproximat de la partició en dos del primer fragment de 2500 mots. Els resultats van ser els que es poden veure a la taula 2, de la pàgina següent, en què s’indiquen els percentatges obtinguts en la classe magistral i en la prosa acadèmica per a cada fragment i les desviacions dels percentatges calculades prenent com a referència els resultats del fragment de 5000 mots.

Les dades són, com es pot veure, contundents. Les desviacions per als fragments de 2500 mots han estat en tots els casos inferiors a l’1 per cent, mentre que en els fragments de 1300 mots n’hi ha hagut de superiors a aquesta xifra —els he assenyalat amb vermell en el quadre. Això fa replantejar els resultats de D. Biber (1990) i fa concloure que és més segur treballar amb fragments de 2500 mots que de 1000, com ell proposa<sup>120</sup>.

---

<sup>120</sup> Cal pensar que els marges d’error obtinguts en els fragments de 1300 mots són suficients com per jugar alguna mala passada. El millor exemple d’això són els resultats en adjectius, en què la desviació de + 0.9 del fragment 1300a de la classe magistral afegida a la de - 1.9 del fragment 1300b, també de la classe, ha comportat un desnivell de 2.8 punts en unes xifres que no superen el 10 % i ha provocat el resultat decebedor de 1300a, en què han coincidit la classe i la prosa en el mateix percentatge (8.4 %), mentre que en el fragment 1300b la diferència entre ambdós gèneres ha estat considerable (del 5.6 % al 8.2 %).

TAULA 2	NOMBRE DE MOTS	RESULTATS		DESVIACIONS	
		CLASSE	PROSA	CLASSE	PROSA
NOMS	5000	18.8 %	25.8 %		
	2500a	18.4 %	25.9 %	- 0.4	+ 0.1
	2500b	19.3 %	25.7 %	+ 0.5	- 0.1
	1300a	18.3 %	24.7 %	- 0.5	- 1.1
	1300b	19.0 %	27.8 %	+ 0.2	+ 2.0
ADJECTIUS	5000	7.5 %	9.2 %		
	2500a	7.0 %	8.3 %	- 0.5	- 0.9
	2500b	7.8 %	9.4 %	+ 0.3	+ 0.2
	1300a	8.4 %	8.4 %	+ 0.9	- 0.8
	1300b	5.6 %	8.2 %	- 1.9	- 1.0
VERBS	5000	13.8 %	10.2 %		
	2500a	14.5 %	11.0 %	+ 0.7	+ 0.8
	2500b	13.4 %	9.4 %	- 0.4	- 0.8
	1300a	14.1 %	11.3 %	+ 0.3	+ 1.1
	1300b	14.4 %	10.6 %	+ 0.6	+ 0.4
PREPOSICIONS	5000	13.6 %	17.5 %		
	2500a	13.2 %	17.4 %	- 0.4	- 0.1
	2500b	14.1 %	18.4 %	+ 0.5	+ 0.9

(Castellà, 1996c: 180)

Per tant, en aquests quatre trets, que són els més nombrosos quantitativament —i també dels més laboriosos en la recerca manual—, caldrà treballar amb fragments de 2500 mots, quantitat suficient perquè dona lloc a uns marges d'error molt petits. Per als altres trets del treball, que són molt menys nombrosos en els textos (adverbis, nominalitzacions, subordinació, coordinació, marcadors discursius, etc.), considero més aconsellable de treballar amb els fragments de 5000 mots, és a dir, amb la totalitat de la mostra.

### 3.5.2- EXTENSIÓ DE LA MOSTRA

Encara que a primera vista pogués semblar una cosa senzilla, comptar el nombre de mots d'un text és una operació bastant problemàtica. Els còmputos automàtics de què he disposat són el del processador de textos *Word 97* i el del programa d'anàlisi de textos *TACT* (al punt 3.6.2 exposo les característiques d'aquest programa). El criteri de tots dos és ortogràfic: es considera *mot* qualsevol caràcter o agrupació de caràcters precedits i seguits d'un espai en blanc o d'un signe de puntuació, en els textos escrits, o bé d'una barra entonativa, en la transcripció dels textos orals<sup>121</sup>. Els dos programes difereixen, però, en la manera de tractar els apòstrofs i els guionets. Així, *Word 97* en fa cas omís i, per tant, compta agrupacions com *l'home*, *agafa'm* o *desfer-se'n* com un sol mot. Els criteris que segueix *TACT* (*Makebase F2*) resulten més convenients, perquè és capaç de discernir en casos com els esmentats, cosa que facilita el recompte del total de mots gramaticals:

**1– L'apòstrof separa en dos mots.** Per tant, els articles o la preposició *de* apostrofsats amb un nom i els pronoms apostrofsats amb el verb o entre si, són comptats com a mots independents (*l'home*: dos mots; *agafa'm*: dos mots; *n'hi*: dos mots).

**2– El guionet separa en dos mots.** Per tant, els pronoms febles units al verb amb un guionet són comptats com a mots independents (*imposar-se*: dos mots).

**3– Els parèntesis, signes de puntuació i barres d'entonació no són mots** ni afecten per a res el càlcul.

**4– Els nombres en xifres són mots.**

En conseqüència, he adoptat el nombre de mots a *TACT* com a xifra base per a tots els càlculs de la tesi. D'acord amb aquest còmput, la mostra està formada per un total de 80096 mots, subdividits en 15 fragments, cinc de cadascun dels gèneres estudiats. El nombre de mots per a cada fragment, amb el nom dels arxius corresponents, és el següent<sup>122</sup>:

---

<sup>121</sup> S'hi inclouen com a mots, per tant, en el cas de la llengua oral, els omplidors (*eee*, *emm*), les interjeccions (*eh/*) i els mots inacabats o estroncats (*la prop-*).

<sup>122</sup> El fet que el nombre de mots sigui només aproximadament el mateix d'un fragment a un altre i d'un gènere a un altre, atès que la diferència no és gaire gran, no té cap conseqüència de cara als càlculs, ja que normalment els expressaré en percentatges —per ser més exactes, en tants per cent i en tants per mil.

<b>TAULA 3</b>	<b>NOM DEL FITXER</b>	<b>NOMBRE DE MOTS</b>
<b>Conversa 1</b>	1C5000.DOC	5557 mots
<b>Conversa 2</b>	2C5000.DOC	5359 mots
<b>Conversa 3</b>	3C5000.DOC	5361 mots
<b>Conversa 4</b>	4C5000.DOC	5516 mots
<b>Conversa 5</b>	5C5000.DOC	5461 mots
<b>TOTAL CONVERSES</b>		<b>27254 mots</b>
<b>Classe individu A</b>	AO5000.DOC	5265 mots
<b>Classe individu B</b>	BO5000.DOC	5316 mots
<b>Classe individu C</b>	CO5000.DOC	5299 mots
<b>Classe individu D</b>	DO5000.DOC	5241 mots
<b>Classe individu E</b>	EO5000.DOC	5253 mots
<b>TOTAL CLASSES</b>		<b>26374 mots</b>
<b>Escrit individu A</b>	AE5000.DOC	5265 mots
<b>Escrit individu B</b>	BE5000.DOC	5303 mots
<b>Escrit individu C</b>	CE5000.DOC	5341 mots
<b>Escrit individu D</b>	DE5000.DOC	5286 mots
<b>Escrit individu E</b>	EE5000.DOC	5273 mots
<b>TOTAL ESCRITS</b>		<b>26468 mots</b>
<b>TOTAL DE LA MOSTRA</b>		<b>80096 mots</b>



## 3.6- ALTRES ASPECTES METODOLÒGICS

### 3.6.1- SOBRE LA TRANSCRIPCIÓ I SOBRE LA PREPARACIÓ DELS TEXTOS

He treballat amb un sistema de transcripció mínim, basat en els criteris exposats per L. Payrató (1993 i 1995), però limitat als elements més imprescindibles. Per evitar problemes de reconeixement informàtic i atès que aquest treball abraça només aspectes lèxics i sintàctics, he estandarditzat totes les variants fonètiques i morfològiques dels mots. Les convencions mínimes de transcripció adoptades han estat les següents:

- ( / ) Entonació final ascendent (interrogativa) : sí/
- ( \ ) Entonació final descendent (declarativa) : sí\
- ( \_ ) Entonació final sostinguda (neutra) : sí\_
- ( - ---, - -, - ) Mot o enunciat estroncat: *es pot esperar un al- --*
- ( e, ee, eee ) Mot-crossa o omplidor, propi del discurs oral monologat.
- ( ??? ) Parlant no identificable.

La transcripció l'he efectuada jo mateix en el cas de les classes<sup>123</sup>. Per a les converses, en canvi, he pogut disposar de les transcripcions ja existents en el *COC* del *CUB*, però m'ha calgut simplificar-les per no tenir problemes en el còmput de mots amb el programa *TACT*. La simplificació —que ha consistit a reduir tots els símbols i convencions als que acabo d'indicar a la llista anterior— s'ha efectuat automàticament per mitjà d'un programa de

---

<sup>123</sup> La part de la transcripció de les classes on he trobat més dificultat ha estat la determinació de les corbes d'entonació finals ( \_ / \ ) i la delimitació de la longitud de les unitats entonatives. Crec que qualsevol estudi basat en aquests aspectes hauria de partir de transcripcions realitzades alhora per diverses persones que hagin pogut contrastar prèviament els criteris subjectius de transcripció. Per tant, les corbes d'entonació i la longitud de les unitats entonatives de les transcripcions que presento s'han de considerar merament indicatives i, per això mateix, no les he utilitzat per a cap càlcul ni per a cap correlació dins de la tesi.

“neteja” elaborat per Lluís de Yzaguirre (IULA, Universitat Pompeu Fabra), amb un repàs manual posterior per a algunes restes no resoltes pel programa.

D'altra banda, tant els textos orals com els escrits els he hagut de preparar per tractar-los amb *TACT*. Les operacions efectuades han estat les següents:

a- Supressió dels guionets en els mots compostos (si no, *TACT* els comptaria com dos mots).

b- Supressió dels guionets i dels espais en blanc dels numerals expressats en lletres (si no, serien comptats per *TACT* com diversos mots: *tres-cents vint-i-u* serien cinc mots ortogràfics).

c- Supressió del punt en la *ela* geminada (si no, *TACT* hi comptaria dos mots).

d- Supressió, en els textos escrits, de les notes a peu de pàgina.

e- Estandardització, en els textos orals, de les formes diferents d'un mateix mot (p. ex. *aquesta* i *aqueta*, *anar* i *nar*; si no, *TACT* comptaria com dos mots diferents el mateix mot).

f- Repàs global amb el corrector ortogràfic (perquè no interessa que *TACT* compti com dos mots la mateixa paraula escrita per error de transcripció amb dues grafies diferents).

### **3.6.2- SOBRE LA FIABILITAT EN LA CODIFICACIÓ**

Els treballs de tipus quantitatiu recolzen en la codificació com a aspecte bàsic de la fiabilitat dels resultats. El concepte de fiabilitat i el de rendiment metodològic són presentats per M.T. Turell (1995) de la manera següent:

“I per tal que aquesta codificació sigui tan completa i fiable com sigui possible s’han de tenir en compte els principis següents:

3.3.1. D’una banda, el *principi de rendiment metodològic*, ja que interessa saber no només quan ocorre algun fenomen, sinó també quan aquest fenomen no es produeix. Això vol dir que s’ha d’haver observat qualitativament de forma correcta, de manera que es coneguin

mínimament bé les pautes de comportament lingüístic dels membres de la comunitat i es compregui la naturalesa de la variable,

3.3.2. I, d'altra banda, el *principi de fiabilitat*, perquè s'han de comprovar les ocurrencies d'una determinada variable tantes vegades com calgui fins arribar a un 95-100 % de seguretat no només amb relació a les produccions de la variable, sinó també a les de cadascuna de les variants." (Turell, 1995: 36)<sup>124</sup>

Aquests principis es poden posar en perill segons el mètode emprat per a la codificació. En aquest sentit, una decisió fonamental és l'adopció d'un sistema de còmput manual o automàtic de les dades. En el treball de tipus macroscòpic de D. Biber (1988) es va efectuar la recerca de les ocurrencies d'una manera automàtica, per mitjans informàtics, a partir d'algorismes elaborats *ad hoc* (p. ex., quan el mot *x* apareix en el context lingüístic *n*, pertany a la categoria *a*; en cas contrari, pertany a la categoria *b*)<sup>125</sup>. C. Ball (1994) ha fet notar que aquest sistema pot excloure de la recerca fenòmens importants, ja que els algorismes de què es disposa actualment en lingüística computacional estan encara lluny de ser prou acurats —ho demostra amb un tipus de clàusules en anglès (subordinades amb *that-deletion*) que han restat amagades en l'algorisme aplicat per D. Biber (1988)<sup>126</sup>. Aquesta autora proposa que la recerca automàtica de dades lingüístiques vagi sempre acompanyada de mètodes manuals de codificació i que, mentrestant, els treballs automatitzats sobre grans corpus es facin amb ple coneixement de les limitacions que tenen<sup>127</sup>. C. Ball acaba afirmant la necessitat dels estudis de tipus microscòpic:

---

<sup>124</sup> Aquest principi m'ha fet desistir, com es veurà al lloc corresponent (vg. § 5.1.2), del còmput del nombre total de clàusules complexes en els textos orals, ja que la codificació que en feia jo mateix en passades successives donava resultats massa diferents.

<sup>125</sup> La majoria d'aquests algorismes estan explicats a D. Biber (1988: 221-245). A més, una explicació més detallada dels procediments que D. Biber utilitza en la seva recerca automatitzada ha aparegut posteriorment a D. Biber, S. Conrad i R. Reppen (1998), especialment en la part IV, "Methodology boxes", pàg. 243-280.

<sup>126</sup> Una deficiència similar és la que es dona a D. Biber i B. Altenberg (1992), en què es calcula la presència d'elements de referència en el text sense diferenciar entre els oracionals i els textuals, i sense distingir tampoc els endofòrics dels exofòrics o díctics. L'arrel del problema és també en el còmput automatitzat de les dades, ja que l'ordinador no pot fer la distinció entre els diferents tipus de referència que hi ha en els usos dels pronoms personals, dels pronoms demostratius, etc.

<sup>127</sup> Una altra observació interessant de C. Ball (1994) és que l'ordinador difícilment pot ser exhaustiu en la recerca dels elements lingüístics de realització zero, és a dir, pot ser que no compleixi el principi de rendiment metodològic enunciat per M.T. Turell (1995: 36) que acabo de citar.

“Large-scale macroscopic studies of multiple linguistic phenomena such as those conducted by Biber are thus premature. For the results to be reliable and obtained within a reasonable amount of time, such studies require first the support of adequate microscopic studies of individual phenomena, and secondly the development of robust text analysis tools of large parsed corpora. (...) The development of robust systems for processing unrestricted text is still a long-term goal, as Church (1991) points out and most computational linguists would probably agree: ‘In the long term, at least, we have the responsibility to deliver a large grammar with broad coverage for unrestricted text. We need to start thinking now about how we could ever hope to achieve this long-term goal.’ (Church, 1991, p.102).” (Ball, 1994: 301)

En el meu estudi m’he trobat també en la necessitat de decidir quina part del recompte podia ser automàtica i quina havia de ser manual. En la majoria dels trets analitzats, el treball de codificació ha estat manual, ja que ha calgut fer un tipus de distincions que els programes informàtics aborden de manera incompleta. Així, per exemple, vaig realitzar una prova amb el lematitzador i el desambiguador elaborats per L. de Yzaguirre (IULA, Universitat Pompeu Fabra) per a la codificació de noms, verbs, adjectius i adverbis, i es van presentar dos problemes importants. En primer lloc, el desambiguador funcionava amb un marge d’ambigüitat raonable en els textos escrits, però en els textos orals i especialment en les converses, el nombre de lemes dubtosos o ambigus augmentava fins a una proporció inassumible —el motiu era la brevetat de molts dels enunciats, que limitava molt la capacitat de desambiguar a partir de regles sintàctiques que posen en relació un mot amb els mots circumdants. En segon lloc, el lematitzador estava mancat d’algunes etiquetes que eren clau per al treball, com ara la distinció entre verbs copulatius i verbs predicatius, o entre adjectius atributius i adjectius predicatius, i això obligava igualment a una codificació manual posterior.

Un altre bon exemple de la problemàtica del recompte automàtic d’alguns trets és el verb *ser*. Per trobar totes les ocurrències d’aquest verb calia fer recerques individuals amb *TACT* per a cadascuna de les seves arrels (*sóc, ets, és, som, sou, són, er-, ér-, fui, for-, fou, fór-, ser-, sigu-, fos- fós-, sent, estat*) i restava el problema de l’ambigüitat de la forma *estat*, que també pot ser el participi del verb *estar*. El sistema, per tant, era tan laboriós com la recerca manual. Si a això s’hi afegeix que calia distingir entre els usos copulatius i els usos intransitius de *ser*, es pot entendre que s’hagi imposat l’opció de la codificació manual en

aquest i en molts altres casos, no només per raons de fiabilitat, sinó, fins i tot, per raons pràctiques de rapidesa en l'execució<sup>128</sup>.

El programa d'anàlisi del discurs *TACT* (*Textual Analysis Computing Tools*) desenvolupat per I. Lancashire i M. Stairs (Universitat de Toronto)<sup>129</sup>, tampoc no resolía el problema. *TACT* és una eina útil per al maneig de textos, permet de seleccionar mots, fer-ne el recompte i presentar-los en un context amb facilitat. Però a *TACT* hi ha d'arribar el text prèviament codificat (p. ex., no sap distingir quan el mot *que* és una conjunció, un pronom relatiu, un adverbí o un adjectiu determinatiu exclamatiu, ni quan, essent conjunció, introdueix una subordinada substantiva en funció de subjecte, d'objecte directe o de complement del nom, etc.). L'únic càlcul que és totalment automàtic i que *TACT* ofereix d'una manera còmoda i immediata és l'índex de variació lèxica (*type / token*) i el de repetició lèxica (*token / type*), així com el còmput del nombre total de mots del text. Al llarg de la tesi, en l'exposició dels resultats, aniré indicant quan he treballat amb *TACT*, quina part de la feina ha facilitat i quina part ha hagut de ser manual.

D'altra banda, però, el risc de la codificació manual és l'error humà, que podria posar en entredit la fiabilitat dels resultats. En el meu cas, he emprat dos procediments combinats per assegurar un marge d'error acceptable. El primer ha estat la codificació de diversos trets relacionats entre si en passades successives sobre els mateixos exemplars dels textos. Per exemple, he fet en primer lloc la recerca dels noms i en una segona mirada la dels adjectius, i així he descobert alguns noms que havien passat inadvertits en la primera ronda. Una operació semblant és la que he fet amb els nexes oracionals, els connectors textuais i les interjeccions<sup>130</sup>.

---

<sup>128</sup> El fet de treballar amb quinze fragments i voler disposar de dades independents per a cadascun per tal de veure la distribució de resultats per textos i per individus —que després he sumat per obtenir els totals de cada gènere— feia molt laboriós el processament automàtic del verb *ser*, ja que la recerca de totes les arrels calia fer-la quinze vegades, una per a cada fragment.

<sup>129</sup> Per al maneig de *TACT* he aprofitat les indicacions i els consells d'I. Lancashire i M. Stairs (1994), d'A. Oosterhoff (1995) i de J. Pérez Guerra (1998).

<sup>130</sup> En el cas dels connectors textuais i de les interjeccions, la codificació manual, amb una observació directa de tot el text, ha permès de trobar connectors i interjeccions que no eren a les llistes de referència que havia consultat. Com assenyalen C. Ball (1994) i M.T. Turell (1995), una recerca manual i qualitativa ha augmentat el rendiment metodològic i ha enriquit els resultats.

El segon procediment ha estat de fer una segona codificació, amb unes setmanes de diferència, d'alguns fragments dels textos, triats a l'atzar, que han representat aproximadament una cinquena part del total. El resultat comparatiu entre les dues codificacions ha estat excel·lent, ja que el marge d'error ha estat inferior, en la majoria de trets, a l'1.6 %. En alguns trets isolats en què el marge d'error ha estat superior, aquesta prova m'ha servit per adonar-me de la debilitat dels criteris de codificació que havia emprat i ha fet que els descartés de l'anàlisi (és el cas dels usos d'*això*, *aquest/a/s/es* i *ell/a/s/es* com a referents intraoracionals, com a díctics discursius i com a díctics exofòrics: els criteris de codificació no eren prou nítids, el marge d'error es disparava i, per això, he exclòs aquestes distincions de l'anàlisi)<sup>131</sup>.

La codificació manual que he fet ha estat, doncs, tan detallada que el present treball, que té un caràcter quantitatiu preferent, es pot considerar que té alhora un component qualitatiu destacat. De fet, la meua codificació és una combinació de l'anàlisi purament quantitativa (vàlida només en els trets monofuncionals) i de l'anàlisi qualitativa (possible en els trets monofuncionals i imprescindible en els plurifuncionals).

Una altra qüestió per comentar en relació amb la codificació és l'existència de diverses unitats lingüístiques de difícil delimitació, algunes de les quals han estat definides de manera discrepant per diferents autors i tendències. Això, que ja ocorre amb alguns elements de la sintaxi oracional (p. ex., la categoria adverbi, la distinció entre els tipus d'oracions compostes, el concepte mateix d'oració, etc.), es dona encara amb més intensitat amb les unitats de tipus discursiu, que han començat a estudiar-se en època relativament recent i presenten moltes dificultats per a l'anàlisi. En aquests casos, he aplicat un principi de consistència: atès que l'objectiu de la recerca no és la definició d'aquestes unitats, sinó el paper que tenen en el contrast entre els gèneres que es volen caracteritzar, l'important és adoptar un criteri únic i explícit de codificació, mantenir-lo de manera sistemàtica al llarg de tot el treball i aplicar-lo de manera idèntica a cadascun dels textos i a cadascun dels

---

<sup>131</sup> Cal excloure del marge d'error de l'1.6 % la distinció entre els usos oracionals i textuals de les conjuncions de subordinació, de coordinació i d'interordinació, que he afrontat malgrat la dificultat en la codificació.

gèneres. Aquest és el punt de vista mantingut per M.A.K. Halliday (1985b) i D. Biber, S. Conrad i R. Reppen (1998):

“For purposes of comparing spoken and written English it does not matter exactly where we draw the line provided we do it consistently.” (Halliday, 1985b: 63)

“A more complicated question, however, is what precisely to include as a noun or as a verb. There is no single correct answer to this question. Rather, the important thing is to make a principled decision and then be consistent in making counts.” (Biber, Conrad i Reppen, 1998: 66)

### 3.6.3- SOBRE LA SIGNIFICACIÓ ESTADÍSTICA DE LES DADES: EL TEST $\chi^2$ (CHI QUADRAT)<sup>132</sup>

En estudis de tipus quantitatiu com el present és habitual d'aplicar el test  $\chi^2$  (chi quadrat). Aquesta prova estadística està enfocada a resoldre un problema habitual en la interpretació de les dades: ¿quan es pot afirmar que els resultats presenten entre si unes diferències prou significatives? ¿A partir de quines xifres es pot considerar que el grau de contrast entre dues variables —en el cas present, els gèneres discursius— és suficient per ser donat com a bo? El test  $\chi^2$  permet, doncs, d'assegurar la significació estadística dels resultats, sempre que sigui en les dades que s'anomenen *nominals*, és a dir, les que no estan subjectes a gradació o escala, a diferència de les *ordinals* i de les *d'interval*:

“Some variables are nominal —that is, the attribute either is or is not present, a person either possesses or does not possess a particular attribute. You can't measure *how much* left-handedness a person possesses. There are, of course, exceptions where a nominal variable could be changed to an interval variable. For example, we can say a person either is or is not bilingual. Bilingualism is then a nominal variable. However, we can measure just how much bilingualism a person has; in that case, it's an ordinal or even interval variable. (...) When we

---

<sup>132</sup> He adaptat al català com a *chi* el nom de la lletra grega  $\chi$ , d'acord amb el criteri que m'ha ofert el Departament de Grec de la Facultat de Filologia de la Universitat de Barcelona (consulta 28.04.00). Aquesta lletra, que en castellà s'anomena *ji* i en anglès *chi* —pronunciat, d'acord amb la fonètica del mot en grec clàssic, [kai] i no [t i] ni [t ai] —, podria haver-se transcrit també *khi*, però és una grafia inusual en català i, a més, m'ha semblat que seria més una transliteració que una transcripció. En congruència amb això, la pronúncia catalana de *chi* ha de ser la velar [ki] —com a *Bosch* o com la *ch* de l'italià— i no pas la palatal [t i] que generalment s'associa a aquest dígraf a partir de l'ús que se'n fa en llengua castellana.

measure nominal variables, we are concerned not with *how much* but with *how many* or *how often*.” (Hatch i Farhady, 1982: 165)

Les dades d'aquesta tesi són nominals perquè compto els percentatges d'aparició d'uns trets lingüístics de manera que, per a cada tret, en un gènere donat, només hi ha dues possibilitats: aparèixer-hi o no aparèixer-hi.

El test  $\chi^2$  dóna, de fet, el grau de probabilitat que les freqüències observades siguin degudes a l'atzar i, per tant, la mesura en què poden ser tingudes en compte perquè responen a una causa no atzarosa. Generalment, en estudis de tipus sociològic, se sol considerar satisfactori un grau d'atzar inferior al 5 % o al 2 %, segons el casos<sup>133</sup>. La mecànica del càlcul de  $\chi^2$  es basa en la distinció entre les freqüències observades en una mostra i les freqüències teòriques o esperades en el supòsit d'independència:

“Supongamos que en una muestra particular un conjunto de sucesos posibles  $E_1, E_2, E_3, \dots, E_k$  (...) se observa que ocurren con frecuencias  $o_1, o_2, o_3, \dots, o_k$ , llamadas *frecuencias observadas*, y que según las leyes de las probabilidades, se espera que sucedan con frecuencias  $e_1, e_2, e_3, \dots, e_k$ , llamadas *frecuencias esperadas* o *teóricas*. (...) A menudo deseamos saber si las frecuencias observadas difieren significativamente de las esperadas. (...) Una medida de la discrepancia existente entre las frecuencias observadas y esperadas viene proporcionada por el estadístico  $\chi^2$  (léase ji-cuadrado) dado por

$$\chi^2 = \frac{(o_1 - e_1)^2}{e_1} + \frac{(o_2 - e_2)^2}{e_2} + \dots + \frac{(o_k - e_k)^2}{e_k} = \sum_{j=1}^k \frac{(o_j - e_j)^2}{e_j}$$

(...) Si  $\chi^2 = 0$ , las frecuencias observadas y teóricas coinciden completamente; mientras que si  $\chi^2 > 0$ , no coinciden exactamente. A valores más grandes de  $\chi^2$ , mayor discrepancia entre las frecuencias observadas y esperadas.” (Spiegel, 1988: 268-269)<sup>134</sup>

---

<sup>133</sup> Aquesta informació m'ha estat facilitada per Frederic Udina, professor d'Estadística del Departament d'Economia i Empresa de la Universitat Pompeu Fabra. La seva orientació i el seu suport han resultat essencials per al desenvolupament d'aquest punt 3.6.3.

<sup>134</sup> El càlcul de  $\chi^2$  es pot fer automàticament amb el programa *Excel 97* de *Microsoft Office 97*, que ofereix els resultats directament en percentatges de probabilitat d'atzar. Si no, caldria fer els càlculs del sumatori un per un i contrastar-los amb les taules “Valores percentiles ( $C^2$ ) para la distribución de ji-cuadrado con  $\nu$  grados de libertad” de C.M. Thompson que apareixen a M.R. Spiegel (1988: 538, “Apéndice IV”), cosa que faria molt laboriós el càlcul (els graus de llibertat  $\nu$  s'obtenen per  $\nu = (h-1)(k-1)$ , essent  $h$  el nombre de columnes i  $k$  el nombre de files en una disposició de les dades en forma de taula).



Per a una comprensió millor de les operacions i de les relacions que es posen en joc en aquesta prova estadística, n'ofereix uns exemples graduats en diverses magnituds. Imaginem que per a una mostra d'un total de 750 mots, repartits a parts iguals en tres gèneres A, B i C (250 mots per cada gènere), s'han obtingut uns resultats (= *resultats observats*), per a un tret lingüístic  $\alpha$ , de 40 ocurrencies en el gènere A, de 50 en el gènere B i de 60 en el gènere C. Alhora, d'acord amb un repartiment probabilístic proporcional, els resultats esperats haurien estat de 50 ocurrencies en cada gènere. Aquest conjunt de dades es poden representar en les dues taules següents:

**A- TAULA DE RESULTATS OBSERVATS (*o*) (Exemple hipotètic)<sup>135</sup>**

		1	2	3	
		A	B	C	TOTALS
1	a	40	50	60	150
2	no a	210	200	190	600
TOTALS		250	250	250	750

<sup>135</sup> Aquest quadre i el següent incorporen unes columnes de totals perquè el nombre de mots no té per què ser el mateix a cada columna. De fet, és ben possible —i és el cas de la meua mostra—, que el nombre total de mots dels textos de cada gènere sigui aproximadament igual, però no exactament el mateix. En aquesta circumstància, el càlcul de freqüències esperades es fa en una proporció entre els totals de cada fila i els de cada columna —se segueix la fórmula  $E_{ij} = n_i n_j / N$ , on  $E$  és la freqüència esperada,  $i$  és la fila i  $j$  és la columna— cosa que complica el còmput global.

**B- TAULA DE RESULTATS ESPERATS (*e*) (Exemple hipotètic)**

		1	2	3		
		A	B	C	TOTALS	
1	a	50	50	50	150	
2	no a	200	200	200	600	
TOTALS		250	250	250	750	

El procediment aritmètic que se segueix per a l'aplicació de la fórmula —i que, afortunadament, com ja he dit, es calcula de manera automàtica amb *Excel 97*— és el que s'exemplifica al quadre de la pàgina següent<sup>136</sup>.

D'acord amb les operacions efectuades al quadre,  $\chi^2$  és de valor 5.0, xifra que contrastada amb les taules de C.M. Thompson (Spiegel, 1988: 538) dona un valor final del 8.2058 %<sup>137</sup>. Això significa que la probabilitat que els resultats obtinguts hagin estat causats per l'atzar i no pel contrast entre els gèneres A, B i C és d'un 8.2 % —xifra superior al 2 % i al 5 %, que són la mesura habitual en les ciències humanes— i, en conseqüència, faríem bé de descartar aquests resultats de la recerca.

<sup>136</sup> He elaborat aquests quadres a partir dels que presenten E. Hatch i H. Farhaday (1982: 169, *Table 13.3, Computation of  $\chi^2$* ).

<sup>137</sup> Les taules de valors de  $\chi^2$  amb dos graus de llibertat —calculats, com he dit, per  $v = (h-1)(k-1)$ , és a dir, d'acord amb el nombre de columnes menys un multiplicat pel nombre de files menys un [ $v = (3-1)(2-1) = 2$ ]— presenten un valor de 4.61 per a  $\chi^2_{.90}$  i de 5.99 per a  $\chi^2_{.95}$ , marge dins el qual s'obté, proporcionalment,  $\chi^2_{.917942}$  0, el que és el mateix, un percentatge de probabilitat d'atzar del 8.2058 %.

FILA	COL.	O (resultats observats)	E (resultats esperats)	O – E	(O – E) <sup>2</sup>	(O – E) <sup>2</sup> / E
1	1	40	50	-10	100	2.0
1	2	50	50	0	0	0
1	3	60	50	10	100	2.0
2	1	210	200	10	100	0.5
2	2	200	200	0	0	0
2	3	190	200	-10	100	0.5
$\chi^2 = \sum (o-e)^2/e =$						5.0

(Exemple hipotètic)

El punt clau del càlcul de  $\chi^2$  és en l'operació  $(O - E)^2$ , és a dir, el quadrat de la diferència entre els valors observats i els esperats. A mesura que augmenta la diferència entre O i E, el quadrat creix progressivament i això fa que les mostres més grans, en les quals es treballa amb xifres més altes, donin amb més facilitat uns resultats satisfactoris de  $\chi^2$  i, per tant, siguin més fiables. Per exemple, he calculat que si una mostra de 250 mots per gènere necessita, per a una probabilitat d'atzar inferior al 5 %, un diferencial mínim de 22 ocurrències entre els tres gèneres ( $\pm 11$ , en una relació d'1 és a 11 entre 22 i 250), una de 2500 mots per gènere en necessita un de només 70 ocurrències ( $\pm 35$ , en una relació d'1 és a 36 entre 70 i 2500) i una de 25000 mots per gènere necessita un diferencial de 219 ocurrències ( $\pm 109.5$ , en una relació d'1 és a 114 entre 219 i 25000). És a dir, com més

grans es fan les xifres, calen diferencials proporcionalment menors per acomplir amb unes mateixes expectatives de probabilitat d'atzar<sup>138</sup>.

En el meu cas, que he treballat amb un total de més de 80000 mots, repartits en uns 26000 mots per gènere, els contrastos entre les ocurrències de la majoria dels trets que he analitzat han situat la probabilitat d'atzar per sota del 0.0001 %, molt lluny del 5 % o del 2 % exigint habitualment i, per tant, en un nivell de fiabilitat molt alt<sup>139</sup>.

### 3.6.4- SOBRE ELS DENOMINADORS I SOBRE LES REPETICIONS CONTIGÜES

Una qüestió important en el càlcul dels percentatges amb què s'expressen la majoria dels resultats és la dels denominadors utilitzats per a la divisió. No és el mateix calcular el percentatge de nominalitzacions sobre el nombre total de mots, sobre el nombre de clàusules o sobre el nombre de noms, ni tampoc no pot tenir el mateix valor comptar els nexes de subordinació substantiva sobre el total de mots, sobre el nombre de clàusules o sobre el nombre de nexes de subordinació. La tria del denominador més adequat en cada cas és crucial per tenir una visió adequada del comportament i del pes real de cada unitat lingüística. Aquesta idea ha estat assenyalada per M.A.K. Halliday (1985b) en tractar sobre la proporció de mots lèxics i mots gramaticals en el text (vg. citació en el punt 4.6.1) i per

---

<sup>138</sup> Una altra manera d'explicar el mateix és comparar el valor de  $\chi^2$  per a quantitats proporcionals en mostres progressivament més grans. Així, si en una mostra de 25000 mots per gènere uns resultats de 4800, 5000 i 5200 ocurrències d'un tret lingüístic  $\alpha$  en els gèneres A, B i C, respectivament, donen una probabilitat d'atzar del 0.0045 %, en una mostra de 2500 mots uns resultats proporcionals de 480, 500 i 520 ocurrències la donen del 36.8 % i en una mostra de 250 mots uns resultats de 48, 50 i 52 ocurrències donen una probabilitat d'atzar del 90.5 %. Per tant, és clar que amb una mateixa proporció numèrica, el càlcul del quadrat de les diferències entre  $o$  i  $e$  dona uns resultats de  $\chi^2$  progressivament més satisfactoris com més àmplia és la mostra.

<sup>139</sup> He calculat el test  $\chi^2$  en totes les taules que s'expressen en nombre absolut d'ocurrències o en percentatge de nombre d'ocurrències, excepte aquelles en què es detallen els resultats text per text (conversa per conversa i individu per individu), que sempre han estat incloses en taules anteriors. No he fet el càlcul, lògicament, en les taules que reflecteixen el resultat de la mitjana d'una proporció (*type / token*, longitud de mot, etc.). En el cas dels percentatges, he hagut de treballar amb les dades absolutes a partir de les quals els havia obtingut, ja que  $\chi^2$  s'ha de comptar sempre a partir de xifres absolutes d'ocurrències. A cada apartat dels capítols 4 i 5 aniré oferint un comentari del resultat de  $\chi^2$  i de les seves conseqüències per a la fiabilitat de les dades.

C. Ball (1994), que defensa, per exemple, la clàusula com a denominador més adequat per a les oracions escindides:

“(...) relative frequency should be a measure of the number of times something occurs within the number of opportunities for it to occur. (...) Choosing the appropriate class is, of course, crucial and requires considerable thought. (...) the denominator in a frequency ratio must be of the same class as the numerator, and syntactic constructions such as the cleft sentence are not words, but clauses. Their relative frequency should therefore be measured in terms of number of clauses.” (Ball, 1994: 297-299)

Un altre aspecte que he de destacar és el de les repeticions contigües del mateix mot, del mateix sintagma o, fins i tot, de la mateixa clàusula, que m’ha fet dubtar a l’hora de fer els càlculs en la llengua oral. Finalment, he decidit, amb caràcter general, d’incloure-hi tots els mots, sense eliminar del còmput els repetits. M’han convençut principalment tres arguments. En primer lloc, el fet que les repeticions desenvolupen un paper important en la construcció del discurs oral i eliminar-les implicaria falsejar-lo. En segon lloc, la dificultat d’establir un límit a la repetició: ¿fins a quants mots repetits caldria descomptar? Quina distància màxima hi hauria entre els mots?, etc. I, en tercer lloc, no sembla que el nombre de repeticions sigui tan alt que hagi de fer variar substancialment les conclusions a què es pugui arribar. Com a exemple d’aquest darrer argument, hi ha el cas del verb *ser*. He calculat les repeticions contigües d’aquest verb a la classe de l’individu A i s’han produït en 12 ocasions de les 157 ocurrències registrades. En cas de descomptar-les, minvaria el percentatge de verbs copulatius sobre el total de verbs del fragment del 22.4 % al 20.8 %, canvi que tampoc no seria transcendental. A més, també caldria calcular les repeticions contigües de verbs predicatius i descomptar-los del total, de manera que la reducció en el percentatge de verbs copulatius encara seria menor.

Aquest criteri general, però, no l’he aplicat a les repeticions contigües dels elements subordinants i coordinants, tant si es tracta de preposicions com de conjuncions. He cregut que seria també falsejar els resultats si, per exemple, un sintagma preposicional introduït per dues ocurrències contigües de la preposició *de* fos comptat com dos sintagmes o si una subordinada substantiva amb dos *que* fos comptada dues vegades. Per tant, en el cas de

dues preposicions o dues conjuncions contigües, n'he sumat només una (això afecta el punt 4.7.2 i tots els apartats del capítol 5).

### 3.6.5- SOBRE LA RECERCA BIBLIOGRÀFICA

L'examen de la bibliografia sobre l'oposició entre llengua oral i llengua escrita m'ha reportat algunes sorpreses que voldria destacar<sup>140</sup>. He observat la tendència dels autors de cada àmbit lingüístic i geogràfic (francòfon, anglòfon —subdividit en britànic i nord-americà—, hispanòfon, germanòfon, etc.) a buscar les referències bibliogràfiques dins del seu àmbit. Que hi hagi una certa preferència és una cosa que s'entén, per la proximitat no només geogràfica, sinó també cultural o, fins i tot, de cercles científics afins. Ara bé, el que m'ha sobtat és que en alguns casos aquesta sigui una tendència tan accentuada, de manera que hi ha autors que esmenten quasi exclusivament les recerques efectuades en la seva llengua, sobre la seva llengua i en el seu entorn immediat.

Així, alguns estudiosos de l'àmbit de la filologia hispànica —amb l'excepció destacada d'A. Briz i S. Pons—, citen en una proporció elevada els treballs escrits en i sobre l'espanyol, i sovint obliden obres importants sobre el mode del registre publicades en altres llengües<sup>141</sup>. En els autors francòfons, entre els quals destaca C. Blanche-Benveniste, les citacions d'altres tradicions també són escadusseres. Però, fins i tot dins de l'àmbit anglòfon, potser l'únic que es podria permetre el luxe d'ignorar la resta —si això no fos

---

<sup>140</sup> No és només una sorpresa sinó una veritable curiositat bibliogràfica que el text d'Elinor Ochs (1979), "Planned and Unplanned Discourse", aparegut dins Talmy Givon (ed.), *Discourse and Syntax* i aparegut novament sense cap canvi dins Elinor Ochs i B. Schieffelin (1983), *Acquiring Conversational Competence*, sigui exactament el mateix que havia publicat Elinor O. Keenan dos anys abans, el 1977, amb el mateix contingut paraula per paraula, però amb un títol diferent, "Why Look at Unplanned and Planned Discourse?", dins del volum Elinor O. Keenan i Tina Bennett (eds.), *Discourse across Time and Space*. Per algun motiu personal aquesta autora va canviar entre 1977 i 1979 el cognom amb el qual firmava les seves publicacions i, a més, va decidir tornar a publicar el mateix text amb un canvi de títol (i sense avisar...). Aquesta operació d'autèntic camuflatge ha provocat el desconcert, la curiositat i el goig de la troballa —per aquest ordre— en el pobre doctorand que escriu aquestes ratlles, esdevingut ja irremissiblement detectiu de bases de dades bibliogràfiques —o, pitjor, *ratolí de biblioteca virtual*.

<sup>141</sup> Es dona fins i tot el cas de L. Cortés (1992: 58), que comet la "imprudència" de citar un autor anglòfon, M. Gregory (1967), i li atribueix les nocions de *camp*, *mode* i *to*, passant per alt les aportacions anteriors, originals de M.A.K. Halliday i T. Reid.

sempre un error—, perquè disposa d'una línia extensíssima de treballs de recerca en aquest camp, es produeixen exclusions i oblitats difícils d'entendre. Per posar només uns exemples: L.W. Smith (1994), en un treball bastant aprofundit sobre el contrast oral / escrit, no esmenta per a res l'obra fonamental de M.A.K. Halliday, i G.B. Sherman (1993), en una comparació de les estratègies de cohesió en discursos expositius orals i escrits, no cita en cap moment els treballs, també imprescindibles, de D. Biber. Encara més estrany és el cas de S. Greenbaum i G. Nelson (1995), que es declaren lingüistes del corpus, citen W. Chafe, K. Beaman i M.A.K. Halliday, i també deixen de banda D. Biber, precisament l'únic d'aquests autors que treballa amb extensos corpus preconstituïts i ha editat obres específiques sobre lingüística de corpus<sup>142</sup>. Així, fora de F.N. Akinnaso (1982 i 1985) i J. Luetkemeyer, C. Van Antwerp i G. Kindell (1984), que són precisament articles de revisió bibliogràfica, l'únic que presenta una visió més àmplia i completa és D. Biber (1986a i 1988).

La recerca catalana sobre la variació funcional s'ha caracteritzat, en canvi, per l'intent de repassar la major part de la bibliografia existent en les llengües més accessibles i oferir així una bona panoràmica sobre la qual fonamentar-se (vegeu, per exemple, els treballs de L. Payrató i de M.T. Turell que esmento en la bibliografia)<sup>143</sup>. Jo he volgut inserir-me en aquesta jove però ferma tradició i he procurat de fer una recerca bibliogràfica àmplia del contrast oral / escrit i dels aspectes que hi estan relacionats. Per a això he partit d'una mirada exhaustiva a les bases de dades *MLA Bibliography 1963-1981* i *MLA Bibliography 1981-1999* (MLA = *Modern Language Association of America*, Nova York), la segona de les quals és oberta i s'actualitza trimestralment. A més, he utilitzat les referències que es troben en obres de conjunt com D. Biber (1986a i 1988), D. Tannen i R. Freedle (1982 i

---

<sup>142</sup> Em refereixo a D. Biber (1986a i 1988) i a D. Biber i B. Altenberg (1992). Hi ha, encara, una darrera obra de D. Biber que S. Greenbaum i G. Nelson no poden citar, perquè és posterior al seu treball, però que acaba de demostrar la intervenció d'aquest autor en la línia de recerca a la qual s'adscriuen: es tracta de D. Biber, S. Conrad i R. Reppen (1998): *Corpus Linguistics. Investigating Language Structure and Use*.

<sup>143</sup> Aquesta obertura al que prové de l'exterior, que no sé si es tracta d'una característica del nostre país en tots els terrenys, ho és almenys en les publicacions d'àrees com l'anàlisi del discurs o la lingüística del text. Això es pot veure amb un simple cop d'ull a les referències bibliogràfiques de V. Salvador (1984 i 1989), L. Payrató (1988, 1992, 1993, 1994, 1995, 1998a i 1998b), M.T. Turell (1988 i 1995), A. Tusón (1988, 1991 i 1995), H. Calsamiglia (1990), J.M. Cots, L. Nussbaum, L. Payrató i A. Tusón (1990), M.J. Cuenca (1990a, 1996, 1998, 2000c, 2000d, en premsa a i en premsa b, etc.), H. Calsamiglia i A. Tusón (1991 i 1999), J.M. Castellà (1992 i 1995) i A. Cros (1995 i 1999).

1984) i en els articles de F.N. Akinnaso (1982 i 1985) i J. Luetkemeyer, C. Van Antwerp i G. Kindell (1984) que ja he esmentat. Aquestes fonts m'han garantit l'accés a la producció en anglès, que és la més copiosa, i a bona part de l'escrita en francès. Les obres de L. Cortés, per la seva banda, m'han donat entrada a la producció en espanyol. Amb això crec que he recollit la part més important —mai no pot ser “tot”— del que s'ha escrit recentment sobre el tema objecte d'aquesta tesi en les llengües que em resulten més pròximes i que puc llegir amb més facilitat.



## **4. RESULTATS DE LA RECERCA (I): DENSITAT LÈXICA, ESTIL NOMINAL I ESTIL VERBAL**

### **4.1- LA DENSITAT LÈXICA, L'ESTIL NOMINAL I L'ESTIL VERBAL**

En els treballs que he analitzat anteriorment hi ha un acord bastant generalitzat sobre el paper del lèxic en la diferenciació llengua oral / llengua escrita. Segons els autors citats, la llengua escrita tendeix a basar la construcció del discurs en la densitat lèxica i en l'organització dels mots al voltant dels noms, amb una presentació de la informació més compacta. La llengua oral, en canvi, és menys densa i articula els significats sobretot al voltant dels verbs.

Un dels primers estudis que fan referència a la densitat lèxica és el de J. Ure (1971), que calcula, en 34 discursos orals i 30 d'escrits, el nombre de mots amb contingut lexical respecte del total de mots de cada discurs i obté el resultat següent<sup>144</sup>:

“The most striking feature of the results is immediately obvious: the spoken texts, all except two have a lexical density of under 40 %; the written ones, all except two, have a density of 40 % and over. (...) Lexical density was largely a matter of the choice of spoken or written medium.” (Ure, 1971: 445-446)

Les conclusions de J. Ure podrien ser qüestionades metodològicament, ja que la mostra de textos que emprà no ha estat controlada pel que fa a paràmetres com la formalitat o el

---

<sup>144</sup> J. Ure (1971) no especifica quines categories de mots tenen contingut lexical i, per tant, quines són les que inclou en el còmput.

gènere. Però, malgrat això, aquesta autora té l'encert de plantejar un tema i de donar-ne unes conclusions que, després, en els treballs d'altres investigadors, s'han pogut confirmar.

Des del punt de vista de la psicologia del llenguatge, D.P. Hayes (1988) presenta un estudi també relatiu a la densitat lèxica. Segons aquest investigador, la llengua oral i l'escrita segueixen diferents patrons a l'hora de triar els mots que s'hi utilitzen. El discurs oral, tant si és entre adults com entre infants i tant si és en la conversa quotidiana com si es tracta de programes televisius, prefereix l'ús de mots gramaticals i dels mots lexicals més comuns. La llengua escrita, en canvi, presenta una proporció menor de mots gramaticals i major de mots lexicals poc comuns, tendència que té la màxima expressió en els textos científics:

“In this entire corpus, samples of science writing are unique, forming one extreme of word choice. Conversations involving children, books for preschool children and television shows designed for preschool children constitute the other extreme. (...) What makes science texts unique is their unusual use of function words, the reduced use of the common content terms and the extensive use of uncommon technical content words. (...) The texts of the most popular television shows [sèries televisives] simulate conversations. (...) In producing scripts for the actors, writers have followed the word choice pattern of natural conversation (...). These scripted conversations are difficult to distinguish (lexically) from natural conversations.”<sup>145</sup> (Hayes, 1988: 576-578)

Les raons que D.P. Hayes esgrimeix per a aquests dos patrons de comportament són les que ja coneixem d'altres autors, si bé amb una argumentació més pròpia de l'àmbit de la psicologia:

---

<sup>145</sup> D.P. Hayes fa bé de puntualitzar que la seva observació afecta només el lèxic, perquè en matèria de vacil·lacions, anacoluts, reformulacions, etc., la conversa natural difereix considerablement dels diàlegs cinematogràfics, teatrals i novel·lescós. D'altra banda, D.P. Hayes generalitza les seves conclusions sobre la llengua oral i la llengua escrita sense incloure en l'anàlisi cap gènere oral expositiu comparable amb la classe magistral o amb els *Prepared speeches* de D. Biber. De fet, com a mostra d'oral formal analitza les declaracions de testimonis en judicis realitzats en el Senat dels Estats Units, i considera sorprenent que siguin similars en la tria de mots a les converses quotidianes, sense tenir en compte que són discursos relativament breus, sense preparació prèvia, en una situació d'una certa tensió i sense temps per a l'elaboració lèxica —bastant diferents, per tant, dels gèneres orals expositius més habituals.

“Speech and writing normally occur under different time pressures. Differential access to words in our productive / recognition vocabularies has long been suspected and is now known to be frequency dependent (...). The less common the term, the longer the latency for its recognition and retrieval from memory (...). The implication is that under the time pressures of natural conversation, and especially where control of the floor is problematic, speakers are compelled to rely upon the most accessible (common) words. (...) In writing, there is normally more time to reflect on the most suitable (generally a less common) word. (...) Finally, because writing is more decontextualized than conversation, words must be chosen to overcome the advantages of a variety of nonverbal and contextual cues available during face-to-face conversation. Explicit reference is more important to comprehension in writing.”<sup>146</sup> (Hayes, 1988: 580-581)

M.A.K. Halliday (1985b) tracta la mateixa qüestió, si bé des d'un punt de vista més ampli. Per a ell, la densitat lèxica es materialitza en diversos aspectes interrelacionats. En primer lloc, la proporció de mots lexicals i de mots gramaticals en el discurs<sup>147</sup>. En segon lloc, l'estil nominal o verbal del discurs. Així, la llengua escrita es caracteritza perquè conté més densitat de mots lexicals i per un estil més nominal, i la llengua oral perquè conté més mots gramaticals i per un estil més verbal. L'estil nominal implica una major presència, òbviament, de noms, però també d'adjectius, de sintagmes preposicionals complements del nom i de cert tipus de subordinades nominals<sup>148</sup>:

“The difference between written and spoken language is one of DENSITY: the density with which the information is presented. (...) So the structure of the modern world and the structure of the language combine together to make the written language what it is: a

---

<sup>146</sup> F. Ferreira, J.M. Henderson, M.D. Anes, P.A. Weeks i D.K. McFarlane (1996) fan un experiment sobre el temps necessari per a la comprensió de mots d'alta i baixa freqüència dins de discursos orals i arriben també a la conclusió que la freqüència és un tret rellevant per a la dificultat de processament dels mots:

“(...) we have found evidence that lexical frequency affects the time required to process a word in spoken connected sentences. (...) High-frequency words in spoken sentences require less time to process than do low-frequency words.” (Ferreira, Henderson, Anes, Weeks i McFarlane, 1996: 324-334)

<sup>147</sup> M.A.K. Halliday considera, d'acord amb la definició clàssica de la lingüística estructural, que els mots lexicals són aquells que formen part d'un repertori obert, mentre que “*a grammatical item enters into a closed system*” (Halliday, 1985b: 63).

<sup>148</sup> Es tracta, com he avançat en l'explicació del treball de K. Beaman (vg. § 2.3.4), de les subordinades substantives amb verb en forma no personal, que en anglès poden ser amb formes d'infinitiu i de participi de present, i que en català es donen únicament en infinitiu. Aquest tipus de construccions, tot i que també afecten la densitat lèxica, les tractaré més endavant en el capítol de la complexitat sintàctica.

language with a high lexical density, measured in the number (and informational load) of lexical items per clause, and a strong tendency to encode this lexical content in a nominal form: in head nouns, other items (nouns and adjectives) in the nominal group, and nominalised clauses. It is these nominal structures that give the clause its enormous elasticity.” (Halliday, 1985b: 62-75)

D’aquesta manera, el tractament de la densitat lèxica s’eixampla i incorpora aspectes de l’organització sintàctica i semàntica dels mots en el discurs per a la transmissió del significat<sup>149</sup>. W. Chafe (1982), per la seva part, relaciona també la densitat informativa amb l’ús de nominalitzacions i d’adjectius atributius, i D. Biber (1988) hi afegeix aspectes com la longitud de mot i la variació lèxica (*type / token ratio*), que també tractaré en aquest capítol.

D’acord amb les idees exposades fins aquí, al llarg d’aquest capítol desenvoluparé l’anàlisi de la densitat lèxica i dels estils nominal i verbal en els apartats següents:

#### **4.2- Els noms**

#### **4.3- Els verbs**

#### **4.4- Els adjectius**

#### **4.5- Els adverbis**

#### **4.6- Els mots lexicals i els mots gramaticals**

#### **4.7- Els sintagmes preposicionals**

#### **4.8- Les nominalitzacions**

#### **4.9- La variació lèxica, la repetició lèxica i la longitud de mot**

#### **4.10- La densitat lèxica, l’estil nominal i l’estil verbal: visió de conjunt**

---

<sup>149</sup> El concepte que planteja M.A.K. Halliday va, de fet, encara més enllà. A parer meu, “*the density with which the information is presented*” és una qüestió que depassa l’àmbit estricte del lèxic o de la sintaxi i que demanaria un tractament global de l’organització de la informació en el discurs. La llengua oral i l’escrita no difereixen només en el percentatge de mots lèxics o en el nombre de sintagmes preposicionals complement del nom, sinó que tota l’estructuració informativa (repeticions, perífrasis, aclariments, recursos retòrics, etc.) té un caràcter diferent segons cada mode. L’estudi d’aquesta vessant de la coherència textual no admet, probablement, un tractament merament quantitatiu, sinó que exigiria una anàlisi qualitativa profunda que escapa de l’abast d’aquesta tesi.

## 4.2- ELS NOMS

### 4.2.1- FONAMENTACIÓ CONCEPTUAL I HIPÒTESIS

Per a diversos autors, com ja he dit, la llengua escrita té un estil nominal i la llengua oral té un estil verbal. Això significa que els escriptors confien més en els noms i en les estructures nominals per a la construcció del discurs i prefereixen una forma d'expressió més abstracta i estàtica, mentre que els parlants recolzen més en els verbs i en les estructures verbals, i tendeixen a una expressió més concreta i més mòbil. Dels trets lingüístics en què es materialitzen aquests estils, en aquest apartat tractaré els noms i en els apartats següents aniré desenvolupant els verbs, els adjectius, els adverbis, els sintagmes preposicionals i les nominalitzacions.

D. Biber (1988) ha comptabilitzat el nombre de noms del text. En el seu càlcul, aquest autor ha pres en consideració “*all nouns included in the dictionary, excluding those forms counted as nominalizations or gerunds*” (1988: 228). Que no hi hagi comptat les nominalitzacions l'allunya del criteri que utilitzaré en aquest treball, però que hagi exclòs els gerundis l'apropa, ja que en català aquesta forma no personal no té el valor nominal característic de l'anglès. El resultats obtinguts per D. Biber han estat els següents:

<b>TAULA 4 *</b>	<i>Face-to-face conversation</i>	<i>Prepared speeches</i>	<i>Academic prose</i>
<i>Nouns</i>	13.7 %	18.9 %	18.8 %

(\* Percentatges sobre el total de mots - Taula elaborada a partir de les dades de Biber, 1988: 255, 265 i 269)

Les xifres presentades a la taula 4 indiquen un percentatge menor de noms a *face-to-face conversation* i una igualtat entre els altres dos gèneres, de manera que la distinció de camp general o específic preval sobre la de mode oral o escrit<sup>150</sup>.

C. Gelpí (1998) calcula el nombre de noms en relació amb el mode del registre i conclou —en coincidència, en aquest cas, amb E.O. Keenan / E. Ochs (1977 / 1979) (vg. § 2.2.3)— que la freqüència d'aparició dels noms depèn del tret espontani / preparat i del tret informatiu / interactiu<sup>151</sup>:

“(2) El mode condiona el nombre de noms del text: si el text és preparat en el mode, predominaran les unitats amb categoria gramatical nom. I el tenor complementa el mode: si el text té tenor informatiu, té més possibilitats que es confirmi la idea prèvia que si el text té el tenor interactiu.” (Gelpí, 1998: 139)<sup>152</sup>

En una altra direcció argumentativa —que no afecta només els noms sinó qualsevol categoria lexical— M.A.K. Halliday (1985b) destaca que en la diferenciació entre mot lexical i mot gramatical hi ha un *continuum*: alguns mots lèxics tenen poca aportació de contingut o poc pes informatiu i estan en el límit amb els mots gramaticals. Aquests

---

<sup>150</sup> W.N. Francis i H. Kuèera (1982), en un recompte de la freqüència dels mots en un corpus ampli de llengua anglesa escrita (*Frequency Analysis of English Usage: Lexicon and Grammar*), troben un percentatge de noms del 26.8 %. Si bé aquesta xifra no s'ha obtingut només sobre la prosa acadèmica sinó sobre el conjunt de la llengua escrita, el resultat és més congruent que el de D. Biber i s'aproxima bastant al meu, que, com veurem més endavant, s'ha situat en el 24.5 % per a la prosa acadèmica. Cal tenir en compte, també, que W.N. Francis i H. Kuèera (1982) i D. Biber (1988) no han treballat sobre el mateix corpus (els primers ho han fet sobre el *Brown Corpus*, d'anglès americà, i el segon ho ha fet amb el *LOB Corpus*, d'anglès britànic). Malgrat aquest factor, però, el resultat de D. Biber resulta sorprenent i difícil de comprendre.

<sup>151</sup> C. Gelpí (1998) adopta el concepte de densitat lèxica de M.A.K. Halliday i fa el recompte de les ocurrences de diverses categories lexicals en català d'acord amb diversos paràmetres dels registres. Tot i que al llarg d'aquest capítol aniré citant les conclusions d'aquesta autora, no podré utilitzar les seves dades com a base de comparació quantitativa perquè no estan presentades en forma de percentatges i, a més, la seva anàlisi no es basa en el contrast oral / escrit. D'altra banda, el corpus del seu treball és de dimensions relativament reduïdes i amb molt poca presència de la llengua escrita.

<sup>152</sup> La composició de la meua mostra, per les característiques que té, no permet de contrastar aquesta afirmació i decidir si és a causa del canal, del grau de preparació o del tenor que apareixen més noms en els textos. Malgrat això, ja he dit abans que la distinció espontani / preparat, que és central en el corpus sobre el qual treballa C. Gelpí, pot ser discutible. A més, en aquest corpus la presència de textos escrits és molt minsa —només ho són dos dels 16 textos analitzats (vg. Tomàs, 1998: 47). De fet, en la intenció que en va presidir la constitució, els textos havien de ser tots orals o, almenys, emesos oralment, en una gradació àmplia des de l'espontaneïtat a la planificació (L. Payrató, comunicació personal).

mots fronterers, que fan una contribució menor a la densitat lèxica, són els que apareixen amb més freqüència en els textos:

“The vocabulary of every language includes a number of highly frequent words, often general terms for large classes of phenomena. Examples from English are *thing, people, way, do, make, get, have, go, good, many*. These are lexical items, but on the borderline of grammar. (...) This is equivalent to recognising three categories rather than two: grammatical items, high-frequency lexical items, and low-frequency lexical items.” (Halliday, 1985b: 64-5)

D’acord amb aquesta observació i en la línia ja apuntada per D.P. Hayes (1988) i per G. Brown i G. Yule (1983b) (vg. el fragment citat en el punt 2.3.2) sobre l’ús de mots més o menys comuns en el discurs, he inclòs en la codificació una distinció entre els noms menys habituals o de més pes informatiu i els noms més habituals o de menys pes informatiu. Anomeno aquests darrers *noms de freqüència alta*.

Les aportacions teòriques comentades permeten d’abordar, doncs, aquesta part de la recerca amb la hipòtesi que el nombre de noms serà més alt en la prosa acadèmica que en la classe magistral, i en la classe més que en la conversa informal. A més, resulta congruent de pensar que en els dos gèneres orals hi haurà més noms de freqüència alta que en l’escrit.

#### **4.2.2- UNITATS, PROCEDIMENTS I CRITERIS DE CODIFICACIÓ**

He considerat nom qualsevol mot classificat en aquesta categoria pel diccionari general. No hi he inclòs les substantivacions d’adjectius determinatius (*els primers, les altres*), llevat del cas dels numerals usats per donar nom als anys, que sí que semblen aportar una densitat al text, tant si són usats d’una manera absoluta (*el mil nou-cents vuitanta-quatre, el 1984*), com si ho són en aposició (*l’any mil dos-cents dos, l’any 1202*).

Pel que fa a la distinció entre els noms més o menys habituals, he partit del *Diccionari de freqüències*, dirigit per J. Rafel (1996 i 1998), que estableix les freqüències d’ús dels

mots en català contemporani sobre un corpus de grans dimensions<sup>153</sup>. El procediment seguit per a aquests *noms de freqüència alta* ha estat el següent:

**1- Selecció dels 80 primers noms** de la llista total de mots del diccionari ordenats per freqüències, a partir de la més alta (apartat “Ordenació per freqüència amb dispersió i ús”, pàg. 437 i sgs.)<sup>154</sup>. La llista de noms obtinguda és la següent:

1- <i>any</i>	21- <i>aigua</i>	41- <i>pare</i>	61- <i>mà</i>
2- <i>part</i>	22- <i>casa</i>	42- <i>persona</i>	62- <i>classe</i>
3- <i>home</i> <sup>155</sup>	23- <i>senyor</i>	43- <i>sant</i>	63- <i>idea</i>
4- <i>dia</i>	24- <i>món</i>	44- <i>canvi</i>	64- <i>efecte</i>
5- <i>cosa</i>	25- <i>nom</i>	45- <i>fill</i>	65- <i>història</i>
6- <i>temps</i>	26- <i>treball</i>	46- <i>problema</i>	66- <i>compte</i>
7- <i>manera</i>	27- <i>exemple</i>	47- <i>element</i>	67- <i>art</i>
8- <i>vida</i>	28- <i>país</i>	48- <i>raó</i>	68- <i>caràcter</i>
9- <i>cas</i>	29- <i>moment</i>	49- <i>acció</i>	69- <i>camí</i>
10- <i>vegada</i>	30- <i>ciutat</i>	50- <i>figura</i>	70- <i>tipus</i>
11- <i>lloc</i>	31- <i>paraula</i>	51- <i>grup</i>	71- <i>rei</i>
12- <i>forma</i>	32- <i>sentit</i>	52- <i>valor</i>	72- <i>objecte</i>
13- <i>terra</i>	33- <i>dret</i>	53- <i>sistema</i>	73- <i>principi</i>

<sup>153</sup> El *Diccionari de freqüències* s’ha elaborat a partir del *Corpus textual informatitzat de la llengua catalana*, constituït per 3387 textos, amb un total de 52375044 mots, dels quals 29266353 corresponen a la llengua no literària, que és la que serveix de punt de referència per al present estudi. Aquest corpus ha estat dissenyat per ser representatiu dels textos escrits en llengua catalana entre 1833 i 1988.

<sup>154</sup> En el *Diccionari de freqüències* apareixen ordenats tots els mots de la llengua, sigui quina sigui la categoria gramatical. Per a la present recerca ha calgut seleccionar els noms i, en els apartats següents, els verbs, els adjectius i alguns adverbis.

<sup>155</sup> Un o una analista crítica del discurs faria notar que el mot *dona* no figura entre els vuitanta més freqüents de la llengua catalana i en trauria conclusions de tipus ideològic. Efectivament, si *home* és el tercer nom més usat, amb 37727 ocurrences, el nom *dona* se situa en el lloc 102è de la llista, amb només 8336 ocurrences. Pel que fa a la meua mostra, puc aportar dades complementàries a aquestes: el nombre total d’aparicions d’*home* en els fragments analitzats ha estat de 62, mentre que les aparicions de *dona* han estat només 6, en una proporció, per tant, de 10 és a 1. Ocorre el mateix, si bé en una proporció més moderada, amb els mots *pare*, que és el 41è de la llista, amb 12783 ocurrences, i *mare*, situat en el lloc 135è de la llista, amb 7113 ocurrences. Una relació quantitativa semblant, i fins i tot en algun cas més gran, és la que he trobat entre les formes *ell/s* i *ella/es*, *aquest/s* i *aquesta/es*, *algun/s* i *alguna/es* i, en general, entre les formes masculines i femenines dels mots gramaticals. D’altra banda, els mots *home* / *hombre* i *dona* / *mujer* són abordats per M.J. Cuenca i M. Torres (en premsa) pel que fa a l’ús que tenen com a marcadors discursius, propi sobretot de la llengua oral i més relacionat amb les funcions de les interjeccions que es tractaran en l’apartat 5.6.



14- <i>déu</i>	34- <i>relació</i>	54- <i>llibre</i>	74- <i>terme</i>
15- <i>obra</i>	35- <i>força</i>	55- <i>estudi</i>	75- <i>realitat</i>
16- <i>poble</i>	36- <i>llengua</i>	56- <i>causa</i>	76- <i>carrer</i>
17- <i>fet</i>	37- <i>hora</i>	57- <i>societat</i>	77- <i>camp</i>
18- <i>punt</i>	38- <i>església</i>	58- <i>condició</i>	78- <i>època</i>
19- <i>segle</i>	39- <i>lleï</i>	59- <i>nombre</i>	79- <i>banda</i>
20- <i>estat</i>	40- <i>esperit</i>	60- <i>cos</i>	80- <i>mena</i>

Molts d'aquests noms presenten unes característiques que permeten de considerar-los mots de pes informatiu baix, perquè tenen un significat molt general o perquè se solen usar en expressions que estan en un procés de gramaticalització avançat. Per concretar una mica més això, he assajat de fer-ne la classificació següent<sup>156</sup>:

**a/ Noms de significat amplíssim o de referència vaga<sup>157</sup>:** *cas, cosa, element, fet, manera, mena, persona, vegada.*

**b/ Noms referits a conceptes bàsics:** *acció, aigua, art, camí, camp, canvi, caràcter, carrer, casa, causa, ciutat, classe, condició, cos, déu, dret, efecte, església, esperit, estat, estudi, figura, fill, força, forma, grup, història, home, idea, lleï, llengua, llibre, lloc<sup>158</sup>, mà, món, nom, nombre, objecte, obra, país, paraula, pare, persona, poble, principi, problema, raó, realitat, rei, relació, sant, senyor, sistema, societat, temps, terme, terra, tipus, treball, valor, vida.*

**c/ Noms relatius a períodes de temps:** *any, dia, època, hora, moment, segle.*

<sup>156</sup> Val a dir que aquesta classificació té només un valor indicatiu, ja que en el *Diccionari de freqüències* no apareix especificat quin és el significat o l'ús concret de cadascun dels mots. Així, per exemple, les aparicions ressenyades del mot *classe* poden referir-se tant al significat concret d'aquest mot (sinònim d'*aula*, etc.), com al més general (sinònim de *tipus*), o les del mot *efecte* poden ser amb el significat ple (sinònim de *conseqüència*) o amb el significat afeblit, parcialment gramaticalitzat, de l'expressió *en efecte*.

<sup>157</sup> Seria difícil traçar una línia divisòria exacta entre aquests mots i els de l'apartat següent. Potser els casos més clars són *cosa, fet* i *persona*, que s'ajusten més al concepte de *superhiperònim* i al de les anomenades *proformes* de la lingüística textual (Bernárdez, 1982: 105-108). Els altres mots que he posat en l'apartat a/ s'acostarien més o menys a aquesta noció.

<sup>158</sup> I el castellanisme *puesto*, que també apareix en les converses de la mostra.

**d/ Noms que apareixen associats a verbs, amb els quals constitueixen una unitat de significat:**

- . *banda*: deixar de *banda*
- . *lloc*: tenir *lloc*
- . *part*: formar *part*, prendre *part*
- . *terme*: dur a *terme*, portar a *terme*
- . *compte*: anar amb *compte*, tenir en *compte*

**e/ Noms que apareixen en sintagmes preposicionals amb funcions diverses (circumstancials, connectors discursius, etc.):**

- |  |  |
|--|--|
| . <i>banda</i> : d'una <i>banda</i> , d'altra <i>banda</i> | . <i>canvi</i> : en <i>canvi</i>                           |
| . <i>cas</i> : en tot <i>cas</i> , si de <i>cas</i>        | . <i>causa</i> : a <i>causa</i> de, per <i>causa</i> de    |
| . <i>efecte</i> : en <i>efecte</i>                         | . <i>exemple</i> : per <i>exemple</i>                      |
| . <i>fet</i> : de <i>fet</i>                               | . <i>forma</i> : de <i>forma</i> que, en <i>forma</i> de   |
| . <i>funció</i> : en <i>funció</i> de                      | . <i>hora</i> : a l' <i>hora</i> de                        |
| . <i>lloc</i> : en <i>lloc</i> de                          | . <i>manera</i> : a <i>manera</i> de, de <i>manera</i>     |
| que  |  |
| . <i>mena</i> : una <i>mena</i> de                         | . <i>part</i> : a <i>part</i>                              |
| . <i>punt</i> : fins al <i>punt</i> que                    | . <i>relació</i> : amb <i>relació</i> a, en <i>relació</i> |
| amb  |  |
| . <i>sentit</i> : en el <i>sentit</i> que                  | . <i>vegada</i> : a la <i>vegada</i>                       |

Com es pot comprovar en aquestes llistes, els noms que apareixen amb més freqüència en el català escrit contemporani poden fer-ho per diversos motius. Així, els dos primers grups es caracteritzen per una àmplia generalització del significat. En el grup c/, els noms relatius a períodes de temps veuen afeblit el pes significatiu per l'ús sovintejat en aposicions (*el dia 26 de setembre, l'any 1948, el segle XIII*, etc.) en què informativament destaca més l'expressió numèrica que els acompanya. Els grups d/ i e/, en darrer lloc, presenten graus i tipus diversos de gramaticalització, amb una disminució o bé total desaparició del valor lexical dels mots.

Per tant, una anàlisi ràpida dels noms de la llista els mostra bastant heterogenis, amb l'única característica comuna objectiva i indiscutible de ser els noms més usats en el corpus amplíssim en el qual es basa el *Diccionari de freqüències*<sup>159</sup>.

**2- Elaboració d'una segona llista, més reduïda**, aplicant la condició que, per formar-ne part, els noms han d'aparèixer almenys un cop en algun dels textos de cadascun dels gèneres de la mostra. Els noms exclosos per aquest motiu han estat els següents:

- **Noms exclosos perquè no apareixen a cap de les converses:** *acció, art, camí, caràcter, causa, ciutat, condició, cos, dret, efecte, element, església, esperit, estat, estudi, figura, força, forma, idea, llei, llengua, llibre, mà, món, nom, nombre, objecte, país, principi, realitat, rei, relació, segle, sistema, societat, terme.*

- **Noms exclosos perquè no apareixen a cap de les classes:** *acció, aigua, art, caràcter, carrer, causa, cos, efecte, esperit, estudi, llibre, objecte, obra, pare, treball.*

---

<sup>159</sup> Per part meua, amb anterioritat a l'aparició del *Diccionari de freqüències*, havia elaborat una llista de mots prenent com a criteri no pas la freqüència sinó exclusivament l'escàs pes informatiu. Aquesta llista, preparada d'una manera totalment intuïtiva a partir de l'observació dels textos de la mostra, incloïa una seixantena de mots, la meitat dels quals han coincidit també a ser entre els 80 primers del *Diccionari de freqüències*:

- **Noms de significat general o de referència vaga:** *aspecte, base, caràcter, cas, classe* (com a sinònim de *tipus*), *cop* (com a sinònim de *vegada*), *cosa, dada, dona, element, factor, fenomen, fet, home, idea, lloc, mena* (com a sinònim de *tipus*), *persona, problema, punt* (com a sinònim d'*apartat*), *qüestió, tema, tipus, total, totalitat, tret, vegada.*

- **Noms relatius a períodes de temps:** *any, dia, època, hora, moment, període, segle.*

- **Noms que apareixen en expressions equivalents a adjectius determinatius:** *la majoria de, la meitat de* (i el col·loquialisme *mitat*), *una mica de, una minoria de, una part de, la resta de, un seguit de, una sèrie de.*

- **Noms que apareixen associats a verbs, amb els quals constitueixen una unitat de significat:** *deixar de banda, tenir lloc* (com a sinònim d'*ocórrer*), *formar part, prendre part, dur a terme, portar a terme.*

- **Noms que apareixen en sintagmes preposicionals amb funcions diverses (circumstancials, connectors discursius, etc.):** *d'una banda, d'altra banda, en canvi, en tot cas, si de cas, a continuació, en efecte, per exemple, de fet, a la fi, en fi, per fi, de forma que, en forma de, en funció de, a l'hora de, en lloc de, a manera de, de manera que, una mena de, a part, fins al punt que, en el sentit que, a grans trets, a la vegada.*

- **Noms exclosos perquè no apareixen a cap dels escrits:** *acció, aigua, art, camí, déu, església, esperit, fill, paraula, sant.*

Cal remarcar que la llista dels mots exclosos pel fet de no aparèixer en les converses és la més extensa. Molts d'aquests mots tenen un caràcter bastant formal i són propis dels discurs expositiu escrit. Això no és estrany si es té en compte que el *Diccionari de freqüències* es basa exclusivament en la llengua escrita i que caldria disposar d'una eina semblant per a la llengua oral<sup>160</sup>. La segona llista resulta, per tant, més adequada perquè evita el greuge quantitatiu que hauria representat per al gènere conversa la inclusió en el còmput d'un nombre considerable de mots que no apareixen ni una sola vegada en les converses de la mostra. Els 35 noms que integren aquesta llista són els següents:

1- <i>any</i>	13- <i>poble</i>	25- <i>raó</i>
2- <i>part</i>	14- <i>fet</i>	26- <i>grup</i>
3- <i>home</i>	15- <i>punt</i>	27- <i>valor</i>
4- <i>dia</i>	16- <i>casa</i>	28- <i>classe</i>
5- <i>cosa</i>	17- <i>senyor</i>	29- <i>història</i>
6- <i>temps</i>	18- <i>exemple</i>	30- <i>compte</i>
7- <i>manera</i>	19- <i>moment</i>	31- <i>tipus</i>
8- <i>vida</i>	20- <i>sentit</i>	32- <i>camp</i>
9- <i>cas</i>	21- <i>hora</i>	33- <i>època</i>
10- <i>vegada</i>	22- <i>persona</i>	34- <i>banda</i>
11- <i>lloc</i>	23- <i>canvi</i>	35- <i>mena</i>
12- <i>terra</i>	24- <i>problema</i>	

Com es pot veure, fins a la posició número 11 coincideixen exactament la primera llista i la segona. Sembla, per tant, que es tracta de mots tan freqüents que deuen aparèixer en qualsevol mostra una mica extensa.

Pel que fa als aspectes metodològics, els fragments de textos utilitzats per al còmput dels noms han estat de 2500 mots cadascun, d'acord amb les conclusions sobre l'extensió de

---

<sup>160</sup> A la llista intuïtiva passava el mateix: en pensar mots de significat general i expressions gramaticalitzades havia recorregut sense adonar-me'n a noms que són més propis dels registres escrits, expositius i formals.

la mostra presentades en el punt 3.5.1. El procediment emprat ha estat la codificació manual per als noms en general i la recerca automàtica amb *TACT* per als noms de freqüència alta. Els pronoms personals forts i el demostratiu neutre *això*, que he cercat, com es veurà a continuació, per a la justificació d'una part dels resultats, també han estat comptats automàticament amb *TACT*<sup>161</sup>.

He calculat els percentatges per als quadres de resultats sobre el total de mots de cada fragment i també sobre el total de noms. Aquesta segona mesura de referència és la més adequada per als noms de freqüència alta ja que, com es veurà més avall, aporta més claredat en la interpretació de les dades.

#### 4.2.3- RESULTATS I DISCUSSIÓ

Els resultats han confirmat les hipòtesis d'aquest apartat, especialment pel que fa al nombre total de noms. En els noms de freqüència alta també acaben complint-se les hipòtesis formulades, tot i que l'explicació ha de ser més matisada. Vegem-ne els resultats globals a la taula següent:

<b>TAULA 5 *</b>	<b>CONVERSA INFORMAL</b> (27254 mots)	<b>CLASSE MAGISTRAL</b> (26374 mots)	<b>PROSA ACADÈMICA</b> (26468 mots)
<b>Total noms</b>	13.4 % (3652)	19.9 % (5248)	24.5 % (6485)

(\* Percentatges sobre el total de mots i, entre parèntesis, dades absolutes d'ocurrències – Test  $\chi^2$ : probabilitat d'atzar inferior al 0.0001 %)

<sup>161</sup> Les instruccions de *TACT* per a la llista de noms de freqüència alta han estat *Usebase - F3 Complete Word-List - F9 Writing to disk*, amb la posterior recerca manual dels mots i del nombre d'ocurrències en l'ordenació alfabètica. En canvi, per a la recerca dels noms més habituals en els textos he utilitzat *TactFreq - Frequency Order*, i per als pronoms personals forts i el demostratiu neutre *això*, *Usebase - F2 Query*.

A la taula 5 es pot observar que, efectivament, el percentatge de noms ha estat progressivament més alt, d'un 13.4 % en les converses, d'un 19.9 % en les classes i d'un 24.5 % en els escrits. El contrast és notable, ja que en la conversa informal els noms són menys de la sisena part del total de mots, mentre que en la prosa acadèmica són pràcticament la quarta part. Alhora, el nombre de noms en la prosa és gairebé el doble que en la conversa (el diferencial és de l' 11.1 %). La classe magistral, per la seva part, se situa en una posició central entre els altres dos gèneres, si bé una mica més pròxima a la prosa (a 6.5 punts de les converses i a 4.6 de la prosa). D'altra banda, en una mirada superficial, les diferències entre aquests percentatges podrien semblar petites: hom esperaria resultats més espectaculars, amb diferencials del 10 per cent o més. Però, atès el gran nombre de mots gramaticals d'un text, que en tots els casos supera el 50 % (vg. § 4.6.2), el marge que queda per al contrast entre noms, verbs i adjectius és relativament petit i les diferències són realment de consideració.

Pel que fa als noms de freqüència alta, a la taula 6 s'observen unes xifres més altes en la classe i en la prosa que en la conversa, al contrari del que era esperable. Aquesta tendència aparentment contradictòria es capgira si es fa la proporció respecte del total de noms (taula 8):

<b>TAULA 6 *</b>	<b>CONVERSA INFORMAL</b> (27254 mots)	<b>CLASSE MAGISTRAL</b> (26374 mots)	<b>PROSA ACADÈMICA</b> (26468 mots)
<b>Noms de freqüència alta</b>	2.4 % (654)	2.8 % (738)	2.7 % (715)

(\* Percentatges sobre el total de mots i, entre parèntesis, dades absolutes d'ocurrències – Test  $\chi^2$ : probabilitat d'atzar de l'1.0643 %) <sup>162</sup>

<sup>162</sup> Els resultats són una mica més igualats en el còmput primerenc basat en la meua llista intuïtiva. De totes maneres, la diferència no és pas gaire gran d'una llista a l'altra:

<b>TAULA 8 *</b>	<b>CONVERSA INFORMAL</b> (3652 noms)	<b>CLASSE MAGISTRAL</b> (5248 noms)	<b>PROSA ACADÈMICA</b> (6485 noms)
<b>Noms de freqüència alta</b>	17.9 % (654)	14.1 % (738)	11.0 % (715)

(\* Percentatges sobre el total de noms i, entre parèntesis, dades absolutes d'ocurrències – Test  $\chi^2$ : probabilitat d'atzar inferior al 0.0001 %)

D'acord amb la taula 8 existeix una progressió decreixent en el percentatge de noms de freqüència alta, que representen un 17.9 % del total de noms en les converses i només un 11.0 % en la prosa acadèmica, mentre que la classe magistral torna a quedar en una posició intermèdia. Els noms de freqüència alta, per tant, tenen un paper lleugerament més destacat en els dos gèneres orals que en l'escrit.

Al costat d'aquestes dades sobre els noms que tenen una freqüència alta en el conjunt de la llengua segons el *Diccionari de freqüències*, hom es pot demanar quins han estat els noms més habituals de la mostra present, informació que oferiré a continuació. Els presento text per text, ja que el tema que tracta cadascun resulta pertinent per a l'anàlisi:

**Conversa 1 - Tema principal<sup>163</sup>**: els afers econòmics d'una família, les seves cases i les seves propietats. **Noms més freqüents**: 1- *casa*, 2- *pare*, 3- *dia*, 4- *cosa*, 5- *diner*, 6- *senyor*, 7- *terreny*, 8- *cortines*, 9- *habitació*, 10- *família*.

---

<b>TAULA 7 *</b>	<b>CONVERSA</b>	<b>CLASSE</b>	<b>ESCRIT</b>
<b>Noms de freqüència alta</b>	2.0 %	2.1 %	1.7 %

(\* Percentatges sobre el total de mots)

<sup>163</sup> En les converses es tracten temes ben diversos i pot haver-n'hi alguns que s'allarguin i ocupin més el temps dels contertulians. En les classes i en els escrits, en canvi, hi ha un tema central monogràfic.

**Conversa 2 - Tema principal:** qüestions diverses. **Noms més freqüents:** 1- *Jordi*, 2- *home*, 3- *diumenge*, 4- *gota*, 5- *cop*, 6- *dissabte*, 7- *Jaume*, 8- *lirona*, 9- *sentit*, 10- *any*.

**Conversa 3 - Tema principal:** visita turística al sud de França. **Noms més freqüents:** 1- *home*, 2- *any*, 3- *càtars*, 4- *castell*, 5- *mica*, 6- *catalans*, 7- *cosa*, 8- *estat*, 9- *dia*, 10- *història*.

**Conversa 4 - Tema principal:** la boda d'una de les filles. **Noms més freqüents:** 1- *vestit*, 2- *cosa*, 3- *església*, 4- *gent*, 5- *mica*, 6- *casa*, 7- *pati*, 8- *Carme*, 9- *casament*, 10- *despedida*.

**Conversa 5 - Tema principal:** viatges i malalties. **Noms més freqüents:** 1- *menta*, 2- *Halles*, 3- *Jordi*, 4- *àvia*, 5- *París*, 6- *regla*, 7- *carrer*, 8- *natura*, 9- *orujo*, 10- *cosa*.

**Classe individu A - Tema:** història econòmica d'Espanya en els anys cinquanta del segle XX. **Noms més freqüents:** 1- *economia*, 2- *cinquanta-nou*, 3- *mesures*, 4- *estabilització*, 5- *país*, 6- *indústria*, 7- *situació*, 8- *fet*, 9- *manera*, 10- *divises*.

**Classe individu B - Tema:** història dels inicis de l'islamisme. **Noms més freqüents:** 1- *imam*, 2- *punt*, 3- *islam*, 4- *vista*, 5- *poder*, 6- *part*, 7- *califa*, 8- *comunitat*, 9- *profeta*, 10- *xiisme*.

**Classe individu C - Tema:** història de l'esclavitud. **Noms més freqüents:** 1- *esclau*, 2- *home*, 3- *exemple*, 4- *color*, 5- *esclavitud*, 6- *cosa*, 7- *gent*, 8- *pell*, 9- *superioritat*, 10- *cas*.

**Classe individu D - Tema:** història del Brasil en el segle XIX. **Noms més freqüents:** 1- *Brasil*, 2- *liberals*, 3- *Portugal*, 4- *estat*, 5- *rei*, 6- *moment*, 7- *Pere*, 8- *segle*, 9- *monarca*, 10- *any*.

**Classe individu E - Tema:** història de Xina en els segles I i II. **Noms més freqüents:** 1- *Xiongnu*, 2- *regals*, 3- *política*, 4- *nord*, 5- *xinesos*, 6- *estat*, 7- *imperi*, 8- *quantitat*, 9- *seda*, 10- *Han*.



**Escrit individu A - Tema:** història de la construcció a Barcelona en el segle XIX. **Noms més freqüents:** 1- *promotor*, 2- *casa*, 3- *construcció*, 4- *propietari*, 5- *promoció*, 6- *propietat*, 7- *any*, 8- *Eixample*, 9- *part*, 10- *procés*.

**Escrit individu B - Tema:** història econòmica de Catalunya en el segle XVIII. **Noms més freqüents:** 1- *drets*, 2- *economia*, 3- *segle*, 4- *domini*, 5- *Mediterrània*, 6- *nord*, 7- *població*, 8- *terra*, 9- *treball*, 10- *comerç*.

**Escrit individu C - Tema:** assaig contra l'aplicació de l'anàlisi del discurs en la ciència històrica - Reflexió prospectiva sobre la marxa de l'economia mundial. **Noms més freqüents:** 1- *cosa*, 2- *text*, 3- *cultura*, 4- *història*, 5- *realitat*, 6- *anàlisi*, 7- *discurs*, 8- *estudi*, 9- *mena*, 10- *passat*.

**Escrit individu D - Tema:** història del moviment literari de la Renaixença. **Noms més freqüents:** 1- *literatura*, 2- *poesia*, 3- *llengua*, 4- *català*, 5- *procés*, 6- *Aguiló*, 7- *castellà*, 8- *classe*, 9- *cultura*, 10- *definició*.

**Escrit individu E - Tema:** història dels inicis de l'imperi xinès. **Noms més freqüents:** 1- *Qin*, 2- *imperi*, 3- *món*, 4- *estat*, 5- *regne*, 6- *text*, 7- *fet*, 8- *segle*, 9- *estat*, 10- *Han*.

Els resultats en aquestes llistes són diàfans: els noms més usats a cada fragment estan en relació directa amb el tema principal tractat, excepte alguns com *any*, *dia*, *exemple*, *cosa*, etc., que són de caràcter general i que apareixen també a la llista del *Diccionari de freqüències*. Indiscutiblement, mots com *càtars*, *divises*, *imam*, *Brasil*, *Xiongnu* o *promotors*, que són molt freqüents en aquests textos, no es poden considerar de cap manera mots de pes informatiu baix, sinó que deuen la freqüència únicament a la posició central que ocupen en el tema tractat en cada cas. També és clar que les converses, com a textos de tema general, tenen entre els més freqüents noms també més generals, mentre que les classes i els escrits, que se situen dins de l'àmbit específic dels estudis històrics, tenen com a més freqüents mots també més específics. I, encara, una característica comuna a les converses és que els noms ocupen llocs més endarrerits en la llista total de mots més freqüents, perquè tenen al davant molts mots gramaticals (els deu noms més freqüents estan entre els 100 primers mots). En les classes i, sobretot, en els escrits, els

noms ocupen posicions més avançades (en les classes estan entre els 65 primers i en els escrits entre els 50 primers).

Continuant amb l'anàlisi dels resultats, ara toca el torn d'una mirada detallada a les dades obtingudes a cadascun dels fragments estudiats. Aquesta informació és a les taules 9 i 10 (vg. pàgina següent), en la primera la corresponent a les cinc converses i en la segona la de les classes i els escrits, aparellats d'acord amb els cinc individus que els han produït. Com es pot veure en aquests taules, els fragments analitzats presenten una certa dispersió en els resultats, que és de 5.4 punts en les converses, de 3.2 punts en les classes i de 3.2 punts en la prosa, de manera que la xifra més alta de les converses (conversa 5: 17.6 %) s'aproxima a la més baixa de les classes (classe de l'individu A: 18.8 %) i la més alta de les classes (classe de l'individu D: 22.0 %) també s'aproxima a la més baixa dels escrits (escrit de l'individu C: 22.6%). Això significa que algunes converses s'han situat en un nivell de densitat nominal similar al d'algunes classes i que algunes classes han estat també pròximes a alguns escrits, sense que en cap cas, però, no s'hagi sobrepassat l'índex corresponent a l'altre gènere. I, d'altra banda, la diferència entre la conversa amb l'índex més baix (conversa 3: 9.5 %) i l'escrit amb l'índex més alt (escrit de l'individu A: 25.8 %) ha estat notable, de 16.3 punts. Així, la distribució de les dades conforma, malgrat la proximitat d'alguns resultats, unes bandes perfectament delimitades. La representació gràfica d'aquest fenomen pot veure's a la taula 11.

La taula 10 permet, a més, una comparació entre les produccions de cada individu, en la classe magistral i en la prosa acadèmica. Des d'aquest punt de vista, s'observa que la prosa presenta un percentatge sempre més alt: bastant més alt en els individus A, B i E (7.0, 5.3 i 6.7 punts de diferència, respectivament), una mica més a C (3.1 punts) i molt poc a D (0.7 punts). Així, l'individu D és l'excepció, i ben remarcable, ja que presenta uns percentatges gairebé iguals en tots dos gèneres, cosa que indica que els noms tenen un paper destacat també en les seves produccions orals formals<sup>164</sup>.

---

<sup>164</sup> Curiosament, l'individu D és de tots el qui presenta més vacil·lacions, més mots trencats i més repeticions de mots gramaticals (especialment preposicions) per fer reformulacions. Així, si en aquest aspecte —que he volgut remarcar, malgrat que no sigui centre d'atenció en aquest treball— té un caràcter més oralitzant, en el nombre de noms és l'individu que produeix una classe més pròxima a l'escrit. Aquest és un contrast que no he pogut comprendre i que demanaria recerca suplementària: potser l'explicació

<b>TAULA 9 *</b>	<b>NOMS</b>	<b>CONVERSA</b>
<b>CONVERSA 1</b>	Total noms .....	15.0 %
<b>CONVERSA 2</b>	Total noms .....	12.2 %
<b>CONVERSA 3</b>	Total noms .....	9.5 %
<b>CONVERSA 4</b>	Total noms .....	12.7 %
<b>CONVERSA 5</b>	Total noms .....	17.6 %

(\* Percentatges sobre el total de mots)

<b>TAULA 10 *</b>	<b>NOMS</b>	<b>CLASSE</b>	<b>PROSA</b>
<b>INDIVIDU A</b>	Total noms .....	18.8 %	25.8 %
<b>INDIVIDU B</b>	Total noms .....	20.4 %	25.7 %
<b>INDIVIDU C</b>	Total noms .....	19.5 %	22.6 %
<b>INDIVIDU D</b>	Total noms .....	22.0 %	22.7 %
<b>INDIVIDU E</b>	Total noms .....	18.9 %	25.6 %

(\* Percentatges sobre el total de mots)

---

seria simplement que aquests dos fenòmens no tenen per què anar lligats. (Per a més dades sobre les classes d'aquest individu, vg. § 4.5.3 i 4.10.6)

<b>TAULA 11 * - NOMS (Dispersió de les dades)</b>				
26				
25	Escrit E	Escrit B	Escrit A	<b>BANDA DE LA PROSA ACADÈMICA</b>
24				
23	Escrit C	Escrit D		
22	<b>Classe D</b>			
21				<b>BANDA DE LA CLASSE MAGISTRAL</b>
20	<b>Classe B</b>			
19	<b>Classe C</b>			
18	<b>Classe A Classe E</b>			
17	<b>Conversa 5</b>			
16				<b>BANDA DE LA CONVERSA INFORMAL</b>
15	<b>Conversa 1</b>			
14				
13	<b>Conversa 4</b>			
12	<b>Conversa 2</b>			
11				
10				
9	<b>Conversa 3</b>			

(\* Percentatges sobre el total de mots)

Les dades que he analitzat fins ara condueixen a la conclusió general que, d'acord amb M.A.K. Halliday, la llengua oral informal tendeix a un ús menor dels noms per a l'articulació del discurs. L'explicació d'aquest fet pot complementar-se, a parer meu, amb l'anàlisi de dos fenòmens importants: en primer lloc, la major utilització de pronoms en lloc dels noms corresponents i, en segon lloc, la construcció de moltes oracions amb un dels recursos de més força cohesiva, l'elisió sintàctica, especialment la de subjecte.

Per a la justificació del primer aspecte, la major utilització dels pronoms, he cercat en els textos, a tall indicatiu, les aparicions dels pronoms personals forts —taula 12— i del demostratiu neutre *això* —taula 13— (a diferència de les altres d'aquest apartat, les taules 12 i 13 estan calculades sobre fragments de 5000 mots, és a dir, sobre la totalitat de la mostra):

<b>TAULA 12 *</b>	<b>CONVERSA INFORMAL</b>	<b>CLASSE MAGISTRAL</b>	<b>PROSA ACADÈMICA</b>
<i>jo</i>	295	5	1
<i>tu</i>	62	4	0
<i>ell/a</i>	74	20	9
<i>nosaltres</i>	37	7	0
<i>vosaltres</i>	31	2	0
<i>ells/es</i>	34	9	17
<b>Total pronoms personals forts</b>	533	47	27

(\* Dades absolutes d'ocurrències)<sup>165</sup>

<sup>165</sup> Els resultats aportats per D. Biber (1988) sobre els pronoms personals són els següents:

<b>TAULA 14 *</b>	<i>Face-to-face conversation</i>	<i>Prepared speeches</i>	<i>Academic prose</i>
<i>1st person pronouns</i>	5.8	4.2	0.5
<i>2nd person pronouns</i>	3.1	0.5	0.0
<i>3rd person pronouns</i>	2.9	3.7	1.1
<i>Total personal pronouns</i>	11.8	8.4	1.6

(\* Percentatges sobre el total de mots - Taula elaborada a partir de les dades de Biber, 1988: 255, 265 i 269)

El contrast en els textos de D. Biber és molt menor, però cal pensar que hi inclou tots els pronoms personals, forts i febles, i això el fa poc comparables amb el present treball.

La taula 12 reflecteix una desproporció enorme en l'ús dels pronoms personals forts. La relació que s'estableix és d'11 a 1 entre la conversa i la classe magistral, i de 19 a 1 entre la conversa i la prosa acadèmica. De fet, aquests resultats no poden ser una sorpresa, ja que tots els tractadistes assenyalen la major presència d'aquests pronoms motivada pel caràcter interactiu de la conversa. Sí que és remarcable, en canvi, que en aquest aspecte la classe s'alineï clarament amb l'escrit. Això reforça la idea que l'aparició d'aquesta mena de pronoms no està associada al canal oral (en el sentit d'*auditiu*) sinó a la interactivitat, és a dir, a la participació de diversos interlocutors en la construcció conjunta de la conversa.

TAULA 13 *	CONVERSA <i>INFORMAL</i>	CLASSE MAGISTRAL	PROSA ACADÈMICA
<i>això</i>	200	143	56

(\* Dades absolutes d'ocurrències)

La taula 13 mostra més abundància del pronom demostratiu neutre *això* en els gèneres orals i especialment en la conversa, tot i que en una mesura menor que la que acabem de veure en els pronoms personals forts. Així, encara que en el text hi hagi també altres pronoms (els pronoms personals febles i altres pronoms determinatius), les dades de les taules 12 i 13 resulten ja indicatives, i poden explicar, almenys en part, el nombre reduït de noms que apareixen en la conversa.

Pel que fa al segon aspecte, l'elisió de subjecte, l'exemplificaré amb textos reals, al mateix temps que donaré unes mostres de la presència de noms a cada gènere. Es tracta de dos fragments, un de la conversa 1 i un de l'escrit de l'individu E (de 266 i 269 mots, respectivament), en els quals he assenyalat els noms amb color vermell i els subjectes explícits subratllats:

“MAR: amb lo qual\_  
i eren **veïns**\

i llavòrens un bon **dia**\_  
aquest **senyor**  
 els hi va oferir\_  
 de compra'ls-hi la **casa**\  
 de compra'ls-hi la **casa** i el **terreny**\  
 i s'ho van estar pensant i no sé què\  
 i els hi va dir que els hi canviaria\_  
 per un **terreny** que ell tenia a **V**\  
 IMM: ah\_  
 MAR: que era seu\  
 que els hi canviaria\_  
 i els hi donaria **diners** si ells\_  
 consideraven\_  
 oportú\_  
 que els hi havia de donar **diners**\_  
 perquè valia més la **casa** i el **terreny** aquell  
 que el **terreny** de **V**\  
 que em sembla que sí\  
 IMM: clar\  
 MAR: perquè clar\  
 era **casa** i **terreny**\  
 encara que ell ho volia per tirar-ho\  
 IMM: bueno però això no importa\  
 però ells tenien allà una **casa**\  
 MAR: exacte\  
 i aleshores el **cas** és que van\_  
 IMM: això és valorar la **finca**\_  
 i llavon-  
 la **finca**\_  
**finca** i **casa** val tant\_  
 i llavòrens\_  
 MAR: mhm\  
 IMM: això\_  
 es valora el **terreny** i la **resta** s'ha d'afegir\  
 MAR: el **cas** és que van anar a mirar el **terreny** de **V**\_  
 IMM: i els hi va agradar\  
 MAR: i van decidir que sí\  
 que se n'anaven\  
 IMM: bueno\_  
 és que allà estarien sols\  
 MAR: que se n'anaven a **V**\_  
 perquè\_  
 perquè bueno\_  
 perquè aquella **casa**\_  
ells se l'havien trobat feta\_  
 i aquesta se la farien a **mida**\  
 i se la farien al seu **gust**\_  
 i no sé què\_  
 i com que\_  
 no tenien **pressa**\_  
 IMM: i a més que són\_

MAR: perquè ademés\_  
 IMM: ah\_  
 no tenien **pressa**\  
 MAR: no\_  
 tenien\_  
**pressa**\  
 o sigui\_  
el **senyor** aquest  
 els hi deia que\_  
 IMM: haguessin pogut estar\_  
 fins que hagués acabat la **casa**\  
 MAR: sí\  
 sí\  
 o sigui que sí\_  
 RUT: sí”

(Conversa 1, pàg. 4-5, lín. 25-92)

“Aquesta **existència** prèvia de múltiples **estats** es va consolidar durant la segona **etapa** de la **dinastia Zhou**, en el **període** denominat **Primaveres i Tardors** (771-453 ane), on **uns 170** petits **principats** es van multiplicar per la gran **plana** central i les **terres** muntanyoses **que** voregen els **rius Huanghe i Yangzi**. **Les Il·lites entre ells** van anar reduint aquest **nombre** i, a **finals** del **període** dels **Regnes Combatents** (453-221 ane), **set d'ells** es dividien la major **part** del **territori**.

I això ens porta a la segona **consideració**: la **idea imperial** no era immanent al **món** xinès. Certament, la **tradició confuciana** parlava d'una **unitat** primigènica, i concebia el **passat** com una **edat d'or** en què **emperadors virtuoses** governaven per damunt de tots. Però cal no oblidar que **tots aquests textos** van ser elaborats pocs **segles** abans de la **implantació** de l'**imperi** i que **tots ells** van ser seriosament manipulats durant els **Han**, és a dir, quan l'**imperi** era ja una **realitat**. **El cert** és que a **principis** del **segle** IV ane, amb set poderosos **regnes** ben instal·lats en **porcions** geogràficament ben distintes del **territori**, la **idea d'una unitat política** era sobretot una **figura** literària. Però el **tombant de segle** marcarà una clara **diferència**: i no és **casualitat** que els primers **enunciats** de la **necessitat d'un poder únic** apareguin en els **escrits** de **Xunzi** - **el que** seria **mestre** de **Li Si** i **Han Fei** - i en l'**obra** recopilada per **Lü Buwei** - **el que** seria el primer gran **canceller** del **rei** de **Qin** -, el **Lushi Chunqiu**: **tots dos** eren **persones** properes a **Qin**, l'**estat que** realitzaria la **idea imperial**.”

(Escrit individu E, pàg. 286, lín. 13268-13289)

Amb el fragment de la conversa 1 s'exemplifica perfectament el paper destacat de l'elisió de subjecte en la construcció oracional. Així, hi he comptat 42 verbs en forma personal —susceptibles, per tant, de concordar amb un subjecte explícit— i només 17 subjectes explícits. Aquests subjectes han estat, per ordre d'aparició, els següents: *aquest senyor; ell; ells; la casa i el terreny aquell; el terreny de V.; que sí; casa i terreny; ell; això; ells; el cas; això; finca i casa; la resta; el cas; ells; el senyor aquest*. Com es pot veure, dels 17 subjectes assenyalats, n'hi ha cinc que són pronominals forts (*ell* i *ells*), dos que són el demostratiu neutre *això*, tres són substantius molt generals com *el*



*cas* o *la resta*, un és la clàusula subordinada *que sí* i només sis són sintagmes amb un cert gruix lèxic: 1- *aquest senyor*, 2- *la casa i el terreny aquell*, 3- *el terreny de V.*, 4- *casa i terreny*, 5- *finca i casa* i 6- *el senyor aquest*. I el fet més destacat és que dels 42 verbs, n'hi ha 25 amb el subjecte implícit, a base d'una successió d'oracions que comparteixen un mateix subjecte temàtic, en una progressió de *tema constant*, en la qual hi ha un *tema* (informació coneguda) que es va repetint i un *rema* (informació nova) que va canviant a cada oració (Combettes, 1983: 91-97 i Castellà, 1992: 182-183).

Per tant, és fàcil hipotetitzar que el baix índex de noms de la conversa és en part causat per una elisió de subjecte més que habitual. Una altra cosa que s'observa ja a primera vista és que la construcció lingüística de la conversa no recolza en els noms —que han estat només 34—, sinó que s'estructura bàsicament a partir dels verbs, que suporten el principal pes sintàctic i organitzen el desenvolupament de la informació (vg. més comentaris sobre aquest fenomen en els apartats següents sobre els verbs, els adjectius i les preposicions). En el segon fragment, el de l'escrit de l'individu E, es pot veure l'altre tipus d'arquitectura discursiva, la que recolza primordialment en els noms per transmetre el contingut. Com es pot observar fàcilment en els sintagmes subratllats, ha canviat profundament la naturalesa categorial dels subjectes:

*Aquesta existència prèvia de múltiples estats; uns 170 petits principats; que; Les lluites entre ells; set d'ells; això; la idea imperial; la tradició confuciana; emperadors virtuosos; tots aquests textos; tots ells; l'imperi; El cert; la idea d'una unitat política; el tombant de segle; els primers enunciats de la necessitat d'un poder únic; el que; el que; tots dos; que.*

S'observa, en primer lloc, que en general són més llargs que en la conversa. La majoria no solament estan constituïts per mots lèxics plens sinó que consisteixen en sintagmes complexos amb un o diversos noms complementats per adjectius i per sintagmes preposicionals. La situació pel que fa a l'elisió també ha canviat radicalment respecte de la conversa, ja que del total de 24 verbs del fragment, n'hi ha 19 amb subjecte explícit —i, per tant, només 5 el tenen implícit—, que només en un cas és el demostratiu neutre *això*, en quatre casos les formes de relatiu *que* o *el que*, i en tres ocasions una combinació de mots gramaticals com *set d'ells*, *tots ells* i *tots dos*. Els 11 subjectes

restants són sintagmes que han arribat fins als deu mots i han acumulat diverses paraules lexicals de significat ple.

Però, a més, els noms —i els conjunts complexos de noms— no apareixen només exercint la funció de subjecte, sinó que són a tot arreu: com a complements d'objecte directe i com a complements circumstancials de menes diverses, fins a arribar a un total de 67, quasi el doble que en el fragment de conversa. És a dir, en aquest text els noms agafen el protagonisme i són els principals portadors del significat i articuladors de la informació. Les tornes han canviat i l'escrit ha adoptat un estil predominantment nominal.

#### **4.2.4- CONCLUSIONS DE L'APARTAT**

En resum, els resultats de la recerca en aquest apartat indiquen que els noms són significativament més abundants en la prosa acadèmica que en la classe magistral i més en la classe magistral que en la conversa informal, en una progressió que va de l'escrit a l'oral, de la formalitat a la informalitat i del monòleg a la interacció. Els noms de freqüència alta tendeixen a aparèixer relativament més en la conversa que en la prosa, si bé presenten unes xifres menys contundents. La distribució de les dades segons els textos mostra una certa dispersió, però en conjunt els resultats s'organitzen en unes bandes clares, amb alguns índexs pròxims entre gèneres que en cap cas arriben a superposar-se.

La menor presència de noms en la conversa queda explicada en part per un índex elevat de pronoms personals forts i, de forma més moderada, del demostratiu neutre *això*, i pel recurs generalitzat a l'elisió de subjecte en sèries d'enunciats amb progressió de tema constant. Tot plegat confirma que, com era d'esperar i com refermaré amb més dades en els apartats següents, la prosa acadèmica s'elabora amb un estil més nominal que la conversa informal, i que la classe magistral ocupa una posició intermèdia entre els altres dos gèneres.

## 4.3- ELS VERBS

### 4.3.1- FONAMENTACIÓ CONCEPTUAL I HIPÒTESIS

Respecte als verbs no es pot afegir gaire a les recerques que ja he comentat en l'apartat anterior per al comportament conjunt de noms i verbs. D'acord amb la majoria d'autors, ha d'haver-hi més verbs en la llengua oral que en l'escrita. Les dades aportades per D. Biber (1988) sobre aquesta qüestió són les següents<sup>166</sup>:

<b>TAULA 15 *</b>	<i>Face-to-face conversation</i>	<i>Prepared speeches</i>	<i>Academic prose</i>
<i>Past tense</i>	3.7 %	4.8 %	2.2 %
<i>Present tense</i>	12.8 %	7.0 %	6.4 %
<i>Infinitives</i>	1.4 %	1.6 %	1.3 %
<i>Gerunds</i>	0.5 %	0.5 %	0.8 %
<i>Total verbs</i>	18.4 %	13.9 %	10.7 %

(\* Percentatges sobre el total de mots - Taula elaborada a partir de les dades de Biber, 1988: 255, 265 i 269)

---

<sup>166</sup> Les xifres sobre els verbs són presentades per D. Biber en el seu treball d'una manera fragmentada, en quatre temps verbals diferents, i incompleta, ja que hi falten els de futur —sense que aquest fet sigui justificat per l'autor. Per això, les he agrupat i en dono la suma global, que no apareix enlloc a les seves taules. Per comparar aquestes dades amb les meves caldrà tenir en compte l'absència dels temps de futur en la xifra global de D. Biber, tot i que, atesa la baixa freqüència d'aquest temps, el buit deu ser poc rellevant.

Així, del treball d'aquest autor, en resulta una progressió descendent en l'ús dels verbs des de la conversa fins a la prosa, passant per l'oral preparat (*Prepared speeches*), que hi ocupa una posició intermèdia<sup>167</sup>.

Per la seva banda, P. Fraisse i M. Breyton (1959) i C. Gelpí (1998) posen l'èmfasi no exactament en el nombre de verbs sinó en la proporció entre els verbs i els adjectius presents en un text<sup>168</sup>. Aquests autors ho expliquen de la manera següent:

“Le rapport verbes / adjectifs est très nettement plus grand dans l'expression orale que dans l'expression écrite, ce qui signifie que le langage oral est plus simplifié, le verbe étant toujours nécessaire, l'adjectif apportant un enrichissement plus gratuit à l'expression.” (Fraisse i Breyton, 1959: 70)<sup>169</sup>

“(3) El tenor condiciona el nombre d'adjectius en relació amb el nombre de verbs: si el tenor és informatiu, el nombre de verbs serà superior al nombre d'adjectius. I el mode complementa el tenor, de manera que el text no hauria de ser escrit i planificat.” (Gelpí, 1998: 139-140)

El segon enunciat de la citació de C. Gelpí pot resultar una mica críptic; d'acord amb el conjunt del seu treball, l'he interpretat de la manera següent: “Perquè el nombre de verbs sigui superior al d'adjectius, el tenor informatiu i el mode oral seran els més avinents”.

---

<sup>167</sup> W.N. Francis i H. Kuèera (1982), en el treball ja esmentat per als noms, troben en el conjunt de la llengua escrita un 18.2 % de verbs, xifra que difereix de la de D. Biber i, també, de la meua, que serà del 9.4 %. L'explicació més immediata d'aquestes diferències radica en el fet que W.N. Francis i H. Kuèera compten dins de la categoria verb “*all verbs, including auxiliaries and modals*” (1982: 547), i el nombre d'auxiliars i de modals en anglès és molt alt, prou com per explicar la discrepància en els resultats.

<sup>168</sup> En la present recerca he estat més interessat, d'acord amb M.A.K. Halliday, en la proporció verbs / noms, més bàsica en la construcció del text i més explicativa del contrast entre l'estil nominal i l'estil verbal. Els adjectius, des d'aquest punt de vista, són una categoria ambivalent, ja que poden desenvolupar una funció atributiva o predicativa, al servei d'una construcció sintàctica més centrada en els noms o en els verbs. (vg. § 4.4.1).

<sup>169</sup> En el punt 2.3.2, ja he esmentat aquests autors entre els qui destaquen la simplicitat de la llengua oral. Les idees que presenten en aquest fragment crec que són un bon exemple de la subjectivitat i de la poca consistència amb què a vegades s'ha tractat la qüestió. La seva argumentació m'ha suggerit les preguntes següents: per què un text amb més verbs que adjectius s'ha de considerar més simple? ¿Realment, el fet que els verbs siguin “*toujours nécessaires*” en l'oració provoca que l'ús sigui menys complex? ¿Qualsevol adjectiu pot suposar “*un enrichissement plus gratuit*” o això depèn del tipus d'adjectiu, de l'ús, de la intenció, del tipus de text, etc.? La resposta a aquestes preguntes, que supera l'abast del meu treball, podria fer trontollar la posició defensada per P. Fraisse i M. Breyton.

Així, encara que els primers ho relacionin amb el contrast oral / escrit i que C. Gelpí ho atribueixi al tenor, hi ha una coincidència a afirmar que en la llengua oral hi haurà més verbs que adjectius, amb la precisió de C. Gelpí que això ocorre en l'oral quan el tenor és informatiu<sup>170</sup>.

Pel que fa als mots de freqüència alta, el criteri de M.A.K. Halliday que he aplicat per als noms també serà vàlid aquí, de manera que esperaré de trobar més verbs de freqüència alta —o de pes lexical baix— en la conversa informal i menys en la prosa acadèmica, ja que aquesta mena de verbs contribueixen a alleugerir la densitat del text. La classe magistral, lògicament, hauria d'ocupar, també en aquest tret, una posició intermèdia.

#### 4.3.2- UNITATS, PROCEDIMENTS I CRITERIS DE CODIFICACIÓ

En aquest treball, he codificat els verbs en quatre grups segons la forma, d'acord amb els criteris següents:

**A/ Formes verbals conjugades**, simples i compostes. He considerat formes compostes les integrades per l'auxiliar *haver* + participi i per l'auxiliar *vaig, vas* (o *vares*), *va, vam* (o *vàrem*), *vau* (o *vàreu*), *van* (o *varen*) + infinitiu.

**B/ Perífrasis verbals**, constituïdes per un verb modal i una forma no personal del verb. Atès que no es disposa d'una única llista de verbs modals comuna a totes les gramàtiques, he seleccionat els casos següents: *caler* + infinitiu, *deure* + infinitiu, *haver de* + infinitiu, *poder* + infinitiu, *soler* + infinitiu, *voler* + infinitiu, *estar* + gerundi i *anar* + gerundi. Els he codificat considerant tota la perífrasi com un sol verb; per tant, aquests infinitius i gerundis no els he comptat en els apartats C i D.

---

<sup>170</sup> Com es veurà en aquest apartat i en el següent, els resultats confirmaran la primera afirmació, que en la llengua oral hi ha més verbs que adjectius i en la llengua escrita hi ha més adjectius que verbs (encara que, en aquest cas, amb poca diferència). La segona, la de C. Gelpí (1998) quedarà especialment confirmada pel comportament de la classe magistral, que presenta un 13.9 % de verbs i un 5.8 % d'adjectius i que es tracta d'un discurs oral amb una intenció clarament informativa.

**C/ Infinitius subordinats** en construcció de subordinada substantiva o de subordinada adverbial (de temps amb la preposició *en*, de finalitat amb les preposicions *per* o *per a*, etc.).

**D/ Gerundis subordinats** en construcció de subordinada adverbial de qualsevol mena (de temps, de manera, etc.).

De la mateixa manera que per als noms, he cercat en el *Diccionari de freqüències* (Rafel, 1996) els verbs més habituals. La llista que n'he obtingut està formada pels deu verbs següents<sup>171</sup>:

- |                       |                  |
|-----------------------|------------------|
| 1- <i>ser / ésser</i> | 6- <i>dir</i>    |
| 2- <i>haver</i>       | 7- <i>anar</i>   |
| 3- <i>fer</i>         | 8- <i>donar</i>  |
| 4- <i>poder</i>       | 9- <i>estar</i>  |
| 5- <i>tenir</i>       | 10- <i>veure</i> |

Per a la present recerca, he fet dos canvis en la selecció. En primer lloc, l'entrada *haver* del diccionari de freqüències inclou indistintament els diversos usos gramaticals i lexicals d'aquest verb: a/ el més habitual, com a auxiliar en temps compostos (pretèrit indefinit, pretèrit plusquamperfet, pretèrit anterior, futur compost i condicional compost en el mode indicatiu i pretèrit perfet i pretèrit plusquamperfet en el mode subjuntiu, més el gerundi compost i l'infinitiu compost); b/ com a verb locatiu i existencial, sempre acompanyat de *hi* en la forma fossilitzada *haver-hi* (*hi ha*, *hi havia*, *hi haurà*); i c/ com a verb lèxicament ple, més rar i obsolet, sinònim de *tenir* (*la guineu, quan no les pot haver, diu que són verdes*). D'aquests tres usos he seleccionat el segon, *haver-hi*, en la hipòtesi que és una fórmula descriptiva bàsica que hauria de ser més corrent en la

---

<sup>171</sup> Pot sorprendre que n'hagi seleccionat només deu, ja que en el cas dels noms he treballat amb una llista molt més llarga. La raó és estadística: els 80 primers noms han assolit un percentatge màxim del 2.8 % respecte del total de mots i del 17.9 % respecte dels noms del fragment, mentre que, com es veurà a continuació, els deu primers verbs arriben a un 8.9 % respecte del total de mots, que suposa fins al 50.9 % dels verbs. Amb aquests percentatges, si hagués allargat més la llista, els verbs seleccionats aviat haurien estat la gran majoria de les ocurrences verbals de cada fragment i els índexs obtinguts no haurien permès gaire contrastació.

llengua oral que en l'escrita. La consideració de l'ús com a auxiliar, en canvi, consistiria a la pràctica en el còmput d'un conjunt de temps verbals, la majoria de passat, que no he considerat per a la recerca<sup>172</sup>. L'ús lèxic ple, en darrer lloc, l'he descartat per la freqüència nul·la que té a la mostra.

El segon canvi és el del verb *anar*, que en la setena posició del *Diccionari de freqüències* consisteix exclusivament en les formes d'auxiliar del pretèrit perfet perifràstic, mentre que el mateix verb amb significat lèxic ple ocupa la posició dotzena. Per això he exclòs *anar* de la llista i hi he afegit el verb *trobar*, que és el que apareix a continuació de *veure* en l'ordre freqüencial.

Amb el verb *poder*, en canvi, no han calgut les precaucions que he tingut amb *haver* i amb *anar*, ja que tant en l'ús modal com en el predicatiu aquest verb conserva el mateix significat. Per tant, no ha calgut ni distingir-ne els usos ni excloure'l de la llista. D'acord amb aquest conjunt de criteris, els deu verbs resultants de la selecció són els següents<sup>173</sup>:

- |                       |                   |
|-----------------------|-------------------|
| 1- <i>ser / ésser</i> | 6- <i>dir</i>     |
| 2- <i>haver-hi</i>    | 7- <i>donar</i>   |
| 3- <i>fer</i>         | 8- <i>estar</i>   |
| 4- <i>poder</i>       | 9- <i>veure</i>   |
| 5- <i>tenir</i>       | 10- <i>trobar</i> |

---

<sup>172</sup> El motiu que m'ha portat a no comptar els temps verbals com a tret rellevant està relacionat amb la naturalesa dels gèneres estudiats. Els textos de tema històric solen combinar la narració i l'exposició, de manera que és habitual de trobar-hi verbs en passat (i en present històric o, fins i tot, en futur), cosa que no succeeix en la conversa, que pot ser narrativa o no, segons el tema tractat. Per tant, el contrast entre les dades no s'hauria pogut atribuir a la distinció oral / escrit que ens ocupa, sinó al tipus de text predominant, que és un tret que escapa de l'àmbit de la present recerca.

<sup>173</sup> De la mateixa manera que per als noms, també havia treballat amb una llista intuïtiva abans de la publicació del *Diccionari de freqüències*. En aquest cas, la meua llista era encara més semblant a la del diccionari: 1- *ser / ésser*, 2- *haver-hi*, 3- *fer*, 4- *tenir*, 5- *dir*, 6- *estar*. Els resultats globals, en canvi, han estat diferents, perquè la inclusió de quatre verbs més, amb una freqüència bastant alta, ha fet variar els percentatges.

El verb *ésser* / *ser* ha servit, a més, per al càlcul del nombre de verbs copulatius i predicatius, tret per al qual he prescindit dels altres verbs copulatius<sup>174</sup>. El fet de basar el còmput dels verbs copulatius només en aquest verb l'argumentava a J.M. Castellà (1996c) de la manera següent:

“També s’observa que la gran majoria de les ocurrències de verbs copulatius corresponen al verb *ser* (CM: 150 de 163; CLL: 71 de 83), i que, alhora, els usos intransitius de *ser* són més aviat escassos (CM: 7; CLL: 5). El poc pes d’aquestes quantitats avala la possibilitat de prescindir, en una recerca posterior, de la codificació detallada d’aquests verbs i d’aquestes funcions, i de considerar, a efectes estadístics, la xifra global d’aparicions del verb *ser* com l’indicador general de presència de verbs copulatius en un text (el marge màxim d’error seria del 2.7% respecte del total de verbs, quantitat que no sembla significativa, especialment en aquest cas, ja que els verbs copulatius s’oposen a una única categoria, els predicatius, que ocupen tot el percentatge restant).” (Castellà, 1996c: 70)

Així, queda assegurat que el nombre d’aparicions de verbs copulatius altres que *ser* és bastant reduït (cal tenir en compte, per exemple, que a les ocurrències del verb *estar* de la taula s’hi han de restar les que són d’usos intransitius i les que són en perífrasis verbals) i que aquest nombre es compensa aproximadament amb el d’usos intransitius del verb *ser*. Per tant, he descartat tant les unes com les altres i he utilitzat únicament el verb *ser* com a mesura global.

Pel que fa a la metodologia, he treballat sobre fragments de 2500 mots, d’acord amb les conclusions del punt 3.6.1. He fet una codificació manual per al còmput dels verbs en general i una recerca automàtica amb *TACT* per als verbs de freqüència alta<sup>175</sup>.

---

<sup>174</sup> D’acord amb la concepció més estesa, són copulatius aquells verbs que van acompanyats d’un complement atribuït que es pronominalitza normativament amb el pronom feble *ho*, amb *el*, *la*, *els*, *les* quan l’atribuït és determinat i, excepcionalment, amb el pronom *en*. No són copulatius, per tant, una llista llarga de verbs que duen —o poden dur— un complement predicatiu que pronominalitza amb *hi* i, excepcionalment, amb *en* (*tornar-se*, *anar*, *caminar*, *fer-se*, *dir-se*, etc.). Segons aquest criteri, els verbs copulatius es redueixen a quatre: *ser*, *estar*, *semblar* (amb el valencià *parèixer*) i *esdevenir*.

<sup>175</sup> Les instruccions de *TACT* per a la llista de verbs de freqüència alta han estat *Usebase - F3 Complete Word-List - F9 Writing to disk*, amb la posterior recerca manual dels verbs i del nombre d’ocurrències en l’ordenació alfabètica.



### 4.3.3- RESULTATS I DISCUSSIÓ

Els resultats obtinguts sobre els verbs han coincidit àmpliament amb els esperats. En primer lloc, presentaré les xifres més generals a la taula 16:

<b>TAULA 16 *</b>	<b>CONVERSA INFORMAL</b> (27254 mots)	<b>CLASSE MAGISTRAL</b> (26374 mots)	<b>PROSA ACADÈMICA</b> (26468 mots)
<b>Total verbs</b>	17.5 % (4769)	13.9 % (3666)	9.4 % (2488)

(\* Percentatges sobre el total de mots i, entre parèntesis, dades absolutes d'ocurrències – Test  $\chi^2$ : probabilitat d'atzar inferior al 0.0001 %)

D'acord amb les hipòtesis, el nombre de verbs ha estat més alt en la conversa informal (17.5 %), a continuació en la classe magistral (13.9 %) i l'índex més baix l'ha presentat la prosa acadèmica (9.4 %). Com es podia suposar, la classe s'ha situat en una posició intermèdia, lleugerament més pròxima a la conversa (a 3.6 punts) que a la prosa (a 4.5 punts)<sup>176</sup>.

Com en el cas dels noms, la diferència entre la conversa informal i la prosa acadèmica és notable (8.1 punts), de manera que els verbs de la conversa són gairebé el doble i la relació aproximada entre gèneres és de 2 a 1.5 (conversa / classe) i d'1.5 a 1 (classe / prosa). Malgrat això, aquests 8.1 punts deixen un marge de 3 respecte dels 11.1 que hi havia en el cas dels noms. Així, la menor presència de noms en la conversa no es

---

<sup>176</sup> Els resultats obtinguts per als verbs podrien no ser generalitzables a qualsevol exemplar de prosa acadèmica, ja que els textos de la mostra són de tema històric i, com he dit, això fa que hi abundi més el tipus textual narratiu. Caldria comprovar, per tant, en una recerca ulterior si el nombre de verbs en els textos narratius és menor o, probablement, major que en altres tipus de text, i treure'n les conclusions pertinents.

justifica només amb la presència de més verbs, sinó que s'haurà d'explicar amb altres categories (vg. § 4.6.2).

Els resultats, segons la classificació morfològica anunciada de les formes verbals, són els que s'especifiquen a la taula següent:

<b>TAULA 17 *</b>	<b>CONVERSA INFORMAL</b> (27254 mots)	<b>CLASSE MAGISTRAL</b> (26374 mots)	<b>PROSA ACADÈMICA</b> (26468 mots)
<b>Formes personals</b>	14.5 %	10.4 %	6.8 %
<b>Perífrasis verbals</b>	1.5 %	1.0 %	0.8 %
<b>Infinitius</b>	1.3 %	2.3 %	1.5 %
<b>Gerundis</b>	0.1 %	0.2 %	0.3 %

(\* Percentatges sobre el total de mots)

Pel que fa a les dades d'aquesta taula, en primer lloc cal destacar la semblança que tenen amb la taula 15, en què he mostrat els resultats de D. Biber per a l'anglès. Sense coincidir exactament en cap dels índexs, es veu una gran similitud en la majoria de les línies globals, que resumeixo a continuació:

– Les formes personals del verb presenten una progressió similar a la de les dades globals, amb un nombre menor de verbs a mesura que es va cap a l'escrit.

– Les perífrasis verbals tenen la mateixa evolució, amb xifres progressivament més baixes, de manera aproximadament proporcional a les dades globals.

– Els infinitius, en canvi, són més freqüents en la classe magistral i presenten índexs semblants en la conversa informal i en la prosa acadèmica. L'explicació d'aquest resultat

s'haurà de trobar en el capítol sintàctic d'aquest treball, en l'apartat de la subordinació (vg. § 5.2.3).

– Els gerundis tenen una presència ben minsa a tots tres gèneres. El fet que el nombre augmenti progressivament en direcció a la prosa acadèmica, juntament amb la reducció del total de verbs en aquest gènere, fa que els gerundis siguin més habituals en la prosa en una proporció de més del quàdruple. Malgrat això, la quantitat de gerundis en els textos és tan petita que es demostra que tenen un paper quantitativament molt secundari en la construcció del discurs en català. En el treball de D. Biber (1988), en canvi, els gerundis tenen uns índexs bastant més alts, probablement a causa de la intervenció d'aquesta forma verbal en un temps tan freqüent en anglès com el present continu (*they are writing, I'm going*).

Pel que fa als verbs de freqüència alta, els resultats han estat els que s'indiquen a les taules 18 i 19 de la pàgina següent. Com es pot observar a la taula 18, la tendència ha estat ben nítida: aquesta mena de verbs s'usen més en la llengua oral que en l'escrita. Alhora, si mirem la taula 19 (vg. pàgina següent), el càlcul respecte dels mots de la mateixa categoria torna a aportar més llum per a la interpretació. Respecte del total de verbs, els de freqüència alta són la meitat tant en la conversa (50.9 %) com en la classe (49.6 %) i són només una mica més de la quarta part en l'escrit (29.8 %). Per tant, la classe magistral s'ha alineat amb la conversa informal en la tria de verbs de pes lexical baix per a la transmissió de la informació.

<b>TAULA 18 *</b>	<b>CONVERSA INFORMAL</b> (27254 mots)	<b>CLASSE MAGISTRAL</b> (26374 mots)	<b>PROSA ACADÈMICA</b> (26468 mots)
<b>Verbs de freqüència alta</b>	8.9 % (2426)	6.9 % (1820)	2.8 % (741)

(\* Percentatges sobre el total de mots i, entre parèntesis, dades absolutes d'ocurrències – Test  $\chi^2$ : probabilitat d'atzar inferior al 0.0001 %)

<b>TAULA 19 *</b>	<b>CONVERSA INFORMAL</b> (4769 verbs)	<b>CLASSE MAGISTRAL</b> (3666 verbs)	<b>PROSA ACADÈMICA</b> (2488 verbs)
<b>Verbs de freqüència alta</b>	50.9 % (2426)	49.6 % (1820)	29.8 % (741)

(\* Percentatges sobre el total de verbs i, entre parèntesis, dades absolutes d'ocurrències – Test  $\chi^2$ : probabilitat d'atzar inferior al 0.0001 %)

Tot i aquests índexs generals, els verbs de freqüència alta han mostrat comportaments diversos. La taula 20 ofereix els resultats desglossats verb per verb:

<b>TAULA 20 *</b>	<b>CONVERSA INFORMAL</b>	<b>CLASSE MAGISTRAL</b>	<b>PROSA ACADÈMICA</b>
<b>1- ser / ésser</b>	763	891	348
<b>2- haver-hi</b>	82	94	24
<b>3- fer</b>	346	118	76
<b>4- poder</b>	102	114	87
<b>5- tenir</b>	210	146	76
<b>6- dir</b>	501	261	27
<b>7- donar</b>	46	22	41
<b>8- estar</b>	172	100	26
<b>9- veure</b>	160	60	19
<b>10- trobar</b>	44	14	17
<i>Totals</i>	2426	1820	741

(\* Dades absolutes d'ocurrències - Test  $\chi^2$ : probabilitat d'atzar del 6.6018 % en el verb *poder*, del 0.1334 % en el verb *donar*, i inferior al 0.0001 % en tots els altres i en els totals)

De les dades d'aquesta taula es desprenen diversos patrons de comportament:

. **Patró A:** Els verbs *fer, tenir, dir, estar i veure* són més freqüents en la conversa informal, a continuació en la classe magistral i, en darrer lloc, en la prosa acadèmica i mantenen, per tant, la tendència de les xifres globals.

. **Patró B:** Els verbs *donar i trobar* són també més freqüents en la conversa informal, però a continuació ho són en la prosa acadèmica i, en darrer lloc, en la classe magistral.

. **Patró C:** Els verbs *ser i haver-hi* són més habituals en la classe magistral, a continuació en la conversa informal i, en darrer lloc, en la prosa acadèmica. (El verb *poder* també figuraria en aquest grup, però la xifra de probabilitat d'atzar que té, superior al 5 %, fa aconsellable de deixar-lo de banda.)

El patró A no necessita explicació, ja que és idèntic al comportament general i coincideix amb la hipòtesi inicial. El patró B és més sorprenent, però es dona en verbs amb un nombre d'aparicions bastant petit, cosa que en relativitza la importància en el conjunt.

El patró C és, en canvi, més rellevant, especialment perquè s'hi ajusta el verb *ser*, que és, de lluny, el més usat a tots tres gèneres i el més freqüent en el català escrit contemporani segons el *Diccionari de freqüències*. A més, ja he dit que el verb *ser*, com a principal verb copulatiu, és la mesura més adequada de la proporció d'oracions atributives i d'oracions predicatives en el text. Els resultats pel que fa a aquest aspecte s'especifiquen a la taula 21 i refermen les tendències que ja s'apunten en les dades absolutes. Així, els verbs copulatius tenen un paper clarament destacat en la classe magistral, que arriba al 24.3 %, la quarta part del total, en contrast amb el 16.0 % de la conversa informal i amb el 14.0 % de la prosa acadèmica.

L'explicació del baix índex de la prosa acadèmica ha de trobar-se en la tendència a la precisió lèxica que té la llengua escrita. Per la mateixa naturalesa de l'acte comunicatiu

diferit, l'escriptor disposa de temps per triar els verbs més adequats i de valor lèxic més complet per a cada enunciat, i pot descartar opcions més genèriques com el verb *ser*. En la conversa informal, el percentatge és relativament similar al de la prosa, però el nombre absolut d'ocurrències és molt superior (763 en la conversa informal per 348 en la prosa acadèmica):

<b>TAULA 21 *</b>	<b>CONVERSA INFORMAL</b> (4769 verbs)	<b>CLASSE MAGISTRAL</b> (3666 verbs)	<b>PROSA ACADÈMICA</b> (2488 verbs)
<b>Verbs copulatius</b>	16.0 % (763)	24.3 % (891)	14.0 % (348)
<b>Verbs predicatius</b>	84.0 % (4006)	75.7 % (2775)	86.0 % (2140)

(\* Percentatges sobre el total de verbs i, entre parèntesis, dades absolutes d'ocurrències – Test  $\chi^2$ : probabilitat d'atzar inferior al 0.0001 %)

Cal explicar de manera especial, per tant, el fet que sigui la classe magistral i no la conversa informal la que presenti una proporció més elevada de verbs copulatius. La justificació més plausible pot raure en la naturalesa expositiva i monologada de la classe magistral, en contrast amb el caràcter dialogat de la conversa. El verb *ser* és arquetípic en l'explicació i en la descripció i, usat en tercera persona, especialment del singular, és la marca essencial de l'afirmació científica com a presentació pretesament neutra de la certitud dels enunciats (Castellà, 1992: 228-235). A més, el verb *ser* apareix en algunes construccions que semblen ser típiques de la llengua oral formal. Vegem-ho tot plegat exemplificat en el següent fragment de la classe de l'individu A:

“si les exportacions es mantenen estables\_  
quan el pib està augmentant e fortament\_  
això el què vol dir és doncs que la propensió a exportar de l'economia espanyola\_  
doncs declina\_  
i que de fet doncs el què vol dir doncs és que la capacitat exportadora espanyola\_  
doncs ee és ee és molt reduïda\

i aquí hi ha el punt crític perquè ee aquesta incapacitat per augmentar les exportacions\_  
la incapacitat per augmentar les exportacions\_  
conduirà a un a un dèficit comercial\_  
que resu que serà un dèficit insostenible\”

(Classe individu A, pàg. 172, lín. 8135-8147)

La construcció destacada en verd a la tercera línia, que es repeteix a la sisena, rarament apareixeria en la llengua escrita —que presentaria expressions del tipus *això mostra que* o *aquest fet indica que*, sense el verb *ser*—, però tampoc no es trobaria en l’oral col·loquial. Aquesta construcció es manifesta, més aviat, com el resultat d’un esforç per a la transmissió d’una informació més complexa o més abstracta. En el tercer destacat en verd, el verb *ser* podria aparèixer per escrit o bé ser substituït per un de predicatiu, amb el significat més ple (*es manifesta més reduïda* o *s’evidencia més reduïda*). La substitució sembla que s’apunti en el darrer destacat, en què hi ha el que sembla un verb predicatiu interromput (*resu* podria ser *resultarà*) i finalment s’opta, una vegada més, pel copulatiu *ser*. La classe magistral es mostra, doncs, com un discurs especialment inclinat a l’ús d’aquest verb.

El verb *haver-hi*, d’altra banda, ha estat més abundant en el discurs oral que en l’escrit, sense que la diferència entre la classe i la conversa (només de 6 ocurrències) sembli gaire significativa. Un comentari similar es pot fer per a *poder*, i més perquè en aquest cas la diferència mínima (de 6 i de 7 ocurrències) és entre tots tres gèneres. I, en tots dos, tant *haver-hi* com *poder*, el nombre total d’aparicions és baix i no permet de fer comparacions amb una base tan sòlida com per als altres.

Un altre aspecte interessant pel que fa als verbs de freqüència alta és la comparació amb les dades que forneix el *Diccionari de freqüències* per als mateixos mots. Aquest contrast, que es pot veure a la taula 22, situa els resultats d’aquesta obra entre els de la prosa acadèmica i els de la classe magistral<sup>177</sup>. De fet, resulta congruent que l’amplíssim corpus que serveix de base al diccionari, que inclou tant textos específics com, sobretot,

---

<sup>177</sup> A la taula 22, no hi figura *haver-hi* perquè, com ja he explicat, les xifres del *Diccionari de freqüències* recullen tots els usos del verb *haver*, incloent-hi el d’auxiliar en temps compostos i, per tant, no són en absolut comparables a les meves. Això i la pèrdua de decimals en els arrodoniments és el que explica que les xifres totals d’aquesta taula difereixin en dues o tres dècimes de les de la taula 18.

textos generals, ocupi una posició entre els escrits, d'un caràcter més específic, i les classes que, tot i ser d'un grau d'especificitat similar, estan produïdes amb un grau d'espontaneïtat més gran. En les classes, per la condició que tenen de textos produïts sobre la marxa, sense temps per a un control del lèxic tan estricte com en la llengua escrita, els verbs utilitzats són més generals. La conversa informal, d'altra banda, es desmarca clarament dels resultats del *Diccionari de freqüències*, confirma el caràcter escrit d'aquesta obra i referma la necessitat que ja he esmentat de disposar d'instruments semblants elaborats sobre la base de corpus orals.

<b>TAULA 22 *</b>	<b>CONVERSA INFORMAL</b> (27254 mots)	<b>CLASSE MAGISTRAL</b> (26374 mots)	<b>DICCIONARI DE FREQUÈN.</b> (29266353 mots)	<b>PROSA ACADÈMICA</b> (26468 mots)
<b>1- ser</b>	2.79 %	3.34 %	1.99 %	1.33 %
<b>3- fer</b>	1.26 %	0.44 %	0.52 %	0.27 %
<b>4- poder</b>	0.37 %	0.43 %	0.44 %	0.32 %
<b>5- tenir</b>	0.77 %	0.55 %	0.41 %	0.27 %
<b>6- dir</b>	1.81 %	1.00 %	0.30 %	0.10 %
<b>7- donar</b>	0.16 %	0.08 %	0.18 %	0.14 %
<b>8- estar</b>	0.64 %	0.37 %	0.18 %	0.10 %
<b>9- veure</b>	0.58 %	0.23 %	0.16 %	0.11 %
<b>10- trobar</b>	0.16 %	0.05 %	0.15 %	0.11 %
<b>Totals</b>	8.54 %	6.49 %	4.70 %	2.75 %

(\* Percentages sobre el total de mots)



Com abans per als noms, en aquest apartat també presento uns resultats detallats per a cadascun dels fragments inclosos en l'estudi, per tal de calibrar la dispersió de les dades, que es poden veure en les taules 23 i 24:

<b>TAULA 23 *</b>	<b>VERBS</b>	<b>CONVERSA</b>
<b>CONVERSA 1</b>	Total verbs .....	16.9 %
<b>CONVERSA 2</b>	Total verbs .....	17.7 %
<b>CONVERSA 3</b>	Total verbs .....	17.0 %
<b>CONVERSA 4</b>	Total verbs .....	19.3 %
<b>CONVERSA 5</b>	Total verbs .....	16.8 %

(\* Percentatges sobre el total de mots)

<b>TAULA 24 *</b>	<b>VERBS</b>	<b>CLASSE MAGISTRAL</b>	<b>PROSA ACADÈMICA</b>
<b>INDIVIDU A</b>	Total verbs .....	13.8 %	10.2 %
<b>INDIVIDU B</b>	Total verbs .....	12.9 %	7.5 %
<b>INDIVIDU C</b>	Total verbs .....	15.5 %	10.9 %
<b>INDIVIDU D</b>	Total verbs .....	11.0 %	9.2 %
<b>INDIVIDU E</b>	Total verbs .....	16.2 %	9.3 %

(\* Percentatges sobre el total de mots)

D'acord amb el que s'observa en les taules 23 i 24, el grau de dispersió en el cas dels verbs és similar al dels noms. Hi ha unes bandes bastant àmplies (2.5 punts en les converses, 5.2 punts en les classes i 3.4 punts en els escrits) i una gran proximitat entre els extrems de cada banda: l'escrit amb l'índex més alt (escrit de l'individu C: 10.9 %) és a tocar de la classe amb el percentage més baix (classe de l'individu D: 11.0 %) i la classe amb l'índex més alt (classe de l'individu E: 16.2 %) és pròxima a la conversa que el té més baix (conversa 5: 16.8 %). Però, com en el cas dels noms, es poden delimitar unes franges clares que, per més que s'aproximin, no arriben a superposar-se. Vegem-ho representat gràficament a la taula següent:

<b>TAULA 25 * - VERBS (Dispersió de les dades)</b>		
6		
7	Escrit B	
8		<b>BANDA DE LA</b>
9	Escrit D Escrit E	<b>PROSA ACADÈMICA</b>
10	Escrit A Escrit C	
11	Classe D	
12	Classe B	
13	Classe A	<b>BANDA DE LA</b>
14		<b>CLASSE MAGISTRAL</b>
15	Classe C	
16	Classe E	
17	Conversa 1 Conversa 5	
18	Conversa 2 Conversa 3	<b>BANDA DE LA</b>
19	Conversa 4	<b>CONVERSA INFORMAL</b>
20		

(\* Percentatges sobre el total de mots)

Si es compara la taula 25 sobre els verbs amb la taula 11 sobre els noms (vg. § 4.2.3), es troba una correspondència que, encara que no sigui total, és bastant significativa. Així, en la banda de la conversa ocupen la posició més alta les converses 5 i 1 en totes dues taules; en la banda de la classe, els textos D i B són el més alts i el text E és el més baix a ambdues; i en la banda de la prosa, el text C ocupa una posició molt baixa i el text B molt alta també en totes dues taules. Com és lògic, la correspondència no podia ser general per a tots els textos,

ja que en aquests càlculs hi intervenen alhora diverses categories (noms, verbs, adjectius, adverbis i mots gramaticals), però ha resultat prou indicativa de la correlació bàsica entre noms i verbs a l'hora de conformar un estil més nominal o un estil més verbal per a cadascun dels textos.

Per completar aquest punt falta només veure com a exemple un parell de fragments de textos reals i comprovar de manera palpable les tendències que les xifres han indicat. Per a això, he seleccionat els dos textos més extrems de la mostra: la conversa 4, que és la que té més verbs (19.3 %) i l'escrit de l'individu B, que és el que en té menys (7.5 %). Vegem a continuació dos fragments de la conversa i un de l'escrit, amb els verbs assenyalats amb color verd:

“HIL: però ara es fan padrí i padrina\_  
o madrina\_  
i després fan\_  
dos testimonis\  
NAA: padrí i padrina també/  
HIL: hi ha gent que ho fa\  
i gent que en porta quatre\  
hi ha gent que només en porta dos\_  
SUS: padrí i padrina\_  
testimoni i testimoni\  
NAA: i la padrina què fa\  
SUS: s'ho mira\  
HIL: jo què sé\  
no ho sé\  
FAN: pos va a buscar el nuvi\  
NAA: no\  
HIL: deu anar a casa el nuvi\  
jo què sé\  
no ho sé\  
no ho sé\  
hi ha gent que en porta quatre\  
hi ha gent que només en porta dos\  
el p- el padrí\_  
i dos testimonis\_  
i el padrí fa de testimoni i de padrí no/  
FAN: mhm\  
HIL: que és lo normal\  
hi ha gent que en porta tres i un de padrí\_  
més dos testimonis\_  
això\_  
NAA: clar\  
depèn”(…)

“SUS: ah al restaurant\  
 NAA: a l'església no\  
 HIL: jo estava pensant l'església\_  
 dic a veure si\_  
 NAA: a l'església no\  
 al res- --  
 FAN: a mi és que si em comencen a dir que se besen\_  
 els hi faig dir en català eh jo\  
 NAA: no et preocupis que\_  
 la gent que vindrà a la nostra boda dubto que digui que se besen\  
 HIL: diran altres coses potser\_  
 però que se besen no\  
 NAA: els foterem fora els que diguin que se besen\  
 FAN: ja podeu marxar\  
 HIL: sense postres\  
 NAA: posarem una pancarta allà davant\_  
 català sisplau\  
 HIL: o amb una senyera\_  
 NAA: ai\  
 quines coses d'anar a parlar\  
 oi/  
 aquí sembla que ens hàgim de casar la setmana que ve\  
 HIL: la setmana que ve no em va bé a mi\  
 NAA: no/  
 HIL: no”

(Conversa 4, pàg. 122 i 126, lín. 5807-5838 i 5981-6005)

“L'embranchida industrial que durant el segle XIX singularitzava cada cop més l'economia catalana respecte al seu entorn regional tenia, naturalment, antecedents que poden escorcollar-se fins a èpoques remotes. Però aquí no caldrà retrocedir més enllà del 1600. Al voltant d'aquesta data, en efecte, es pot situar el començament d'un cicle que porta directament fins a l'inici de la Catalunya fabril.

Cap al 1600 l'economia rural ja s'havia refet del llarguíssim període de crisi i de reorganització obert per les escomeses de la pesta en el segle XIV i prolongat després pel conflicte remença. Entorn del 1600, igualment, s'havia acabat de desarticlar el vell sistema de relacions de l'economia catalana amb l'exterior, molt girat cap a la Mediterrània, i començava una reorientació decisiva. Una i altra cosa, amb les implicacions que se'n desprenen, seran successivament considerades a continuació; després s'examinaran els impulsos determinants de la trajectòria de l'economia catalana en el setcents, ja immediatament significativa per a la història de la industrialització del segle XIX.

L'exposició que segueix no ha estat concebuda com a síntesi descriptiva, sinó com a esquema argumentat; per això destaca només els factors i les tendències considerats fonamentals en l'evolució de l'economia catalana durant el període esmentat. El caràcter de la publicació no permet fer les referències oportunes als nombrosos treballs de recerca i propostes d'interpretació que s'han fet durant els darrers decennis, a partir de l'obra capital de Pierre Vilar. No cal dir que he manllevat a aquests treballs les informacions a partir de les quals he bastit l'argumentació que presento.”

(Escrit individu B, pàg. 255, lín. 11948-11971)

Havent destacat els verbs amb color verd, s'observa a simple vista el contrast entre els dos textos<sup>178</sup>. Tot i que tenen pràcticament el mateix nombre de mots (255 i 252, respectivament), els verbs són més del doble en la conversa (54 ocurrences) que en l'escrit (22 ocurrences). Els verbs de freqüència alta són també proporcionalment més en la conversa (23 ocurrences) que en l'escrit (només 8 ocurrences).

He seleccionat dos fragments diferents de conversa per il·lustrar millor el context d'aquesta abundància de verbs. En el primer fragment s'hi veuen els grups tonals bastant curts, de manera que, fins i tot dins de la intervenció d'un mateix interlocutor, les clàusules se succeeixen sense gaire complementació verbal i sovint, com ja he destacat en l'apartat anterior, sense subjecte explícit. En el segon fragment, en canvi, els grups tonals són més llargs, però hi apareix un altre element que és l'elevat grau de subordinació substantiva (vg. § 5.2.3) i, per tant, els verbs s'hi succeeixen també molt ràpidament, amb poca complementació o amb complements que no són noms sinó pronoms o, precisament, altres verbs en posició subordinada. En definitiva, a tots dos fragments les clàusules són molt curtes i els verbs es constitueixen en les peces clau al voltant de les quals es desenvolupa la informació i s'articula el discurs.

Per contrast, en el fragment escrit el pes primordial és suportat pels noms i els grups nominals, i les clàusules s'estiren, tot incorporant subjectes llargs i diversos complements també nominals. Hi ha subordinació, però no és un recurs per enllaçar verbs entre si, sinó una eina més en una construcció compacta de complements al voltant dels noms, dins d'unes clàusules més llargues i jerarquizades.

---

<sup>178</sup> En l'escrit de l'individu B he assenyalat dos casos de verb passiu (*seran successivament considerades* i *no ha estat concebuda*). La construcció de passiva morfològica és poc freqüent en català i ha aparegut només en rares ocasions en els textos (cap vegada en la conversa informal, 22 en la classe magistral i 30 en la prosa acadèmica, xifres que són minúscules respecte del conjunt de la mostra). El fet que no n'hagi trobat ni una en les converses confirma el caràcter formal i escrit d'aquesta forma verbal. La comparació amb l'anglès no té, en aquest cas, cap sentit, ja que en aquesta llengua l'ús de la passiva, especialment en els textos escrits, és molt habitual.

L'anàlisi d'aquests fragments porta, finalment, a la qüestió de la longitud de clàusula<sup>179</sup>. Així, si es divideix el nombre total de mots dels textos en relació pel nombre total de clàusules, s'arriba als resultats següents:

<b>TAULA 26 *</b>	<b>CONVERSA INFORMAL</b> (4769 clàusules)	<b>CLASSE MAGISTRAL</b> (3666 clàusules)	<b>PROSA ACADÈMICA</b> (2488 clàusules)
<b>Nombre de mots per clàusula</b>	5.74	7.27	10.59

(\* Dades mitjaneres sobre el total de clàusules)

Les dades reflectides a la taula 26 no poden ser més rotundes: confirmen absolutament les observacions que acabo de fer sobre els exemples en el sentit que la conversa informal, amb pocs mots per clàusula, es construeix amb un estil essencialment verbal, mentre que la prosa acadèmica, amb el doble de mots per clàusula, té un estil molt més nominal. Cal tenir en compte, a més, que una part —petita, però no menyspreable— dels mots són, en la classe i encara més en la conversa, repeticions contigües, interjeccions, mots-crossa, mots escapçats o estroncats, etc., de manera que el nombre efectiu de mots per clàusula lèxicament significatius en aquests dos gèneres és encara menor<sup>180</sup>.

<sup>179</sup> El nombre de clàusules, com argumentaré més endavant, l'he calculat a partir del nombre total de formes verbals, conjugades i no conjugades, de cada fragment. Per a una discussió sobre el paper de la unitat *clàusula* en aquest treball, vegeu el punt 5.1.2.

<sup>180</sup> Aquests resultats coincideixen en gran manera amb els de R. O'Donnell (1974), el qual els obté en una recerca amb discursos orals formals televisius i amb escrits periodístics argumentatius (vg. § 3.4.1):

<b>TAULA 27</b>	<i>Spoken</i>	<i>Written</i>
<i>Average number of words per clause</i>	8.33	10.54

(Taula elaborada a partir de les dades de O'Donnell, 1974: 108)

#### 4.3.4- CONCLUSIONS DE L'APARTAT

En aquest apartat s'ha comprovat la tendència clara a un nombre més alt d'ocurrències dels verbs en la conversa informal i a un ús decreixent d'aquesta categoria en la classe magistral i en la prosa acadèmica. En una classificació morfològica, les formes personals dels verbs han estat amb diferència les més habituals, seguides a gran distància per les perífrasis verbals, els infinitius i els gerundis. Pel que fa als verbs de freqüència alta, que fan una aportació menor a la densitat lèxica, s'han distribuït amb preferència en els discursos orals i han tingut un paper molt menor en la prosa. El verb amb una freqüència d'aparició més alta ha estat a tots els gèneres el copulatiu *ser*, en coincidència amb les dades estadístiques del *Diccionari de freqüències*.

Una anàlisi final de dos fragments de textos ha evidenciat el paper primordial dels verbs en el discurs oral, amb una estructuració a partir de clàusules molt curtes, en contrast amb el text escrit, elaborat amb predomini de noms i amb una mitjana de mots per clàusula més alta. D'aquesta manera es va perfilant la caracterització del contrast entre l'estil nominal i l'estil verbal, que és a la base de la distinció entre els modes oral i escrit i que s'anirà completant en els apartats següents.

## 4.4- ELS ADJECTIUS

### 4.4.1- FONAMENTACIÓ CONCEPTUAL I HIPÒTESIS

En la bibliografia consultada s'estableix una diferència important entre dos usos dels adjectius. D'una banda, els anomenats *adjectius atributius*, que complementen un nom dins del sintagma nominal i, d'altra banda, els *adjectius predicatius*, que depenen d'un verb, copulatiu o predicatiu, dins del sintagma verbal<sup>181</sup>. W. Chafe (1982) tracta aquestes dues funcions de l'adjectiu i atorga als atributius un paper primordial en la caracterització del llenguatge escrit:

“We found predicative adjectives, as in ‘the house was *old*’, to occur with approximately equal frequency in spoken and written language: 15.8 occurrences per thousand words in speaking, and 20.2 in writing. Such predicative adjectives are the basic device for expressing states in English. The use of attributive adjectives, as in ‘the *old* house’, allows states to be expressed as modifiers rather than assertions and in thus another integrative device. (...) Although attributive adjectives were fairly common even in speaking, they appeared about four times as often in writing (Spoken: 33.5, Written: 134.9). Thus, while nominalizations show the greatest proportional difference across the two kinds of language, attributive adjectives are the single most prevalent feature of written language.” (Chafe, 1982: 41-42)

També D. Biber (1988) troba un nombre superior d'adjectius atributius en el gènere *academic prose*, en comparació amb els gèneres *prepared speeches* i *face-to-face conversation*. La diferència, però, no és tan gran com en la citació anterior:

---

<sup>181</sup> Aquestes denominacions, habituals en la bibliografia anglòfona, difereixen de les tradicionals en el nostre àmbit, en què se sol parlar de *complement atribut* —o de *predicat nominal*— per als adjectius dependents d'un verb copulatiu, que aquí anomeno *predicatiu*, de manera idèntica als que depenen d'un verb no copulatiu i desenvolupen la funció de *complement predicatiu*. De fet, és cert que tant en un complement atribut com en un complement predicatiu l'adjectiu es troba en una posició *predicativa* en contrast amb la que desenvolupa al costat d'un nom, que aquí anomeno *atributiva*.



TAULA 28 *	<i>Face-to-face conversation</i>	<i>Prepared speeches</i>	<i>Academic prose</i>
<i>Attributive adjectives</i>	4.1 %	4.9 %	7.7 %
<i>Predicative adjectives</i>	0.4 %	0.4 %	0.5 %
<i>Total adjectives</i>	4.3 %	5.3 %	8.2 %

(\* Percentatges sobre el total de mots - Taula elaborada a partir de les dades de Biber, 1988: 255, 265 i 269)<sup>182</sup>

En coherència amb aquests resultats cal esperar, doncs, una presència més gran d'adjectius atributius en la llengua escrita i un nombre més o menys equivalent d'adjectius predicatius en els tres gèneres. Malgrat això, i encara que en la majoria de les recerques ressenyades el nombre d'adjectius predicatius sigui lleugerament superior en la llengua escrita que en la llengua oral, la concepció de M.A.K. Halliday que relaciona els usos nominals amb l'escriptura i els usos verbals amb l'oralitat podria portar a la hipòtesi que els adjectius predicatius, associats sempre a un verb, haurien de ser més abundants en el discurs oral.

#### 4.4.2- UNITATS, PROCEDIMENTS I CRITERIS DE CODIFICACIÓ

En aquest treball, he classificat com a adjectius els anomenats *adjectius qualificatius* per la gramàtica tradicional i els participis amb funció adjectival. No hi he inclòs els

<sup>182</sup> En consonància amb les tendències d'aquesta taula, però amb diferències importants en els percentatges, hi ha els estudis de M.E. Poole i T.W. Field (1976) i de L.W. Smith (1994). Els primers troben un 8.6 % d'adjectius per a la llengua oral i un 11.2 % per a l'escrita (1976: 308), per sobre dels resultats de D. Biber. El segon distingeix entre tres gèneres i hi troba unes quantitats d'adjectius atributius —no en cerca de predicatius— inferiors a les de D. Biber: 1.8 % per al gènere *informal conversation*, 2.2 % per a *formal interview* i 6.9 % per a *written essay* (1994: 58). També hi ha, com en el cas dels noms i dels verbs, els resultats de W.N. Francis i H. Kuèera (1982) sobre la llengua escrita, que indiquen un 7.1 % d'adjectius, xifra bastant pròxima, en aquest cas, a la de D. Biber. I, encara, L.S. Golub (1969), que, sense donar xifres concretes, troba més adjectius predicatius —que anomena *linking verb plus adjective combinations* (1969: 82)— en la llengua oral que en l'escrita, en un contrast sorprenent amb les dades de W. Chafe i D. Biber.

adjectius determinatius —els *determinants* d'algunes gramàtiques estructuralistes—, que es caracteritzen per la conformació de repertoris tancats i, per tant, són mots gramaticals que no contribueixen a la densitat lèxica<sup>183</sup>. També s'han descartat, òbviament, els participis que tenen un valor verbal, acompanyats per l'auxiliar *haver* en els temps compostos, que ja han estat comptats amb les formes verbals conjugades de l'apartat anterior.

Els adjectius i els participis s'han dividit en els dos grups establerts pels estudis precedents:

**A/ Adjectius i participis atributius**, en funció de complement del nom dins d'un sintagma nominal.

**B/ Adjectius i participis predicatius**, en funció de *complement atribut* d'un verb copulatiu o bé en funció de *complement predicatiu*, que concorda en gènere i en nombre amb el subjecte o amb el complement directe d'un verb predicatiu del tipus *tornar-se*, *anar*, *caminar*, *fer-se*, *dir-se*, etc., i que pronominalitza amb el pronom *hi* i, en casos comptats, amb *en*. (Aquests verbs ja els he esmentat en el punt 4.3.2, tot indicant que no els comptaria com a copulatius.)

El comentari de M.A.K. Halliday sobre els mots de freqüència alta, que ja s'ha aplicat als noms i als verbs en els apartats anteriors, afecta també els adjectius qualificatius. Els mots d'aquesta categoria més habituals segons el *Diccionari de freqüències* són els següents<sup>184</sup>:

---

<sup>183</sup> He comptat com a adjectius qualificatius els mots *cert/a* (un **cert** rebuig), *determinat/da* (una **determinada** tendència) i els superlatius *major* (= més gran), *menor* (= més petit), *millor* (= més bo) i *pitjor* (= més dolent). No he inclòs en el càlcul els adjectius *general*, *particular* i *definitiva* dins de les expressions *en general*, *en particular* i *en definitiva*, en les quals no funcionen com a qualificadors.

<sup>184</sup> Si per als verbs partia d'una llista de 10 perquè ja abraçaven la meitat de les ocurrències verbals i per als noms d'una llista de 80 perquè no eren tan freqüents, per als adjectius caldria treballar amb llistes d'uns 250 per arribar als percentatges dels noms i d'uns 700 per igualar els dels verbs. Per tant, per raons òbvies de volum de treball, he preferit de fer la recerca, amb valor només indicatiu, sobre una llista més reduïda.

1- <i>gran</i>	8- <i>petit</i>	15- <i>dit</i>
2- <i>bo</i>	9- <i>social</i>	16- <i>possible</i>
3- <i>nou</i>	10- <i>diferent</i>	17- <i>antic</i>
4- <i>català</i>	11- <i>divers</i>	18- <i>següent</i>
5- <i>general</i>	12- <i>major</i>	19- <i>polític</i>
6- <i>cert</i>	13- <i>humà</i>	20- <i>sol</i>
7- <i>propri</i>	14- <i>important</i>	

Per raons de concepte, he exclòs d'aquesta llista els adjectius següents:

a/ L'adjectiu *general*, perquè sol aparèixer més en l'expressió adverbial *en general* que no pas amb un valor ple d'adjectiu qualificatiu.

b/ Els adjectius *diferent* i *divers*, perquè són més freqüents en l'ús com a adjectius determinatius, anteposats al nom (*diverses ocasions, diferents episodis*), que no pas com a adjectius qualificatius, amb el significat ple (*una llista diversa, un episodi diferent*).

c/ El participi *dit* (l'únic participi de la llista), perquè és usat més sovint dins dels temps compostos del verb *dir* (*he dit, havíeu dit, etc.*) que no pas com a complement del nom.

També he exclòs, per raons d'equiparació proporcional entre els gèneres, tal com abans he fet en el cas dels noms, els adjectius següents:

d/ Els adjectius *humà, important, major, polític* i *possible*, perquè no apareixen a cap dels fragments de conversa codificats.

e/ L'adjectiu *català* perquè no apareix a cap dels fragments de classe.

f/ L'adjectiu *següent* perquè no apareix a cap dels fragments de conversa ni de prosa acadèmica.

Per tant, la llista resultant que he cercat en els textos és la composta pels nou adjectius següents:

- |                 |                  |
|-----------------|------------------|
| 1- <i>gran</i>  | 6- <i>petit</i>  |
| 2- <i>bo</i>    | 7- <i>social</i> |
| 3- <i>nou</i>   | 8- <i>antic</i>  |
| 4- <i>cert</i>  | 9- <i>sol</i>    |
| 5- <i>propi</i> |                  |

Cal destacar que els set primers adjectius d'aquesta llista estan entre els nou primers de la llista anterior, és a dir, entre els nou més freqüents en català segons el *Diccionari de freqüències*.

Pel que fa a la metodologia, he fet tots els càlculs amb fragments de 2500 mots, és a dir, sobre la meitat de la mostra, d'acord amb les conclusions exposades en el punt 3.5.1. Per al còmput, he efectuat una codificació manual, distingint entre els adjectius atributius i els adjectius predicatius, i una recerca automàtica amb *TACT* per als adjectius de freqüència alta<sup>185</sup>.

#### 4.4.3- RESULTATS I DISCUSSIÓ

Els resultats coincideixen plenament amb les hipòtesis formulades, fins i tot amb una progressió més accentuada que les dades obtingudes per W. Chafe i D. Biber. Com es pot observar a la taula 29 (vg. pàgina següent), el nombre total d'adjectius augmenta progressivament des de la conversa informal (3.0 %) fins a la prosa acadèmica (10.1 %), amb la classe en la típica posició intermèdia (5.8 %), aquesta vegada més decantada cap a la banda de l'oralitat (2.8 punts de diferència amb la conversa) que cap a la de l'escriptura (4.3 punts de diferència amb la prosa).

---

<sup>185</sup> Les instruccions de *TACT* per a la llista d'adjectius de freqüència alta han estat *Usebase - F3 Complete Word-List - F9 Writing to disk*, amb la posterior recerca manual dels mots i del nombre d'ocurrències en l'ordenació alfabètica. En canvi, per a la recerca dels adjectius més habituals en els textos he utilitzat *TactFreq - Frequency Order*.

<b>TAULA 29 *</b>	<b>CONVERSA INFORMAL</b> (27254 mots)	<b>CLASSE MAGISTRAL</b> (26374 mots)	<b>PROSA ACADÈMICA</b> (26468 mots)
<b>Adjectius atributius</b>	0.8 % (218)	4.7 % (1239)	9.2 % (2435)
<b>Adjectius predicatius</b>	2.2 % (600)	1.1 % (290)	0.9 % (238)
<b>Total adjectius</b>	3.0 % (818)	5.8 % (1529)	10.1 % (2673)

(\* Percentatges sobre el total de mots i, entre parèntesis, dades absolutes d'ocurrències – Test  $\chi^2$ : probabilitat d'atzar inferior al 0.0001 % a totes tres files)

Els responsables d'aquest augment són clarament, en consonància amb les dades de W. Chafe, els adjectius atributius, que mostren un creixement espectacular des del 0.8 % de la conversa al 9.2 % de la prosa. Els adjectius predicatius, dins d'una certa igualtat, presenten la tendència contrària i augmenten en direcció a la llengua oral, fins al punt que en la conversa s'arriben a capgirar les posicions i el nombre d'adjectius predicatius és més del doble que el d'atributius (2.2 % i 0.8 %, respectivament). Així, la vacil·lació de les dades de W. Chafe i D. Biber pel que fa als adjectius predicatius esdevé aquí una clara inversió de la situació en l'oral respecte de l'escrit.

Pel que fa als adjectius de freqüència alta, he trobat els resultats que s'indiquen a la taula 30 (vg. pàgina següent): els adjectius de freqüència alta han tingut un paper més destacat en la conversa informal (15.7 %), mentre que en la classe magistral i en la prosa acadèmica han compartit gairebé el mateix percentatge (10.5 % i 10.6%, respectivament)<sup>186</sup>.

<sup>186</sup> Dono directament els percentatges sobre el total d'adjectius perquè són més significatius i perquè, en aquest cas, les xifres sobre el total de mots són molt petites —0.5 % (128) en la conversa informal, 0.6 % (160) en la classe magistral i 1.1 % (283) en la prosa acadèmica, sobre el total de 27254, 26374 i 26468 mots, respectivament— i donen poc la mesura del paper d'aquesta unitat respecte del conjunt. Cal tenir en compte, a més, que aquests percentatges d'adjectius no són directament comparables ni amb els dels noms ni amb els dels verbs.

<b>TAULA 30 *</b>	<b>CONVERSA INFORMAL</b> (818 adjectius)	<b>CLASSE MAGISTRAL</b> (1530 adjectius)	<b>PROSA ACADÈMICA</b> (2673 adjectius)
<b>Adjectius de freqüència alta</b>	15.7 % (128)	10.5 % (160)	10.6 % (283)

(\* Percentatges sobre el total d'adjectius i, entre parèntesis, dades absolutes d'ocurrències – Test  $\chi^2$ : probabilitat d'atzar del 0.0140 %) <sup>187</sup>

D'altra banda, els adjectius més freqüents de la mostra han estat els següents:

**Conversa 1 - Tema principal:** els afers econòmics d'una família, les seves cases i les seves propietats. **Adjectius més freqüents:** 1- *clar*, 2- *bo*, 3- *gran*, 4- *jubilat*, 5- *exacte*.

**Conversa 2 - Tema principal:** qüestions diverses. **Adjectius més freqüents:** 1- *clar*, 2- *igual*, 3- *nou*, 4- *diabètic*, 5- *estranya*.

**Conversa 3 - Tema principal:** visita turística al sud de França. **Adjectius més freqüents:** 1- *clar*, 2- *francès*, 3- *car*, 4- *espanyol*, 5- *feudal*.

**Conversa 4 - Tema principal:** la boda d'una de les filles. **Adjectius més freqüents:** 1- *clar*, 2- *evangèlica*, 3- *hortera*, 4- *maco*, 5- *pomposo*.

**Conversa 5 - Tema principal:** viatges i malalties. **Adjectius més freqüents:** 1- *clar*, 2- *maco*, 3- *baix*, 4- *diferent*, 5- *igual*.

**Classe individu A - Tema:** història econòmica d'Espanya en la dècada dels anys cinquanta del segle XX. **Adjectius més freqüents:** 1- *espanyola*, 2- *exterior*, 3- *comercial*, 4- *exportador*, 5- *important*.

---

<sup>187</sup> Només he trobat un estudi que faci la recerca dels adjectius segons la freqüència i ha estat el de M.E. Poole i T.W Field (1976: 308), que donen xifres sobre els *uncommon adjectives*. Els percentatges d'aquest treball, però, no es poden contrastar amb els meus ja que els autors no especifiquen quins són aquests adjectius “no comuns”, que constitueixen el 59.7 % en la llengua oral i el 67.7 % en la llengua escrita, en percentatges sobre el total dels adjectius de la seva mostra.

**Classe individu B - Tema:** història dels inicis de l'islamisme. **Adjectius més freqüents:** 1- *legítim*, 2- *musulmà*, 3- *polític*, 4- *amagat*, 5- *únic*.

**Classe individu C - Tema:** història de l'esclavitud. **Adjectius més freqüents:** 1- *blanc*, 2- *clar*, 3- *mongòlic*, 4- *alemany*, 5- *complex*.

**Classe individu D - Tema:** història del Brasil en el segle XIX. **Adjectius més freqüents:** 1- *portuguès*, 2- *brasiler*, 3- *espanyol*, 4- *napoleònic*, 5- *nou*.

**Classe individu E - Tema:** història de Xina en els segles I i II. **Adjectius més freqüents:** 1- *gran*, 2- *magne*, 3- *anglès*, 4- *central*, 5- *important*.

**Escrit individu A - Tema:** història de la construcció a Barcelona en el segle XIX. **Adjectius més freqüents:** 1- *gran*, 2- *llarg*, 3- *social*, 4- *contrari*, 5- *productiu*.

**Escrit individu B - Tema:** història econòmica de Catalunya en el segle XVIII. **Adjectius més freqüents:** 1- *agrari*, 2- *útil*, 3- *català*, 4- *gran*, 5- *atlàntic*.

**Escrit individu C - Tema:** assaig contra l'aplicació de l'anàlisi del discurs en la ciència històrica - Reflexió prospectiva sobre la marxa de l'economia mundial. **Adjectius més freqüents:** 1- *gran*, 2- *popular*, 3- *literari*, 4- *polític*, 5- *acadèmic*.

**Escrit individu D - Tema:** història del moviment literari de la Renaixença. **Adjectius més freqüents:** 1- *popular*, 2- *renaixentista*, 3- *català*, 4- *autòcton*, 5- *escrit*.

**Escrit individu E - Tema:** història de la formació de l'imperi xinès. **Adjectius més freqüents:** 1- *Xinès*, 2- *gran*, 3- *antic*, 4- *diferent*, 5- *social*.

Com es pot veure en aquestes llistes, la correspondència entre els adjectius i el tema principal tractat a cada text no és tan determinant com la que hi havia en el cas dels noms, i s'observa més en les classes i en els escrits. En les converses, en canvi, destaca en primera lloc l'adjectiu *clar*, que deu aquesta posició a l'expressió col·loquial *és clar*, que, amb valor de marcadors discursiu o d'omplidor (mot-crossa), segons el cas, és

habitual en la llengua oral informal (vg. § 5.5.3)<sup>188</sup>. D'altra banda, a la darrera de les classes i a quatre dels escrits, l'adjectiu *gran* se situa entre els cinc primers, cosa que justifica la posició capdavantera que té a la llista del *Diccionari de freqüències*. Altres adjectius com *polític*, *nou*, *social*, *català* i *antic* coincideixen a aparèixer, també, en aquestes llistes i en les del diccionari.

Les taules 31 i 32 (vg. aquesta pàgina i la següent) forneixen dades sobre la distribució dels resultats. Segons aquestes taules, la dispersió és més gran en la classe magistral (3.8 punts) que en la conversa informal i en la prosa acadèmica (2.2 i 1.9, respectivament). Com era d'esperar, una dispersió del 3.8 % en unes xifres més aviat baixes dóna lloc a interseccions (vg. taula 33). Així, les bandes de la conversa i de la classe se superposen, ja que les dues converses amb més adjectius (la conversa 4 i la conversa 1) sobrepassen la classe que en té menys (la de l'individu E). Entre la classe i la prosa, en canvi, no es dóna aquest fenomen i les bandes apareixen ben delimitades.

<b>TAULA 31 *</b>	<b>ADJECTIUS</b>	<b>CONVERSA</b>
<b>CONVERSA 1</b>	Total adjectius .....	3.8 %
<b>CONVERSA 2</b>	Total adjectius .....	1.8 %
<b>CONVERSA 3</b>	Total adjectius .....	2.9 %
<b>CONVERSA 4</b>	Total adjectius .....	4.0 %
<b>CONVERSA 5</b>	Total adjectius .....	2.7 %

(\* Percentatges sobre el total de mots)

<sup>188</sup> La relació entre el tema i els mots més freqüents és, per tant, molt estreta quan es tracta dels noms, segueix essent important, però ho és menys, en els adjectius i es difumina per als verbs que, com s'ha vist en l'apartat anterior, es presenten en un nombre d'ítems més reduït i, correlativament, amb un nombre d'aparicions molt més alt. Els verbs més freqüents són, doncs, mots molt generals que no tenen relació amb el tema principal del text on apareixen.



<b>TAULA 32 *</b>	<b>ADJECTIUS</b>	<b>CLASSE MAGISTRAL</b>	<b>PROSA ACADÈMICA</b>
<b>INDIVIDU A</b>	Total adjectius .....	7.5 %	9.2 %
<b>INDIVIDU B</b>	Total adjectius .....	5.3 %	10.2 %
<b>INDIVIDU C</b>	Total adjectius .....	5.1 %	9.7 %
<b>INDIVIDU D</b>	Total adjectius .....	7.3 %	11.1 %
<b>INDIVIDU E</b>	Total adjectius .....	3.7 %	10.6 %

(\* Percentatges sobre el total de mots)

<b>TAULA 33 * - ADJECTIUS (Dispersió de les dades)</b>		
12		
11	<b>Escrit D</b>	<b>BANDA</b>
10	<b>Escrit B Escrit E</b>	<b>DE LA</b>
9	<b>Escrit A Escrit C</b>	<b>PROSA ACADÈMICA</b>
8		
7	<b>Classe A Classe D</b>	<b>BANDA</b>
6		<b>DE LA</b>
5	<b>Classe B Classe C</b>	<b>CLASSE MAGISTRAL</b>
4	<b>Classe E Conversa 1 Conversa 4</b>	<b>BANDA D'INTERSECCIÓ</b>
3		<b>BANDA</b>
2	<b>Conversa 3 Conversa 5</b>	<b>DE LA</b>
1	<b>Conversa 2</b>	<b>CONVERSA INFORMAL</b>
0		

(\* Percentatges sobre el total de mots)

Per donar uns exemples del grau d'ús dels adjectius en fragments reals, he triat una altra vegada els textos més distants entre si, que en aquest cas són la conversa 2 (1.8 %) i l'escrit de l'individu D (11.1 %), i també he buscat un fragment que ocupi una posició més o menys intermèdia, com la classe de l'individu C (5.3 %). Tots tres fragments són de 326 mots i hi he assenyalat els adjectius atributius amb color vermell i els predicatius amb verd:

“MMM: i llavors li ha mirat e- --  
 li ha posat les gotes\_  
 i al cap d'una estona\_  
 al cap d'una estona li ha mirat els ulls\  
 i ha dit que\_  
 que estava bé\  
 que no tingués d'això\_  
 que fins l'any que ve es pot esperar un al- --  
 esperar a l'any que ve\_  
 que no té re\  
 però que\_  
 que li faran aquesta altra prova\_  
 perquè\_  
 i llavors aquella noia\_  
 diu\_  
 doncs espera't\  
 i ens ha fet entrar en un altre puesto\_  
 CME: i m'ha regalat les gotes que--  
 m'haig de posar unes gotes abans d'anar-hi\  
 MMM: abans d'anar-hi\_  
 aquí casa\_  
 s'ha de posar unes--  
 dugues menes de gotes\  
 que aquella noia ens e les ha donades\  
 PPP: bueno\  
 però això fins al gener ja no\  
 CME: sí\  
 MMM: bueno\  
 però aquella noia\_  
 ha buscat la llista\_  
 i li ha donat lo més aviat que ha p- --  
 que han pogut\  
 PPP: va bé conèixer-la veus\  
 MMM: perquè sinó es tindria que esperar\_  
 i e- --  
 PPP: i t'ha conegut\  
 MMM: sí\_  
 PPP: t'ha conegut\  
 vaia\  
 a tu també\  
 MMM: i ens ha dit\_

i jo li he dit\_  
 dic ai\_  
 la Maria Carme ha estat **ingressada** aquí a Sant Pau\_  
 i d'urgències\_  
 i a Sant Pau diu\_  
 m'ho haguéssiu dit\  
 diu\_  
 jo ja l'haguera anat a veure unes quantes vegades\  
 CME: sí però\_  
 què has de pensar que\_  
 has de dir-li a ella perquè et vingui a veure\  
 PPP: sí\  
 que teniu l'adreça d'ella\  
 CME: home jo tinc el telèfon\  
 me'l va donar una vegada\  
 PPP: és **igual**  
 ja està bé\  
 MMM: bueno\  
 però\_  
 PPP: és d'agrair això\  
 MMM: sí i tant\  
 PPP: conxo\  
 però està molt bé\  
 MMM: i li ha buscat el número més aviat\_  
 perquè--  
 CME: mare\_  
 em deixa dir una cosa\_  
 però no ho digui a ningú eh\  
 MMM: no\  
 CME: m'he trobat el Jordi\_  
 quan m'he trobat el Jordi\  
 de--  
 MMM: sí\  
 CME: m'he trobat el Jordi del cau a Sant Pau\_  
 PPP: qui és el Jordi\  
 MMM: ai\  
 jo haig d'anar a--  
 CME: el Jordi\  
 el Jordi que també és **diabètic**  
 el del cau\  
 el Jordi aquell que viu a Santa Coloma\  
 el Jordi Zabala\”

(Conversa 2, pàg. 46-48, lín. 2119-

2201)

“la veritat és que si poseu un japonès enmig d’un autobús o enmig d’un metro\_  
 i no teniu altre element que el color de la pell\_  
 per jutjar si és japonès o no\_  
 us serà bastant **difícil**  
 perquè no és no acostuma a ser\_  
 ni més ni menys **groc** que bastantes d’altres persones\_

i no solament les que tenen patiments\_  
 patiments de fetge\  
 l'altre element que sí que considerem aleshores que els caracteritza\_  
 és allò que anomenem trets **mongòlics** no/  
 és a dir pòmuls **sortits**\_  
 i el plec aquest que fa que els ulls semblin més **petits**\  
 una altra paradoxa\_  
 els trets **mongòlics** no són **mongòlics**\  
 els mongols inicialment eren un poble **uropeoïde**\_  
 de cabells **rossos** i ulls **blaus**\  
 només han esdevingut **mongòlics** a través de creuaments amb d'altres eh/  
 o sigui que si voleu fins i tot aquesta paradoxa\  
 de fet d'altra manera\_  
 és molt **difícil** de de establir gaire\_  
 gaires elements **objectius** en aquesta qüestió dels colors\_  
 per exemple negre i blanc\_  
 jo no conec ningú **blanc**\_  
 només la gent que a la Rambla\_  
 doncs es pinta per representacions d'aquestes com estàtua\  
 no conec ningú que tingui aquesta blancor eh/  
 i per altra banda on està la frontera entre el blanc i el negre/  
 depèn eh/  
 depèn\  
 per a molts doncs\_  
 eee nosaltres formen part clarament allò ja de gent\_  
 que poden començar a ser considerats com com poc **blancs** i bastant **negres**\  
 per exemple prengueu el cas aquell de Kipling\_  
 el poema **famós** en què parla de la càrrega de l'home **blanc**\_  
 justament celebra els nordamericans perquè\_  
 amb la conquesta de Filipines podran dur a Filipines\_  
 la civilització que correspon a l'home **blanc** de dur al poble **inferior** eh/  
 està **clar** que per Kipling els espanyols no eren **blancs** eh/  
 i és una consideració que la podreu sentir\_  
 que la podríeu sentir **repetida**\_  
 eee amb bastanta amb bastanta facilitat\_  
 per gent que doncs eee tenen menys melanina a la pell encara que nosaltres\_”

(Classe individu C, pàg. 203-204, lín. 9637-9680)

“El mecanisme d'exclusió arrencava, per tant, de consideracions **polítiques** i **ideològiques**,  
 en uns moments en què la revolució, la guerra **civil** i el marc **polític liberal**, per **restrictiu**  
 que fos, afavorien la mobilització i la definició **política** de grups molt **amplis** de la  
 població. Allò que es volia evitar eren les contaminacions d'una cultura **popular** que  
 coneixem molt malament i no pas per casualitat, que s'expressava a través de **múltiples**  
 canals i que, de cap manera, no s'ha d'entendre com un codi **definit**, **coherent** i **sistemàtic**  
 de valors. Hi havia maneres molt **diferents** d'encarar-se amb el fantasma d'un món **popular**  
**convuls** i de vegades molt **violent**, que s'expressava parcialment en la literatura **popular**, i  
 tractar de dominar-lo. Una primera possibilitat era intentar d'integrar el món **popular** o  
 algunes de les seves manifestacions des de dalt, com a primera matèria d'una literatura  
**possible**, tal com es produí en totes les **grans** literatures del segle XIX. També n'hi ha  
 casos a Catalunya. En donaré dos exemples: dos retrats de la **suggestiva** violència dels  
 jamàncios, que dos literats varen intentar de plasmar amb la **necessària** distància **classista**.

Un d'ells va ser el valencià Vicent Boix, l'*íntim* amic del poeta *suïcida barceloní* Joan Arolas, i molt *amic* també tota la seva vida de Víctor Balaguer. Boix dedicà una poesia en castellà a la Jamància que es publicà a la revista *conservadora* La Corona, que dirigia Piferrer. Responia a la filosofia de l'exorcisme de l'adversari tan *característica* del gust *burgès* de mitjan segle XIX. L'altre exemple és una composició del *jove* Sol i Padrís, que llavors encara militava en el progressisme *barceloní*, que Illas i Vidal féu conèixer en la *llarga* necrologia que li dedicà, i que no constava en les antologies que Bofarull i Balaguer donarien a conèixer en aquells anys. Són dues composicions d'un *gran* interès, que ens mostren novament que l'espectre d'opcions *possibles* per a la literatura *catalana* en aquells anys era més *ampli* del que després s'imposà.”<sup>189</sup>

(Escrit individu D, pàg. 278-279, lín. 12972-12998)

Aquests fragments són una bona mostra de les tendències arquetípiques en els textos analitzats. A simple vista ja s'aprecia la quantitat més alta d'adjectius en la prosa acadèmica (38 ocurrencies), a continuació en la classe magistral (23 ocurrencies) i el nombre francament escàs en la conversa informal (només 3 ocurrencies), en la qual s'han succeït passatges ben extensos sense que els parlants hagin utilitzat cap adjectiu.

Pel que fa a la diferenciació entre adjectius atributius i predicatius, el contrast també és considerable. En la prosa, dels 38 adjectius, 36 són atributius i només 2 són predicatius. En la classe, de 23 adjectius, 12 són atributius i 11 són predicatius. I en la conversa, els adjectius que hi apareixen són tots tres predicatius. Podem veure, per tant, com l'individu D, en el seu fragment escrit, ha construït el text amb molts adjectius en funció de complement del nom (*consideracions polítiques i ideològiques; guerra civil; marc polític liberal; codi definit, coherent i sistemàtic*; etc.) i també amb força sintagmes preposicionals complements del nom, com veurem a l'apartat 4.7. La classe de l'individu C, en canvi, no utilitza tant aquest recurs i col·loca la meitat dels adjectius en

---

<sup>189</sup> A la vista d'aquests fragments cal aclarir els criteris amb què he comptat com a adjectius o no alguns dels mots que hi apareixen. Es tracta dels casos següents:

1- El mot *japonès* a l'enunciat “*per jutjar si és japonès o no*” (classe individu C), l'he considerat substantiu, ja que ha estat usat com a tal dues línies més amunt.

2- Els mots *blanc* i *negre* (classe individu C) els he considerat noms quan són la denominació d'un color i adjectius quan són la complementació del mot *color* que s'ha elidit (“*jo no conec ningú blanc*” = *jo no conec ningú de color blanc*). En aquest segon cas, fins i tot apareixen modificats per adverbis (“*que poden començar a ser considerats com poc blancs i bastant negres*”).

3- El mot *ingressada* (conversa 2) a l'enunciat “*la Maria Carme ha estat ingressada aquí a Sant Pau*”, l'he assenyalat com a adjectiu considerant-lo el complement predicatiu depenent del verb *estar* i no pas una construcció de passiva depenent d'un verb *ser*, bastant inversemblant en aquest context i que, a més, implicaria l'aparició d'un participi culte *estat* en lloc del col·loquial *sigut*.

posició predicativa (*els trets mongòlics no són mongòlics* | *els mongols inicialment eren un poble europeïde* | *de cabells rossos i ulls blaus* | *només han esdevingut mongòlics*...). I la conversa 2 és sens dubte el fragment que més sorprèn perquè prescindeix absolutament dels adjectius complements del nom i els poquíssims adjectius que s'hi usen depenen d'un verb.

Aquestes formes d'organització dels sintagmes i de tot el discurs estan en consonància amb el major nombre de noms en la llengua escrita i de verbs en la llengua oral. Així, els adjectius fan una contribució decisiva, tant amb la presència com amb l'absència, al predomini de l'estil nominal o de l'estil verbal en la construcció dels discursos.

#### **4.4.4- CONCLUSIONS DE L'APARTAT**

Els resultats d'aquest apartat han coincidit àmpliament amb els esperats: hi ha hagut més adjectius en la llengua escrita que en l'oral i, dins d'aquesta, més en la classe magistral que en la conversa informal. Aquest contrast ha estat motivat pel comportament dels adjectius atributius, els quals han augmentat en una proporció espectacular des de la conversa fins a la prosa acadèmica. Els adjectius predicatius, en canvi, mostren la tendència contrària, si bé d'una manera més moderada, i són més habituals en la conversa informal que en l'escrit. Els adjectius de freqüència alta, d'altra banda, apareixen més en la conversa, com correspon a uns mots de pes lèxic menor. Els adjectius més freqüents als textos estan en correspondència amb els temes tractats a cada fragment en una mesura menor que els noms i major que els verbs. La dispersió dels resultats ha provocat problemes de superposició entre la conversa i la classe, mentre que a la zona alta, entre la classe i la prosa, les bandes s'han delimitat amb nitidesa.

Finalment, l'anàlisi directa d'uns fragments de textos ha fet evident el paper destacat que té l'adjectiu en la transmissió de la informació en la llengua escrita, per mitjà d'unes estructures sintagmàtiques que inclouen un o diversos adjectius en funció de complement del nom. La conversa informal, en canvi, prescindeix àmpliament de la complementació nominal a través d'adjectius i opta —de fet, també escadusserament—

per l'expressió de les qualitats per mitjà de la predicació verbal i, sobretot, per altres construccions que no impliquen l'adjectivació.

## 4.5- ELS ADVERBIS

### 4.5.1- FONAMENTACIÓ CONCEPTUAL I HIPÒTESIS

La darrera categoria que es pot incloure en el grup dels mots lexicals és l'adverbi. Ara bé, a diferència dels noms, dels verbs i dels adjectius, que són classes de paraules de contingut lèxic indiscutible, els adverbis són de difícil adscripció. És més, en la bibliografia consultada, fins i tot la noció mateixa d'adverbi resulta conflictiva. A continuació exposaré, sobre aquesta qüestió, els punts de vista de P. Fabra (1956), A.M. Badia i Margarit (1994), G. Ferrater (1981), S. Bonet i J. Solà (1986), R. Morant (1987), P. Sancho (1995) i M.J. Cuenca (1996):

**1 –P. Fabra (1956)** planteja la categoria adverbial amb una certa distància i considera que els adverbis “*pròpiament dits*” són només els de manera:

“Els adverbis pròpiament dits (adverbis de manera) juguen respecte al verb el mateix paper que els adjectius respecte al nom: així com per mitjà dels adjectius fem distincions com *paper blanc, paper negre*, etc., així mateix per mitjà d'aquests adverbis fem distincions com *caminar lentament, caminar ràpidament*, etc. I, de fet, la immensa majoria dels adverbis de manera són mots derivats d'adjectius. (...) Tradicionalment s'inclou en la classe dels adverbis tot mot invariable que no serveix normalment de complement determinatiu, sia d'un verb (*caminar lentament*), sia d'un adjectiu o d'un altre adverbi (*més ample, molt més*), i això fa que, al costat dels adverbis pròpiament dits (esmentats en aquest paràgraf: *bé, malament*), hi hagi els adverbis de quantitat o de grau (*molt, ben bé*), de lloc (*aquí, enfora*), de temps, (*avui, després*), de negació (*no*), de dubte (*potser*), etc.” (Fabra, 1956: 81-82)

Les llistes de mots que P. Fabra, després de fer aquestes consideracions, inclou entre els adverbis són les següents:

“ Adverbis de manera (ultra els adverbis en *-ment*): *així, com, bé, millor, pitjor, ensems, corrents, endebades, exprés, etc.*, i locucions adverbials diverses.

. Adverbis de quantitat o de grau: *quant, tant, molt, poc, gaire, bastant, prou, massa, força, més, menys, que, ben, gairebé, quasi, gens*, i locucions adverbials diverses.

. Adverbis de lloc: *davant, damunt, on, ací, aquí, allà, allí, ençà, enllà, deçà, dellà, çà i lla, onsevulla, pertot, enlloc, dalt, baix, endins, enfora, endavant, endarrere, avant, enrera, amunt, avall, endalt, ensota, etc.*

. Adverbis de temps: *quan, ara, adés, avui, demà, demà passat, ahir, abans-d'ahir, anit, llavors, aleshores, sempre, tothora, mai, abans, després, aviat, tard, ja, encara*, i locucions adverbials diverses.

. Adverbis de negació: *no, pas*.

. Altres adverbis: *tanmateix, nogensmenys, però, així, doncs, encara, endemés, més, també, tampoc*, i locucions adverbials diverses.” (Adaptació de Fabra, 1956: 81-84)

**2– A.M. Badia i Margarit (1994)** fa una classificació en la qual aporta la noció d'*adverbi lexemàtic*, que cal suposar referida als adverbis dotats d'un pes lexical. Vegem-la a continuació:

“I - *Adverbis lexemàtics*

1- Adverbis de manera

2- Adverbis de correlació i de referència<sup>190</sup>

II - *Adverbis de nombre*

1- Adverbis de quantitat

2- Adverbis d'ordre

III - *Mots d'afirmació, negació i dubte*

1- Les formes

2- Naturalesa gramatical de *sí* i *no*

3- Adverbis i mots afins de funció polivalent” (Badia i Margarit, 1994: 691)

---

<sup>190</sup> A.M. Badia i Margarit (1994) considera “*de correlació i de referència*” els adverbis de lloc i els de temps.



Tot i això, amb una actitud cauta, el mateix A.M. Badia i Margarit relativitza el valor d'aquesta classificació:

“El Q74 (quadre 74) mostra la classificació dels adverbis que acabo de suggerir, no tant en forma d'una proposta teòrica com amb l'objectiu de presentar-los al lector amb prou garanties d'eficàcia.” (Badia i Margarit, 1994: 690)

**3– G. Ferrater (1981)** fa una de les crítiques més severes a aquesta categoria de mots.

Les seves paraules són les següents:

“A la darrera reunió, amb l'exposició de Mireia Montané, vam inventariar cinc menes bàsiques de funcions adverbials. Parlo de funcions i no d'adverbis, perquè insisteixo que no crec que l'adverbi, en quant classe de mots o part del discurs, tingui cap figura d'existència (si no és que vulguem limitar-lo als adjectivals en *-ment*, i en aquest cas la cosa serà tan trivial que no valdrà la pena d'ocupar-nos-en el temps).” (Ferrater, 1981: 107)

**4– S. Bonet i J. Solà (1986)**, des del punt de vista de la gramàtica generativa, fan una proposta de clarificació de les categories *preposició* i *adverbi*, amb una reducció dràstica del grup de mots considerats adverbis:

“En les consideracions generals sobre la categoria Sprep s'ha apuntat que, per qualificar de Prep un element lèxic (s'entén, diferent de N, Adj i V per hipòtesi), només cal exigir-li que, en algun dels seus usos, sigui capaç de realitzar una relació de subcategorització amb una expansió del tipus SN, SPrep o O. *A contrari*, un bon criteri d'identificació de la categoria Adv podrà consistir, en principi, a caracteritzar els adverbis per una incapacitat absoluta de regir Compl X'. Aquest criteri funciona de manera adequada per confirmar com a Adv la majoria dels tradicionals adverbis de mode, lèxicament obtinguts gràcies a la sufixació per *-ment* de la forma femenina dels adjectius. (...) El criteri funciona, així mateix, per establir una línia divisòria clara en el camp dels adverbis tradicionals, entre els considerables com a Prep usades intransitivament i els Adv pròpiament dits (incapaços de cap mena d'expansió). Entre els Adv poden citar-se, p.e.: *àdhuc*, *aviat*, *ben*, *gairebé*, *inclús*, *potser*, *quasi*, *sovint*, *també*, *tampoc*, *tanmateix*, *tard*, etc.” (Bonet i Solà, 1986: 70-71)

**5– R. Morant (1987)** també redueix el grup dels adverbis, n'exclou la negació com un fenomen a part i hi inclou només els següents:

“Fetes aquestes observacions, entendrem per adverbis:

a) els lexemes «adjectius-adverbis» o adverbis curts: *alt, mal, bé, molt*, etc.

Parla molt.

b) les locucions adverbials lexicalitzades: *a peu coix, a ulls clucs, de bon cor, de bracet, de sobte*, etc.

La Carne botava de genollons.

c) les formes en -ment, que formen un grup molt nombrós: *lentament, humanament, forçosament, impensadament*, etc.

Pep ho negà enèrgicament.” (Morant, 1987: 31)

**6– P. Sancho (1995)**, en una aportació més recent, tracta sobre els límits entre adverbi i preposició. La seva proposta final per als mots que tradicionalment s’han considerat dins de totes dues categories és considerar-los o bé preposicions o bé noms locatius, segons el cas, i mantenir la denominació *adverbi* només a efectes pràctics:

“Així doncs, creiem que a (69a) hi ha una veritable preposició, però a (69b) i (69c) estem davant noms locatius, en què una relació espacial s’ha conceptualitzat com a cosa, i que com a tal nom pot tenir complement (69b) o no (69c). Amb tot, continuarem anomenant adverbis aquests noms locatius d’acord amb la terminologia tradicional.

(69) a. Joan és davant la casa

b. Joan és davant de la casa

c. Joan és davant” (Sancho, 1995: 95)

**7– M.J. Cuenca (1996)**, encara més recent, en una panoràmica completa i aclaridora, fa el següent plantejament:

“L’adverbi és una categoria invariable que, a diferència de la preposició, no admet expansió nominal. La categoria d’adverbi, calaix de sastre de totes les categories invariables que la tradició gramatical no ha pogut definir com a preposicions o com a conjuncions, s’hauria de reduir probablement als adverbis de manera, que tenen significat ple, formen una classe oberta i sintàcticament funcionen com a complements verbals o com a especificadors de la major part de categories (oració, sintagma adjectival, adverbial, preposicional i en algun cas verbal). Tradicionalment s’han inclòs dins de la categoria d’adverbi tots aquells elements invariables que no funcionen com a elements d’enllaç, de

manera que aquesta ha resultat una categoria heterogènia i mal definida. (...) Els adverbis de manera són els prototípics i, per tant, els únics que clarament han de constituir la categoria adverbial. Entre aquests, es poden diferenciar els llargs, o adverbis en *-ment* (*ràpidament*), i amb més reserves els curts, formalment idèntics als adjectius, llevat que no presenten variació de gènere i nombre, i, per tant, són susceptibles de ser tractats com a elements perifèrics de la categoria adjectiu (*ràpid*).” (Cuenca, 1996: 178-179)

Per a M.J. Cuenca, la resta d’adverbis s’han de resituar en altres categories. Els de lloc i de temps en preposicions intransitives (*davant, després, fora, etc.*) i en díctics (*ara, avui, aquí*), que són proformes que equivalen a sintagmes nominals. Els de quantitat s’han d’incloure dins dels quantificadors. Els d’afirmació, de negació i de dubte, i de modalitat s’han de tractar dins de la subcategoria oracional de fragment quan actuen com a mots-frase (p. ex., *no!*) i es poden mantenir com a adverbis, si bé no prototípics, quan actuen com a especificadors verbals (p. ex., *no va venir*).

En definitiva, després d’aquest repàs bibliogràfic, es pot arribar a les tres conclusions següents:

**a/ La categoria adverbial és qüestionada** per pràcticament tots els autors.

**b/ Els únics mots que tothom admet com a adverbis són els de manera.**

**c/ Amb l’ excepció d’A.M. Badia i Margarit, es considera que els únics adverbis amb significat ple són els adverbis de manera en *-ment*.**

Pel que fa al paper dels adverbis respecte dels paràmetres del registre, hi ha una referència a l’ús dels adverbis en *-ment* en el treball de C. Gelpí (1998), que arriba a la conclusió següent:

“(4) El to condiciona el nombre d’adverbis en *-ment*: si el text té to formal, la freqüència d’adverbis serà més alta que si el text té to informal.” (Gelpí, 1998: 140)

Fora d'aquesta autora de l'àmbit català, no he trobat cap altre estudi sobre aquesta unitat que pugui comparar amb el meu<sup>191</sup>. Pel que fa a les hipòtesis de recerca, d'acord amb C. Gelpí cal considerar que els adverbis en *-ment* seran més habituals en registres formals i elaborats. Per tant, s'hauran de trobar més adverbis en *-ment* en la prosa acadèmica i en la classe magistral que en la conversa informal, en la progressió habitual que va de la llengua oral a la llengua escrita —o viceversa.

#### 4.5.2- UNITATS, PROCEDIMENTS I CRITERIS DE CODIFICACIÓ

D'acord amb el que he comentat en el punt anterior sobre la naturalesa dels adverbis, he considerat que l'únic grup que es pot incloure en el còmput de les unitats lèxiques és el dels adverbis de manera en *-ment*<sup>192</sup>. Aquests adverbis són els candidats ideals per formar part de la recerca perquè, a més de ser defensats inequívocament per tots els autors, compleixen amb el requisit estructural de formar un repertori obert —per definició, tan obert com el dels adjectius— i, pel fet mateix de provenir d'un adjectiu, tenen un pes lexical o significat ple indiscutible<sup>193</sup>.

---

<sup>191</sup> D. Biber (1988:224, 238, 240), dins del seu llarg repertori de trets lingüístics, inclou els adverbis, però estableix unes llistes i unes classificacions que fan impossible l'acarament. Així, no només no separa enlloc els adverbis en *-ly* dels altres, sinó que en col·loca amb aquesta terminació dins de les categories *place adverbials* i *time adverbials*, en separa uns altres com a modificadors d'oració i modalitzadors (*hedges*, *downtoners* i *amplifiers*) i fa una categoria final d'*adverbs* que abraça els no inclosos en els grups anteriors.

<sup>192</sup> Pel que fa al caràcter lexical o gramatical dels mots, crec que hi ha una gradació en què els noms, els verbs, els adjectius qualificatius i els adverbis en *-ment* serien categories lèxiques prototípiques en un extrem de la gradació; els pronoms personals, els pronoms relatius, els articles, les preposicions febles i algunes conjuncions serien categories gramaticals prototípiques en l'altre extrem; i hi hauria un grup considerable de mots (adjectius i pronoms determinatius, preposicions fortes, algunes conjuncions, etc.), del qual formarien part molts dels adverbis tradicionals, que estarien situats en una posició intermèdia. El mateix M.A.K. Halliday defensa aquesta gradació i el paper intermedi que hi ocupen alguns adverbis amb l'argumentació següent:

“As you would expect, there is a continuum from lexis into grammar: while many items in a language are clearly of one kind or the other, there are always likely to be intermediate cases. In English, prepositions and certain classes of adverb (for example, MODAL adverbs like *always*, *perhaps*) are on this borderline.” (Halliday, 1985b: 63)

<sup>193</sup> És un repertori tan obert com el dels adjectius amb l'excepció dels noms dels colors, que poden actuar com a adjectius, però no accepten la transformació en adverbi (*blanca*: *\*blancament*), i la dels gentilicis, que tampoc no l'admeten (*francesa*: *\*francesament*).

L'opció per aquest tipus d'adverbi en una recerca de tipus lèxic coincideix, a més, amb la feta per M.T. Cabré (1977) i per C. Gelpí (1998), en unes circumstàncies equiparables<sup>194</sup>.

Per al càlcul en aquest apartat he treballat amb fragments de 5000 mots (la totalitat de la mostra)<sup>195</sup>. He fet una recerca automàtica amb *TACT* de tots els mots acabats en *-ment*, amb una posterior selecció manual dels adverbis respecte dels derivats nominals sufixats amb *-ment*<sup>196</sup>.

#### 4.5.3- RESULTATS I DISCUSSIÓ

Els percentatges generals d'aparició dels adverbis en *-ment* han estat els següents:

<b>TAULA 34 *</b>	<b>CONVERSA INFORMAL</b> (27254 mots)	<b>CLASSE MAGISTRAL</b> (26374 mots)	<b>PROSA ACADÈMICA</b> (26468 mots)
<b>Adverbis en <i>-ment</i></b>	0.2 % (54)	1.2 % (316)	1.0 % (265)

(\* Percentatges sobre el total de mots i, entre parèntesis, dades absolutes d'ocurrències – Test  $\chi^2$ : probabilitat d'atzar inferior al 0.0001 %)

<sup>194</sup> M.T. Cabré, comunicació personal.

<sup>195</sup> En aquest apartat he prescindit de la recerca dels mots de freqüència alta que he fet en apartats anteriors per als noms, els verbs i els adjectius. El motiu ha estat que els adverbis en *-ment* són ben escassos, tant en els meus textos com en les llistes del *Diccionari de freqüències*. Així, el primer adverbi en *-ment* que hi ha al diccionari és *solament*, que ocupa la posició 29a d'entre els mots etiquetats com a adverbis, i la 258a del total de mots. Els adverbis en *-ment* que hi ha a continuació ocupen posicions encara més endarrerides: *especialment*, el lloc 582è; *precisament*, el 829è; *finalment*, el 835è; *generalment*, el 865è i *principalment*, ja més enllà del primer miler, el 1018è. És clar que, en aquestes condicions, qualsevol càlcul de freqüències per part meua seria poc significatiu.

<sup>196</sup> Les instruccions de *TACT* han estat *Usebase - F2 Query . \*ment - F9 Writing to Disk L09*.

Com es pot veure a la taula 34, la conversa informal ha quedat distanciada dels altres dos gèneres i la classe magistral i la prosa acadèmica han obtingut uns índexs bastant acostats, de manera que en l'ús dels adverbis en *-ment* sembla que pesa més el caràcter formal i especialitzat d'un text —hipòtesi que havia adoptat més amunt— que no pas la producció immediata, en temps real, pròpia del discurs oral.

La classe magistral ha registrat, curiosament, el nombre més alt d'ocurrències i ha trencat així la progressió arquetípica entre els tres gèneres. Aquest fet pot explicar-se pel caràcter repetitiu d'alguns adverbis en *-ment* en el discurs oral d'alguns dels individus estudiats. Per exemple, en la classe de l'individu D, de 93 adverbis en *-ment*, n'hi ha hagut dos que han sumat 36 ocurrències, que constitueixen més d'un terç (el 38.7 %) del total del fragment. Es tracta de l'adverbi *evidentment*, amb 19 ocurrències, i de l'adverbi *clarament*, amb 17 ocurrències. Un fenomen semblant, si bé no tan accentuat, s'ha produït amb alguns dels adverbis de les classes dels altres individus: *juntament*, amb 7 ocurrències en l'individu A; *naturalment*, amb 7 ocurrències en l'individu C; *absolutament*, amb 9 ocurrències i *finalment*, amb 7, en l'individu E. L'abundància d'aquests adverbis sorprèn, ja que els altres han aparegut, en les converses, en els escrits i, fins i tot, en les classes mateixes, només una o dues vegades. El fet s'explica, probablement, per una preferència personal d'aquests individus per uns adverbis determinats, que apareixen de manera recurrent en el seu discurs i acaben acostant-se a la funció d'omplidors o mots-crossa. Vegem, a continuació, alguns exemples d'aquesta mena en la classe de l'individu D:

“s'articula **constitucionalment** en uns termes molt ambigus eh/  
**lògicament** amb aquest tipus de formació\_  
però la for la formació d'aquest primer embrió d'estat brasiler\_  
**evidentment** eeem ee planteja problemes molt complexos d'articulació\_  
que són fàcils d'entendre\_  
dos els podeu entendre **claríssima** fa **facilíssimament** per tot el que hem dit\_  
un és el problema **evidentment** de la relació Portugal Brasil\_”  
(...)  
“bueno aqu la hist aquesta història del país ind\_  
de de Brasil com a país independent en el segle XIX\_  
té **evidentment** dos moments mh **clarament fàcilment** identificables eh”  
(...)  
“durant tot el segle XIX es va delimitant **clarament**\_  
**clarament** un un un tipus de visió per part de les classes altes brasileres\_”

sobretot de les més potents\_  
de quin era ee quin havia de ser el futur territorial de Brasil eh/  
això apunta **clarament** als ee cap a uns protagonistes essencials\_  
que són els anomenats paulistas\_”

(Classe individu D, pàg. 216-226, lín. 10248-10254, 10539-10541 i 10760-10765)

Cal dir, en primer lloc, que he seleccionat uns passatges amb una especial proximitat en l'aparició dels adverbis en *-ment* —per poder-los exemplificar sense haver de citar fragments massa llargs— i això no ha de crear una falsa impressió sobre la freqüència que tenen en els textos, que de fet és molt menor. En el primer fragment es pot veure una important acumulació d'adverbis en *-ment* en poques línies, cosa que dóna al discurs un caràcter repetitiu. El primer d'aquest adverbis, *constitucionalment*, fa una aportació de contingut més destacable, però els altres semblen mots amb menys pes significatiu, en els quals el parlant es complau, es recolza per fer discórrer amb més facilitat el pensament o, potser, s'esforça per expressar algun tipus de modalització. Fins i tot, hi ha un cas en què s'estronca l'adverbi *claríssimament* i es reformula, just a continuació, en l'adverbi *facilíssimament*.

En el segon fragment, hi ha tres adverbis en *-ment* dins d'un mateix enunciat i, curiosament, s'hi succeeixen *clarament* i *fàcilment*, en una combinació que ja hem trobat en l'exemple anterior i que deu resultar especialment còmoda per a aquest individu. En el tercer, s'hi observen tres aparicions bastant pròximes del mateix adverbi, *clarament*, que serien evitades en la llengua escrita per una qüestió d'estil i que ajuden a explicar per què aquest adverbi ha aparegut 17 vegades en la classe d'aquest individu.

La distribució dels resultats ha tingut un comportament diferenciat en la conversa informal i en els altres dos gèneres (vg. taules 35 i 36, en les pàgines següents). Així, la dispersió és de només 0.2 punts en la conversa informal, però d'1.0 en la classe magistral i de 0.6 en la prosa acadèmica. Aquesta dispersió, en unes xifres tan baixes i tan pròximes entre la classe i la prosa, provoca una confusió ben accentuada entre les dades d'aquests dos gèneres, tant que esdevenen indistingibles. Com es pot observar a la taula 37, la situació resultant és d'una banda perfectament delimitada per a la conversa, i unes bandes per a la classe i per a la prosa que es barregen fins al punt de fondre's en

una de sola. La dada més significativa d'aquesta fusió és el fet que tant l'índex més alt (1.8 % de la classe de l'individu D) com el més baix (0.8 % de la classe de l'individu B) de la banda comuna corresponen a la classe, i tots els escrits i totes les altres classes queden compresos dins d'aquesta delimitació. La intersecció completa entre aquests dos gèneres, lluny de ser una dada incòmoda, reforça encara més la conclusió a què havia arribat abans sobre el comportament comú que tenen, en un bloc diferenciat de la conversa informal<sup>197</sup>. Així, el pes de la formalitat i de l'especificitat sembla ser tan decisiu per a l'aparició dels adverbis en *-ment* que la classe magistral i la prosa acadèmica es comporten en aquest tret com un sol gènere<sup>198</sup>.

<b>TAULA 35 *</b>	<b>ADVERBIS EN <i>-MENT</i></b>	<b>CONVERSA</b>
<b>CONVERSA 1</b>	Total adverbis en <i>-ment</i> .....	0.3 %
<b>CONVERSA 2</b>	Total adverbis en <i>-ment</i> .....	0.1 %
<b>CONVERSA 3</b>	Total adverbis en <i>-ment</i> .....	0.2 %
<b>CONVERSA 4</b>	Total adverbis en <i>-ment</i> .....	0.2 %
<b>CONVERSA 5</b>	Total adverbis en <i>-ment</i> .....	0.3 %

(\* Percentatges sobre el total de mots)

<sup>197</sup> Hauria estat una dada incòmoda si les bandes de dispersió dels tres gèneres haguessin estat confuses en tots o en la majoria de trets. Com que la superposició de les bandes en el present treball ha estat només ocasional (vegeu-ne uns resums als punts 4.10.6 i 5.7.6), ha tingut el valor positiu de mostrar tendències específiques en moments concrets.

<sup>198</sup> En aquest apartat dels adverbis en *-ment* no ofereixo exemples comparatius de textos de gèneres diferents perquè, com he comentat en donar un exemple de l'ús recurrent en les classes de l'individu D, qualsevol tria d'un fragment on aparegui algun adverbi en *-ment* implica per si mateixa un grau de falsejament de la realitat de la mostra. La freqüència d'aparició dels adverbis en *-ment* és tan baixa que només la casualitat o una recerca massa intencionada farien que en trobés alguns de propers en un fragment del text, i llavors la tria d'aquest fragment resultaria, de fet, poc representativa.



<b>TAULA 36 *</b>	<b>ADVERBIS EN -MENT</b>	<b>CLASSE MAGISTRAL</b>	<b>PROSA ACADÈMICA</b>
<b>INDIVIDU A</b>	Total adverbis en <i>-ment</i> ...	1.2 %	1.4 %
<b>INDIVIDU B</b>	Total adverbis en <i>-ment</i> ...	0.8 %	1.0 %
<b>INDIVIDU C</b>	Total adverbis en <i>-ment</i> ...	1.1 %	0.9 %
<b>INDIVIDU D</b>	Total adverbis en <i>-ment</i> ...	1.8 %	0.9 %
<b>INDIVIDU E</b>	Total adverbis en <i>-ment</i> ...	0.9 %	0.8 %

(\* Percentatges sobre el total de mots)

<b>TAULA 37 * - ADVERBIS EN -MENT (Dispersió de les dades)</b>		
1.9		
1.8	<b>Classe D</b>	
1.7		<b>BANDA</b>
1.6		<b>COMUNA</b>
1.5		<b>DE LA</b>
1.4	<b>Escrit A</b>	<b>PROSA ACADÈMICA</b>
1.3		<b>I</b>
1.2	<b>Classe A</b>	<b>DE LA</b>
1.1	<b>Classe C</b>	<b>CLASSE MAGISTRAL</b>
1.0	<b>Escrit B</b>	
0.9	<b>Classe E</b> <b>Escrit C</b> <b>Escrit D</b>	
0.8	<b>Classe B</b> <b>Escrit E</b>	
0.7		
0.4		
0.3	<b>Conversa 1</b> <b>Conversa 5</b>	<b>BANDA</b>
0.2	<b>Conversa 3</b> <b>Conversa 4</b>	<b>DE LA</b>
0.1	<b>Conversa 2</b>	<b>CONVERSA INFORMAL</b>
0.0		

(\* Percentatges sobre el total de mots)

#### 4.5.4- CONCLUSIONS DE L'APARTAT

La categoria adverbial ha resultat, a diferència de les anteriors, difícil de delimitar. D'acord amb el criteri més generalitzat, només he considerat aptes per al còmput dels mots lexicals els adverbis de manera en *-ment*, derivats d'adjectius, que conformen un repertori obert, amb pes significatiu i amb una clara funció adverbial.

Les hipòtesis d'aquest apartat s'han complert només parcialment. D'una banda, la conversa s'ha destacat, tal com era previsible, pel baix índex d'adverbis en *-ment*, cinc i sis vegades menor que el dels altres gèneres. La classe magistral i la prosa acadèmica s'han agrupat en l'altre pol, amb un ús més freqüent d'aquest tipus d'adverbis, un índex només lleugerament més alt en la classe que en la prosa i una confusió tan gran en la distribució de les dades que els dos gèneres s'han integrat en una banda de dispersió única. Així doncs, la formalitat i l'especificitat del text han pesat més, en aquest cas, que la diferència de canal, acústic o gràfic.

La diferència en l'ús dels adverbis en *-ment* entre la conversa informal d'una banda i la classe magistral i la prosa acadèmica de l'altra, ha estat, en definitiva, ben notable, i constitueix una aportació a la densitat lèxica d'aquests dos gèneres darrers, molt petita si es té en compte el poc pes quantitatiu dels adverbis en *-ment*, però amb la importància que li correspon en la descripció del contrast global entre els gèneres del present estudi.

## 4.6- ELS MOTS LEXICALS I ELS MOTS GRAMATICALS

### 4.6.1- FONAMENTACIÓ CONCEPTUAL, UNITATS, PROCEDIMENTS I CRITERIS DE CODIFICACIÓ

La proporció de mots lexicals i de mots gramaticals o *densitat lèxica* en sentit estricte admet, segons M.A.K. Halliday (1985b), dues maneres de ser calculada: la primera, més general, segons el percentatge de mots plens sobre el total de mots d'un text; la segona, més ajustada, segons el percentatge de mots plens sobre el total de clàusules d'un text:

“In general, the more ‘written’ the language being used, the higher will be the proportion of lexical words to the total number of running words in the text. (...) Lexical density (1): the number of lexical items as a proportion of the number of running words. (...) Words are not packed inside other words; they are packaged in larger grammatical units —sentences, and their component parts. It is this packaging into larger grammatical structures that really determines the informational density of a passage of text. (...) Lexical density (2): the number of lexical items as a ratio of the number of clauses.” (Halliday, 1985b: 64-67)<sup>199</sup>

M.A.K Halliday, per tant, proposa la unitat clàusula com a denominador més adequat per al càlcul de l'índex de mots lexicals en un text<sup>200</sup>. D'acord amb això, estableix unes bandes numèriques per a aquest índex (mots lexicals per clàusula):

“On the basis of various samples I have found that a typical average lexical density for spoken English is between 1.5 and 2, whereas the figure for written English settles down

---

<sup>199</sup> M.A.K. Halliday basa la seva noció de densitat lèxica únicament en la proporció de mots lexicals i de mots gramaticals d'un text. Com veurem més endavant, en l'apartat 4.9, crec que els índexs de variació i de repetició lèxica (*type / token* i *token / type*, respectivament) i la longitud de mot han de tenir també un paper important en la determinació d'aquesta densitat dels textos.

<sup>200</sup> Com ja he assenyalat en parlar de l'índex de verbs per clàusula, al punt 5.1.2 discutiré el paper de la unitat clàusula i hi comentaré amb més detall la visió de M.A.K. Halliday.

somewhere between 3 and 6, depending on the level of formality in the writing.”  
(Halliday, 1985b: 80)

Pel que fa a les unitats que cal tenir en compte, en l'apartat anterior la particular i complexa naturalesa dels adverbis ens ha conduït a la discussió sobre el caràcter lexical o gramatical de les diverses categories de mots. Com es pot deduir del que he comentat sobre aquesta qüestió, a efectes d'aquest càlcul consideraré mots lexicals els de les categories següents: noms, adjectius qualificatius, participis amb funció adjectival, verbs en forma personal i no personal, i adverbis en *-ment*. Complementàriament, consideraré mots gramaticals les preposicions, les conjuncions, els pronoms, els articles, els adjectius determinatius i la resta d'adverbis. Així, les dades d'aquest apartat són, de fet, una suma de les obtingudes en els apartats anteriors<sup>201</sup>.

#### **4.6.2- RESULTATS I DISCUSSIÓ**

La taula 38 (vg. pàgina següent) reflecteix els resultats generals de la proporció de mots lexicals i mots gramaticals. Les relacions que mostra aquesta taula són clares i ben proporcionades, amb l'acostumada progressió de la conversa informal a la prosa acadèmica, progressió ascendent en el cas dels mots lexicals i descendent en el cas dels mots gramaticals, amb la classe a l'entremig dels altres dos gèneres. La diferència entre la conversa i la prosa és de 10.9 punts, i la classe se situa a 4.2 punts de la prosa i a 6.7 punts de la conversa, més decantada, per tant, cap a l'estil propi de la llengua escrita.

Si es comparen aquestes dades amb les de J. Ure (1971) que he comentat a l'apartat 4.1, cal admetre que s'hi acorden plenament: la mitjana de la conversa informal s'ha situat en el 34.1 %, per sota del 40 % fixat per J.Ure com a límit entre la llengua oral i l'escrita.

---

<sup>201</sup> A més de les afirmacions de M.A.K. Halliday (1985b) i de J. Ure (1971) sobre la proporció entre els mots lexicals i gramaticals d'un text, que he comentat a l'inici d'aquest capítol, compto també amb el treball de C. Gelpí (1998), que ja he citat anteriorment i que relaciona l'aparició d'unitats lèxiques amb el to formal del text:

“(6) El to condiciona el nombre total d'unitats lèxiques: si el text té to formal, contindrà més unitats lèxiques, independentment de la seva distribució, que si el text té to informal.” (Gelpí, 1998: 140)

<b>TAULA 38 *</b>	<b>CONVERSA INFORMAL</b> (27254 mots)	<b>CLASSE MAGISTRAL</b> (26374 mots)	<b>PROSA ACADÈMICA</b> (26468 mots)
<b>Mots lexicals</b>	34.1 % (9294)	40.8 % (10760)	45.0 % (11911)
<b>Mots gramaticals</b>	65.9 % (17960)	59.2 % (15614)	55.0 % (14557)

(\* Percentatges sobre el total de mots i, entre parèntesis, dades absolutes d'ocurrències – Test  $\chi^2$ : probabilitat d'atzar inferior al 0.0001 %)

Les classes, per la seva part, ho han fet en el 40.8 %, just al límit i així han demostrat el caràcter intermedi d'aquest gènere entre els altres dos —ahora que el caràcter extrem que té dins de l'oralitat—, i els escrits han arribat fins al 45.0 %, clarament per sobre del límit establert<sup>202</sup>.

A aquestes dades s'hi ha d'incorporar les del segon càlcul, el que M.A.K. Halliday anomena “*Lexical density (2)*”, amb la *ratio* de mots lexicals per clàusula, a la qual he incorporat, complementàriament, la *ratio* de mots gramaticals per clàusula. Es pot veure tot explicitat a la taula 39. D'acord amb aquesta taula, la *ratio* de mots lexicals per clàusula augmenta notablement de la conversa informal (1.96) a la classe magistral (2.97), i de la classe a la prosa acadèmica (4.77), amb un increment més accentuat que en les dades de la taula 38, ja que aquí el denominador ha estat el nombre de clàusules, que decreix progressivament en direcció a l'escrit. La *ratio* de mots gramaticals també augmenta, si bé en una mesura menor, com correspon al menor pes específic d'aquests mots en la classe magistral i en la prosa acadèmica. Les xifres resultants s'integren amb comoditat a les bandes de M.A.K. Halliday (1985b) —oral entre 1.5 i 2; escrit entre 3 i 6—, amb la classe magistral situada al mig, en terra de ningú, si bé molt pròxima a la llengua escrita.

<sup>202</sup> La distribució de les dades dins de cada gènere, que presentaré més avall, també confirma la validesa de la idea postulada per J. Ure.

<b>TAULA 39 *</b>	<b>CONVERSA INFORMAL</b>	<b>CLASSE MAGISTRAL</b>	<b>PROSA ACADÈMICA</b>
<b>Nombre de mots lexicals per clàusula</b>	1.96	2.97	4.77
<b>Nombre de mots gramaticals per clàusula</b>	3.78	4.30	5.82

(\* Índexs sobre el total de clàusules)

Per afinar encara més en l'anàlisi, a la taules 40 i 41 podem veure els resultats de la taula 38 desglossats per categories. La taula 40 mostra el diferent pes quantitatiu de cada categoria en la xifra total de mots lexicals. En la conversa informal, els verbs són la categoria més freqüent, seguits dels noms i, ja molt lluny, els adjectius i els adverbis. En la classe magistral s'inverteix la relació entre els verbs i els noms, i aquests ocupen la primera posició, amb els verbs i les altres categories a continuació. I en la prosa acadèmica, els noms se situen en un primer lloc molt destacat i el percentatge de verbs s'ha reduït tant que els adjectius han passat a ocupar la segona posició.

<b>TAULA 40 *</b>	<b>CONVERSA INFORMAL (27254 mots)</b>	<b>CLASSE MAGISTRAL (26374 mots)</b>	<b>PROSA ACADÈMICA (26468 mots)</b>
<b>Noms</b>	13.4 %	19.9 %	24.5 %
<b>Verbs</b>	17.5 %	13.9 %	9.4 %
<b>Adjectius</b>	3.0 %	5.8 %	10.1 %
<b>Adverbis en <i>-ment</i></b>	0.2 %	1.2 %	1.0 %

(\* Percentatges sobre el total de mots)

<b>TAULA 41 *</b>	<b>CONVERSA INFORMAL</b> (9294 mots lex.)	<b>CLASSE MAGISTRAL</b> (10760 mots lex.)	<b>PROSA ACADÈMICA</b> (11911 mots lex.)
<b>Noms</b>	39.3 %	48.8 %	54.4 %
<b>Verbs</b>	51.3 %	34.1 %	20.9 %
<b>Adjectius</b>	8.8 %	14.2 %	22.5 %
<b>Adverbis en <i>-ment</i></b>	0.6 %	2.9 %	2.2 %

(\* Percentatges sobre el total de mots lexicals)

D'altra banda, si es té en compte que el nombre de noms i el de verbs gairebé es contraresten, ja que segueixen progressions en sentits contraris —hi ha una diferència de només 3 punts a favor dels noms en el còmput global—, els responsables principals del contrast entre la conversa informal i la prosa acadèmica en el nombre de mots lexicals són els adjectius, que aporten 7.1 dels 10.9 punts d'increment que hi ha entre ambdós gèneres i s'erigeixen així en la clau de la diferència quantitativa en la densitat lèxica<sup>203</sup>.

A la taula 41, des d'un altre punt de vista percentual, es pot veure com en la conversa informal els verbs constitueixen el 51.3 % dels mots lexicals, percentatge que es redueix a un 34.1 % en la classe magistral i a un 20.9 % en la prosa acadèmica. En aquest gènere, en canvi, els noms suposen un 54.4 %, i els noms més els adjectius —la majoria dels quals són atributius i, per tant, complements nominal— sumen un aclaparador 76.9 %. Així doncs, enlloc com a les taules 40 i 41 no havíem pogut copsar amb tanta claredat l'evolució quantitativa de l'estil verbal a l'estil nominal.

<sup>203</sup> Especifico que són la clau només quantitativa perquè, com hem anat veient en els apartats anteriors, la diferència entre aquests gèneres és, sobretot, en l'estil, més verbal o més nominal, i en la compactació de la informació. De fet, sense un fort increment en el nombre de noms, no podria haver-hi adjectius que en depenguessin.

Pel que fa al pes significatiu dels mots a cada gènere, cal fer encara una precisió més. D'acord amb els resultats del punt 4.3.3, el grup dels verbs en la conversa informal, que hi són la categoria dominant, està constituït en un 50.9 % per verbs de freqüència alta, cosa que redueix encara més el pes informatiu i, per tant, la densitat lèxica del conjunt. Si a això s'hi afegeix que els noms i els adjectius de freqüència alta, si bé en una proporció menor, també són més habituals en la conversa, s'ha d'acabar conclouent que el desnivell en la densitat lèxica entre la conversa informal i la prosa acadèmica és fins i tot més gran del que indiquen les xifres que fins ara he anat comentant.

La distribució de les dades, que ha estat bastant coherent, es pot veure a les taules 42 i 43 que hi ha a continuació: Els resultats s'han agrupat correctament dins de cada gènere, amb una dispersió notable en el cas de les converses (7.8 punts) i baixa en les classes i en els escrits (2.7 punts a cada gènere). La distància entre la conversa amb l'índex més alt i la classe que el té més baix ha estat de 2 punts, i entre la classe i la prosa acadèmica ha estat d'1.8 punts. Aquesta distribució ha donat lloc, per tant, a unes bandes de dispersió ben delimitades, com es pot veure a la taula 44.

<i>TAULA 42 *</i>	<b>MOTS LEXICALS</b>	<b>CONVERSA</b>
<b>CONVERSA 1</b>	Total mots lexicals .....	36.0 %
<b>CONVERSA 2</b>	Total mots lexicals .....	31.8 %
<b>CONVERSA 3</b>	Total mots lexicals .....	29.6 %
<b>CONVERSA 4</b>	Total mots lexicals .....	36.2 %
<b>CONVERSA 5</b>	Total mots lexicals .....	37.4 %

(\* Percentatges sobre el total de mots)



<b>TAULA 43 *</b>	<b>MOTS LEXICALS</b>	<b>CLASSE MAGISTRAL</b>	<b>PROSA ACADÈMICA</b>
<b>INDIVIDU A</b>	Total mots lexicals .....	41.3 %	46.6 %
<b>INDIVIDU B</b>	Total mots lexicals .....	39.4 %	44.4 %
<b>INDIVIDU C</b>	Total mots lexicals .....	41.2 %	44.1 %
<b>INDIVIDU D</b>	Total mots lexicals .....	42.1 %	43.9 %
<b>INDIVIDU E</b>	Total mots lexicals .....	39.7 %	46.3 %

(\* Percentatges sobre el total de mots)

<b>TAULA 44 * - MOTS LEXICALS (Dispersió de les dades)</b>		
48		
46	<b>Escrit A Escrit E</b>	<b>BANDA DE LA</b>
44	<b>Escrit B Escrit C Escrit D</b>	<b>PROSA ACADÈMICA</b>
42	<b>Classe D</b>	<b>BANDA</b>
40	<b>Classe A Classe C</b>	<b>DE LA</b>
38	<b>Classe B Classe E</b>	<b>CLASSE MAGISTRAL</b>
36	<b>Conv. 1 Conv. 4 Conv. 5</b>	<b>BANDA</b>
34		<b>DE LA</b>
32		<b>CONVERSA INFORMAL</b>
30	<b>Conversa 2</b>	
28	<b>Conversa 3</b>	
26		

(\* Percentatges sobre el total de mots)<sup>204</sup>

<sup>204</sup> En aquest apartat tampoc no donaré fragments de textos, perquè l'essencial que s'hi podria observar ja s'ha mostrat en els exemples dels apartats anteriors, dels quals aquest no és més que una compilació interpretativa.

#### **4.6.3- CONCLUSIONS DE L'APARTAT**

La relació mots lexicals / mots gramaticals ha mostrat, com ja és habitual en aquest treball, una progressió des de la conversa fins a l'escrit, amb la classe magistral situada en una posició intermèdia. Així, la conversa ha tingut el nombre més alt de mots gramaticals i més baix de mots lexicals, entre els quals destaquen en primer lloc els verbs, que coincideixen a ser, en un cinquanta per cent, mots de freqüència alta i, per tant, de menys pes significatiu. La prosa acadèmica, en canvi, ha tingut un percentatge de mots lexicals superior al de la conversa, amb els noms en primer lloc, els adjectius en segona posició i un paper menor dels mots de freqüència alta.

L'índex de mots lexicals per clàusula no ha fet més que reforçar aquestes dades i mostrar un percentatge progressivament superior en la classe magistral i en la prosa acadèmica. Així, amb les dades d'aquest i dels apartats anteriors s'ha mostrat el grau de densitat lèxica dels gèneres estudiats i el paper particular que hi desenvolupen cadascuna de les categories lexicals. En els apartats següents analitzaré aspectes complementaris com els sintagmes preposicionals, les nominalitzacions, la variació lèxica i la longitud de mot, que contribuiran a completar el panorama de la densitat lèxica i dels estils nominal i verbal.

## 4.7- ELS SINTAGMES PREPOSICIONALS

### 4.7.1- FONAMENTACIÓ CONCEPTUAL I HIPÒTESIS

Diversos autors han remarcat el fet que la llengua escrita presenta més sintagmes preposicionals que la llengua oral (Chafe, 1982; Halliday, 1985b i 1985c; Biber, 1986 i 1988; i Chafe i Danielewicz, 1987). Per a M.A.K. Halliday, la justificació d'aquest fet és que l'estil preferentment nominal dels escrits provoca la concentració de sintagmes preposicionals al voltant del nom com a nucli. W. Chafe (1982) en fa una explicació més restringida, que relaciona amb les nominalitzacions. Com es pot veure a continuació, W. Chafe no busca tota mena de sintagmes preposicionals, sinó només aquells que resulten d'una nominalització i són introduïts, per tant, per la preposició *de*:

“Associated with nominalizations is the occurrence of prepositional phrases, usually introduced by *of*, which express the subject or object of the nominalized verb; for example, *treatment of children*, where children are the object of treatment, or *development of proper names*, where the proper names are the subject of development. We found these ‘genitive subjects’ and ‘genitive objects’ occurring significantly in the written material, but hardly at all in the spoken:

	Spoken	Written
genitive subject	0.1	4.1
genitive object	0.1	12.3

” (Chafe, 1982: 40)

D. Biber (1988) també troba més preposicions —en aquest cas, preposicions de tota mena— en l'escrit, lluny, però, de les grans diferències apuntades per W. Chafe:

TAULA 45 *	<i>Face-to-face conversation</i>	<i>Prepared speeches</i>	<i>Academic prose</i>
<i>Prepositions</i>	8.5 %	11.3 %	13.9 %

(\* Percentatges sobre el total de mots - Taula elaborada a partir de les dades de Biber, 1988: 255, 265 i 269)

Tots dos autors, tant W. Chafe com D. Biber, coincideixen amb M.A.K. Halliday a indicar el paper destacat dels sintagmes preposicionals en l'augment de la densitat lèxica i en la conformació del caràcter compacte de la llengua escrita<sup>205</sup>.

#### 4.7.2- UNITATS, PROCEDIMENTS I CRITERIS DE CODIFICACIÓ

D. Biber (1988) busca les preposicions a partir d'una llista extensa<sup>206</sup> i les compta conjuntament, sense distingir entre les funcions que puguin desenvolupar. En el present treball he actuat al contrari en tots dos aspectes. D'una banda, he pres en consideració només els sintagmes preposicionals introduïts per les cinc preposicions febles (*a, amb, de, en i per*), perquè són sens dubte les preposicions més habituals en català i perquè deixar de banda les anomenades *preposicions fortes* evita els problemes sobre la naturalesa preposicional o adverbial de molts d'aquests mots (vg. § 4.5.1).

El segon aspecte en què diferiré de D. Biber és en la distinció de les funcions que realitzen els sintagmes preposicionals, que poden actuar com a complements nominals o

<sup>205</sup> Hi ha també les dades de L.W. Smith (1994), que troba un 5.8 % a *informal conversation*, un 6.8 % a *formal interview* i un 9.1 % a *written essay*. Els tres percentatges, com es pot veure, són notablement inferiors als de D. Biber (1988), tot i que es mostren progressius en la mateixa direcció.

<sup>206</sup> D. Biber (1988) treballa amb una llista de 41 preposicions angleses que extreu de R. Quirk, S. Greenbaum, G. Leech i J. Svartvik (1985: 665-667). Aquestes preposicions són les següents: *against, amid, amidst, among, amongst, at, besides, between, by, despite, during, except, for, from, in, into, minus, notwithstanding, of, off, on, onto, opposite, out, per, plus, pro, re, than, through, throughout, thru, to, toward, towards, upon, versus, via, with, within, without*.

com a complements verbals<sup>207</sup>. Els principals usos dels sintagmes preposicionals introduïts per les preposicions febles en català, d'acord amb la bibliografia que he consultat (Fabra, 1956: 74-77 i 172-178 i Badia i Margarit, 1994: 195-233), són els següents:

- **Preposició *a*:**
  - . Complement indirecte
  - . Complement directe (excepcionalment)
  - . Complement d'objecte preposicional (o de règim verbal)
  - . Complementos circumstancials diversos
  
- **Preposició *amb*:**
  - . Complementos circumstancials diversos
  
- **Preposició *de*:**
  - . Complement del nom
  - . Complement de l'adjectiu
  - . Subjecte en infinitiu (opcionalment)
  - . Complement directe en infinitiu (opcionalment)
  - . Complement d'objecte preposicional (o de règim verbal)
  - . Atribut
  - . Complementos circumstancials diversos
  
- **Preposició *en*:**
  - . Complement d'objecte preposicional (o de règim verbal)
  - . Complementos circumstancials diversos
  
- **Preposició *per*:**
  - . Subjecte agent de la veu passiva
  - . Complementos circumstancials diversos

En resum, segons la llista anterior es podria concloure que les preposicions *a*, *amb*, *en* i *per* introdueixen generalment complements verbals, mentre que la preposició *de* pot introduir indistintament complements verbals i complements nominals. Però, encara que no consti en la bibliografia consultada, és possible de trobar amb tota normalitat sintagmes introduïts per les preposicions *a*, *amb*, *en* i *per* en funció de complement del

---

<sup>207</sup> També poden actuar com a complements oracionals, sovint en funcions d'organització del text relacionades amb els anomenats marcadors discursius, que tractaré en apartats posteriors.

nom o de l'adjectiu (*macarrons a la napolitana, promoció a gran escala, tendència a llarg termini, relació amb el món urbà, literatura en català, pis per llogar, qüestió clau per a l'economia, productes necessaris per fer l'obra*, etc.). Com que aquest fet pot tenir importància en la conformació d'un estil nominal o verbal per al text, l'he tingut en compte en la codificació dels usos de les preposicions, que ha quedat de la manera següent:

**A/ Preposicions a, amb, de, en i per en complements del nom** (dependents o no d'una nominalització).

**B/ Preposicions a, amb, de, en i per en complements de l'adjectiu** (o d'un participi que es comporta com un adjectiu).

**C/ Preposicions a, amb, de, en i per en complements del verb**, de les menes següents:

- Complement directe precedit, excepcionalment, de preposició.
- Complement indirecte.
- Complement d'objecte preposicional (o de règim verbal).
- Complementos circumstancials diversos.

**D/ Altres usos de les preposicions a, amb, de, en i per**, que inclouen els següents:

- Preposició *a* en les expressions *a més, a part de, a veure, al cap de, a lo millor, és a dir, a tot estirar, de cap a peus, portar a terme, a continuació, quant a*.
- Preposició *de* en la perífrasi verbal *haver de + infinitiu*.
- Preposició *de* en les expressions *de fet, de seguida, de manera que, d'una banda, tot d'una, des de, abans de, dins de, lluny de, a causa de, per tal de, a més de*.
- Preposició *en* en les expressions *en conseqüència, en fi, en lloc de, en definitiva, en efecte, en canvi, en tant que, tenir en compte, de nou en nou, avui en dia*.

- Preposició *per* en les expressions *per exemple, per tant, per cert, per tal de, per descomptat, per se, per contra, per bé que, per això*.

En el cas de la preposició *a*, he inclòs en el còmput dels tres primers apartats (A, B i C) els usos en les fórmules compostes *fins a, cap a, respecte a, com a, per a* i *gràcies a*.

He fet els càlculs sobre fragments de 2500 mots, que, tal com assenyalava en el punt 3.5.1, són suficients per a uns resultats fiables en aquesta categoria gramatical. El procediment emprat ha estat la recerca automàtica amb *TACT* de les formes *a, al, als, amb, per, pel, pels, en, de, d', del* i *dels*, amb la posterior distinció manual entre els valors de *en* com a preposició, com a pronom feble i com a article personal, i la codificació, també manual, de cada preposició segons la funció<sup>208</sup>. De les formes de l'adjectiu-pronom demostratiu *aquest/a/s/es* i dels indefinits *un/a/s/es* i *algun/a/s/es*, que he comptat, com es veurà més endavant, per explicar un aspecte parcial de la recerca, n'he fet un còmput automàtic amb *TACT*.

D'altra banda, com ja s'ha assenyalat en el capítol metodològic, he descomptat les ocurrències contigües de qualsevol preposició, possibles en la llengua oral en casos com els següents:

“i **dels dels** perills inherents a una esclavitud\_”

“**en en** aquest moment **de de** transició molt molt clar\_”

“que és un moment molt confús per tot un conjunt **de de** raons\_”

(Classe individu D, pàg. 217-221, lín. 10275, 10495 i 10502)

De tota manera, el nombre d'aquestes ocurrències contigües ha estat relativament baix (l'he calculat de l'ordre de l'1.9 % respecte del total d'ocurrències de les preposicions) i, per tant, tampoc no hauria alterat gaire els resultats en el còmput global.

---

<sup>208</sup> Les instruccions de *TACT* han estat *Usebase - F2 Query - Shift F3 KWIC - F9 Writing to Disc L22 L23*.

### 4.7.3- RESULTATS I DISCUSSIÓ

Les preposicions han seguit la progressió ja habitual en aquesta recerca, de la conversa informal a la prosa acadèmica. En aquest cas ha estat un índex baix en la conversa, que creix en la classe i que augmenta definitivament en la prosa. De totes maneres, una anàlisi detallada de les dades aportarà més informació sobre el comportament d'aquesta categoria. Per als resultats més generals, vegem la taula següent:

<b>TAULA 46 *</b>	<b>CONVERSA INFORMAL</b> (27254 mots)	<b>CLASSE MAGISTRAL</b> (26374 mots)	<b>PROSA ACADÈMICA</b> (26468 mots)
<b>Preposicions febles</b>	8.8 % (2398)	14.7 % (3877)	17.0 % (4499)

(\* Percentatges sobre el total de mots i, entre parèntesis, dades absolutes d'ocurrències – Test  $\chi^2$ : probabilitat d'atzar inferior al 0.0001 %)

D'acord amb la taula 46, el nombre de preposicions augmenta 5.9 punts entre la conversa informal i la classe magistral, i 2.3 punts entre aquesta i la prosa acadèmica, de manera que la classe ocupa la típica posició intermèdia, bastant decantada en aquesta ocasió cap a l'escrit. En relació amb les dades de D. Biber (1988), els resultats que he obtingut en català són gairebé idèntics pel que fa a la conversa i més alts pel que fa a la classe i a la prosa acadèmica. Si es té en compte que D. Biber hi incloïa diverses preposicions fortes, els índexs del català són proporcionalment encara més alts.

La taula 47 (vg. pàgina següent) forneix informació sobre els percentatges d'aparició de cada preposició. Segons les dades d'aquesta taula, la preposició *de* és la més abundant i *amb* la més escassa; l'ordre de freqüència de les cinc preposicions febles en les dades totals és el següent: 1- *de* (6.9 %), 2- *a* (3.2 %), 3- *en* (1.5 %), 4- *per* (1.3 %), 5- *amb* (0.6 %). Gairebé la meitat de les ocurrències de preposicions en la conversa informal (3.8 % del 8.8



%), la meitat en la classe magistral (7.3 % del 14.7 %) i més de la meitat en la prosa acadèmica (9.5 % del 17.0 %), corresponen a la preposició *de*.

<b>TAULA 47*</b>	<b>CONVERSA INFORMAL</b> (27254 mots)	<b>CLASSE MAGISTRAL</b> (26374 mots)	<b>PROSA ACADÈMICA</b> (26468 mots)	<b>TOTAL (mitjana)</b> (80096 mots)
<i>a</i>	3.0 % (817)	3.6 % (949)	3.0 % (794)	3.2 % (2560)
<i>amb</i>	0.6 % (163)	0.5 % (132)	0.8 % (212)	0.6 % (507)
<i>de</i>	3.8 % (1036)	7.3 % (1925)	9.5 % (2514)	6.9 % (5475)
<i>en</i>	0.4 % (109)	2.0 % (527)	2.2 % (582)	1.5 % (1218)
<i>per</i>	1.0 % (272)	1.3 % (343)	1.5 % (397)	1.3 % (1012)

(\* Percentatges sobre el total de mots i, entre parèntesis, dades absolutes d'ocurrències – Test  $\chi^2$ : probabilitat d'atzar inferior al 0.0001 % en les preposicions *a* i *per*, del 0.0051 % en la preposició *amb*, i inferior al 0.0001 % en les preposicions *de* i *en*)

Pel que fa a la progressió entre gèneres, es pot observar que les preposicions *a*, *amb* i *per* es mantenen en unes quantitats bastant similars en tots tres gèneres, amb unes petites diferències a favor de la classe magistral o de la prosa acadèmica. La preposició *en*, en canvi, sorprèn per l'ús escàs que se n'ha fet en la conversa, cosa que intentaré explicar més avall. La preposició *de* és la que presenta unes xifres més rellevants: augmenta considerablement de la conversa (3.8 %) a la classe (7.3 %) i a la prosa acadèmica (9.5 %), i s'erigeix així en la principal responsable de l'augment en el nombre global de preposicions.

La qüestió que cal veure a continuació és la distribució de les preposicions segons les funcions que desenvolupen, cosa que es reflecteix a la taula 48:

<b>TAULA 48 *</b>	<b>CONVERSA INFORMAL</b> (27254 mots)	<b>CLASSE MAGISTRAL</b> (26374 mots)	<b>PROSA ACADÈMICA</b> (26468 mots)
<b>Sintagmes preposicionals complements del nom</b>	2.5 % (681)	6.7 % (1767)	9.4 % (2488)
<b>Sintagmes preposicionals complements de l'adjectiu</b>	0.0 % (0)	0.6 % (158)	1.4 % (371)
<b>Sintagmes preposicionals complements del verb</b>	5.2 % (1417)	5.5 % (1450)	4.8 % (1270)
<b>Altres usos de les preposicions</b>	1.1 % (300)	1.9 % (501)	1.4 % (371)

(\* Percentatges sobre el total de mots i, entre parèntesis, dades absolutes d'ocurrències – Test  $\chi^2$ : probabilitat d'atzar inferior al 0.0001 % a totes les files excepte en els complements del verb, en què és del 0.1295 %)

En la distribució de funcions, les preposicions han mostrat un comportament ben remarcable, que comento a continuació cas per cas:

**A/ La funció complement del nom** ha registrat un gran augment des de la conversa (2.5 %) fins a la prosa acadèmica (9.4 %). Aquest increment s'ha donat en totes les preposicions i, especialment, en la preposició *de*<sup>209</sup>. En molts casos i especialment en les preposicions *a*, *de* i *per*, els noms complementats són derivats de verbs, cosa que dóna un matís semàntic d'acció a la construcció resultant:

“La **transformació d'**aquest estructuralisme (...) hauria donat pas a la **reducció a** una mena de científisme (...)”

“Una cosa és l'ús i l'altra l'abús d'aquesta **preocupació pel** discurs (...)”

“(...) un llibre que comença dient que l'**eixamplament del** món espanyol **per** tal d'abraçar un nou continent requeria una **reordenació intel·lectual de** la realitat (...)”

(Escrit individu C, pàg. 266-269, lín.12417-12419, 12541 i 12554-12556)

<sup>209</sup> Les ocurrències de la preposició *de* en funció de complement del nom han suposat el 2.2 % en la conversa, el 5.6 % en la classe i el 7.8 % en la prosa acadèmica. Aquesta preposició, que ja hem vist que constitueix el gruix principal de les xifres globals, és també la responsable indiscutible de l'augment registrat en aquest tipus de complements.

**B/ La funció complement de l'adjectiu** té els índexs més baixos i, alhora, els contrastos més grans, ja que evoluciona des del 0.0 % de la conversa al 0.6 % a la classe i creix encara més fins a l'1.4 % de la prosa. Sobre aquest tipus de complement cal remarcar que ha estat gairebé inexistent en les preposicions *a* i *de*. En canvi, en la preposició *per* ha estat més freqüent que els complements del nom. En la majoria dels casos, els mots complementats no han estat adjectius sinó participis, que han conservat part de la naturalesa verbal i han donat, també, un matís semàntic d'acció a les construccions on han aparegut<sup>210</sup>:

“(…) s’havia acabat de desarticlar el vell sistema de relacions de l’economia catalana amb l’exterior, molt **girat cap a** la Mediterrània (…)”

“(…) una gran capacitat per a mobilitzar treball amb tractes **adequats a** les circumstàncies de cada moment.”

“Però aquesta situació canvià amb la guerra comercial **iniciada per** la França colbertiana contra Holanda amb mesures aranzelàries discriminatòries i **continuada** després **amb** guerra de debò (…)”

“(…) d’allà procedien més de la meitat de les importacions **fetes pel** port de Barcelona (…)”

(Escrit individu B, pàg. 255-259, lín. 11956-11958, 12034-12035, 12130-12133 i 12150-12151)

**C/ La funció complement del verb** s’ha mantingut bastant estable, amb un lleuger augment en la classe magistral (5.5 %) i un descens, també petit, en la prosa acadèmica (4.8 %). De fet, però, si es consideren les xifres respecte del total de preposicions, que són cada vegada més grans, el que hi ha és una reducció del pes relatiu dels complements del verb.

**D/ Els altres usos de les preposicions febles** han estat més alts en la classe magistral, cosa que podria explicar, almenys parcialment, perquè és en aquest gènere on han aparegut més algunes expressions com *és a dir*, *per exemple*, *per tant*, etc. (vg. § 5.5.3). Es dona la particularitat que la preposició *amb* no ha aparegut en usos d’aquesta mena.

---

<sup>210</sup> Es podria dir, potser, que els sintagmes preposicionals complements del nom i de l’adjectiu es formarien prototípicament amb la preposició *de*, mentre que els precedits per les altres preposicions febles sovint —encara que no sempre— tenen traces d’una naturalesa verbal, transformada però no del tot esvaïda.

En resum, la distribució de funcions ha palesat que el paper fonamental en el contrast entre els tres gèneres correspon als sintagmes preposicionals complements del nom i de l'adjectiu. Aquest fet el podem copsar encara amb més claredat a la taula 49, en la qual he agrupat aquests complements en oposició amb els complements del verb, amb vista a una millor definició de l'estil nominal o verbal dels textos:

<b>TAULA 49 *</b>	<b>CONVERSA INFORMAL</b> (2098 s. prep.)	<b>CLASSE MAGISTRAL</b> (3376 s. prep.)	<b>PROSA ACADÈMICA</b> (4129 s. prep.)
<b>Sintagmes preposicionals complements nominals</b>	32.5 % (682)	57.0 % (1924)	69.2 % (2857)
<b>Sintagmes preposicionals complements verbals</b>	67.5 % (1416)	43.0 % (1452)	30.8 % (1272)

(\* Percentatges sobre el total de sintagmes preposicionals, excepte els de l'apartat "Altres usos de les preposicions" i, entre parèntesis, dades absolutes d'ocurrències – Test  $\chi^2$ : probabilitat d'atzar inferior al 0.0001 %)

D'acord amb la taula 49, el nombre de sintagmes preposicionals complement del nom ha anat augmentant i el de complements del verb ha anat disminuint, fins a conformar una situació percentual inversa. La classe magistral s'ha decantat, en aquest cas, cap a l'escrit, de manera que entre el 32.5 % de la conversa i el 57.0 % de la classe ja s'ha consumat la major part del gir de l'estil verbal cap al nominal en l'ús de les preposicions.

Seguint el criteri de calcular els percentatges amb diversos denominadors per a cada unitat, per tal de millorar-ne l'adequació explicativa, en el cas present he relacionat els sintagmes preposicionals complements del nom amb el nombre total de noms (taula 50) i el conjunt dels sintagmes preposicionals amb nombre total de clàusules (taula 51). Les dades resultants d'aquests càlculs es poden veure a continuació:

<b>TAULA 50 *</b>	<b>CONVERSA INFORMAL</b> (3652 noms)	<b>CLASSE MAGISTRAL</b> (5248 noms)	<b>PROSA ACADÈMICA</b> (6485 noms)
<b>Noms complementats per un sintagma preposicional</b>	23.0 % (840)	55.8 % (2928)	77.3 % (5013)

(\* Percentatges sobre el total de noms i, entre parèntesis, dades absolutes d'ocurrències – Test  $\chi^2$ : probabilitat d'atzar inferior al 0.0001 %) <sup>211</sup>

<b>TAULA 51 *</b>	<b>CONVERSA INFORMAL</b> (4769 clàusules)	<b>CLASSE MAGISTRAL</b> (3666 clàusules)	<b>PROSA ACADÈMICA</b> (2488 clàusules)
<b>Sintagmes preposicionals per clàusula</b>	0.50	1.07	1.81

(\* Dades mitjanes sobre el total de clàusules)

Segons la taula 50, el nombre mitjà de noms que estan complementats per un sintagma preposicional també mostra una evolució considerable: si en la conversa són menys d'una quarta part (el 23.0 %), en la classe ja són més de la meitat (55.8 %) i en la prosa acadèmica superen les tres quartes parts (77.3 %). Per tant, en aquest gènere hi ha d'haver menys del 22.7 % <sup>212</sup> de noms sense complementar, ja que molts d'aquests noms poden estar acompanyats per un o diversos adjectius atributius. Cal pensar, a més, que l'augment dels percentatges d'aquesta taula ha resultat atenuat perquè hi ha actuat com a denominador el nombre total de noms que, en ser creixent cap a l'escrit, ha suavitzat la força ascendent de la progressió.

<sup>211</sup> En el càlcul dels percentatges d'aquesta taula ha calgut fer una operació especial: al nombre total de noms, que actua com a denominador, s'hi ha hagut de restar els noms que són darrere d'una preposició, dins d'un sintagma preposicional complement del nom, i així s'ha obtingut un denominador més ajustat.

<sup>212</sup> Aquesta xifra podria ser una mica més alta, ja que és el resultat d'una ponderació: pot haver-hi noms amb dos sintagmes preposicionals que hagin fet mitjana amb altres que no en tinguin cap.

Les dades de la taula 51, en canvi, han experimentat l'efecte contrari. El denominador, en aquest cas, ha estat el nombre de clàusules, que en ser decreixent de la conversa a l'escrit ha motivat un augment progressivament més alt de l'índex. Així, les xifres han resultat contundents: si en la conversa informal hi ha una mitjana de 0.50 sintagmes preposicionals per clàusula, en la classe magistral n'he trobat 1.07 i en la prosa acadèmica han arribat a 1.81. Es demostra, doncs, que la prosa acadèmica i, en menor mesura, la classe magistral, empen els sintagmes preposicionals com un recurs de construcció dels enunciat en una mesura molt superior a la conversa quotidiana.

Com he dit anteriorment en el comentari dels resultats de cada preposició, caldria explicar l'escassa aparició de la preposició *en* a la conversa informal (0.4 %), en contrast amb els percentatges de la classe (2.0 %) i de la prosa acadèmica (2.2 %). Per buscar-hi una justificació, he comptat les aparicions de l'adjectiu-pronom demostratiu *aquest/a/s/es* i dels indefinits *un/a/s/es* i *algun/a/s/es*, que són mots que provoquen el canvi de *a* per *en* per raons fonètiques en els complements circumstancials de lloc. Els resultats de la recerca d'aquests mots han estat els següents:

<b>TAULA 52 *</b>	<b>CONVERSA INFORMAL</b>	<b>CLASSE MAGISTRAL</b>	<b>PROSA ACADÈMICA</b>
<i>aquest/a/s/es</i>	177	430	200
<i>un/a/s/es</i>	498	765	601
<i>algun/a/s/es</i>	39	55	29
<b>TOTAL</b>	714	1250	830

(\*Dades absolutes d'ocurrències – Test  $\chi^2$ : probabilitat d'atzar inferior al 0.0001 % per a *aquest/a/s/es* i *un/a/s/es*, i del 17.0927 % per a *algun/a/s/es*)

Les dades de la taula 52 poden ajudar-nos poc en la qüestió que tenim plantejada. Efectivament, la relació que s'estableix entre les xifres és ben diferent que la que hi ha en l'ús de la preposició *en* en els diversos gèneres (vg. taula 47). A més, en el cas concret d'*algun/a/s/es*, els resultats han quedat gairebé invalidats per l'alt percentatge obtingut en el test  $\chi^2$ .

Una altra possible justificació podria ser en la naturalesa mateixa de la preposició que ens ocupa. Així, l'ús més habitual de la preposició *en* per indicar lloc figurat, en confluència amb la preposició *a*, que indica lloc concret, podria ser una abstracció pròpia de la llengua escrita i, per extensió, de la llengua oral formal<sup>213</sup>. L'altra funció principal de *en*, la introducció de determinats complements d'objecte preposicional (o de règim verbal), tampoc no sembla donar-se gaire en la conversa, malgrat que tingui un nombre de verbs alt, ja que molts dels verbs que regeixen la preposició *en* pertanyen a un registre bastant formal (*consistir en*, *complaure's en*, *tenir interès en*, etc.). Sigui com sigui, el fet incontestable és que *en* ha aparegut molt menys en la conversa espontània i que això mateix ja pot ser un fet rellevant per a l'estudi d'aquesta preposició<sup>214</sup>.

---

<sup>213</sup> Un fet important a destacar en relació amb aquesta preposició és la dificultat que té el transcriptor per distingir entre *a*, *en* i *amb* en català central, ja que les tres formes tendeixen a neutralitzar-se. La transcripció, en definitiva, el que fa és convertir la llengua oral en un tipus especial de llengua escrita, amb una servitud inevitable a les regles d'aquesta darrera. Des d'aquest punt de vista, els resultats també podrien explicar-se per les convencions emprades pels transcriptors del *Corpus de Català Contemporani* de la Universitat de Barcelona en abordar aquesta qüestió espinosa.

<sup>214</sup> La conclusió que se'n desprèn és que l'ús de la preposició *en* per a complements de lloc podria ser un invent del català escrit. Aquesta preposició consistiria en realitat en la realització de la preposició *a* més una consonant *n* de valor epentètic, com argumenta E. Clua (1996a i 1996b) a partir de l'ús d'aquestes preposicions en tortosí actual:

“Des d'aquesta perspectiva, podem postular una única preposició subjacent *a*, i explicar les realitzacions [ɔn] com a fruit d'una epèntesi consonàntica desencadenada per la necessitat de no violar alguns principis de la GU, concretament el Principi del contorn obligatori i el Principi de l'obertura sil·làbica.” (Clua, 1996a: 166)

“I try to show that the preposition *a* is only appearing in adjunct and indirect object Pps. This forces us to account for the form [an] which this preposition takes in certain contexts: specifically when it is followed by a vowel initial determiner or pronoun.” (Clua, 1996b: 29)

Per la seva banda, P. Sancho (1995) manté que hi ha una diferència de significat entre *a* i *en*, però acaba afirmant el caràcter fonològic de la distinció:

“Amb tot, aquesta distribució basada en diferències semàntiques tendeix a ser substituïda, en major o menor grau segons els casos, per una distribució basada en principis fonològics.” (Sancho, 1995: 201)

Pel que fa a la distribució de resultats per textos, es pot veure a les taules 53 i 54. Les dades d'aquestes taules mostren unes bandes de dispersió bastant reduïdes (de 2.7 punts per a les converses, 3 punts per a les classes i 4.2 punts per als escrits) que tenen, però, un comportament molt diferent de l'una a l'altra. Entre la conversa i la classe apareixen perfectament definides i es mantenen a una bona distància (3 punts de diferència). A l'extrem superior de la taula, en canvi, les bandes de la classe magistral i de la prosa acadèmica se superposen considerablement: hi ha dos escrits, el de l'individu E (14.6 %) i el de l'individu C (15.8 %), que queden per sota de les dues classes amb el valor més alt, la de l'individu B (16.3 %) i la de l'individu D (també 16.3 %). Aquesta indefinició, la més notable que he trobat fins al moment fora del cas dels adverbis en *-ment*, dona lloc a una destacada banda d'intersecció, que queda reflectida a la taula 55 (vg. pàgina següent)<sup>215</sup>:

<b>TAULA 53 *</b>	<b>S. PREPOSICIONALS</b>	<b>CONVERSA</b>
<b>CONVERSA 1</b>	Total s. preposicionals .....	8.1 %
<b>CONVERSA 2</b>	Total s. preposicionals .....	8.5 %
<b>CONVERSA 3</b>	Total s. preposicionals .....	7.6 %
<b>CONVERSA 4</b>	Total s. preposicionals .....	9.5 %
<b>CONVERSA 5</b>	Total s. preposicionals .....	10.3 %

(\* Percentatges sobre el total de mots)

<sup>215</sup> Com en el cas dels adverbis, aquesta confusió de les bandes de la classe magistral i de la prosa acadèmica informa sobre la proximitat entre aquests dos gèneres pel que fa a l'ús dels sintagmes preposicionals, tret en què sembla que també preval la formalitat i el grau d'especificitat associats amb els usos acadèmics de la llengua.



<b>TAULA 54 *</b>	<b>S. PREPOSICIONALS</b>	<b>CLASSE MAGISTRAL</b>	<b>PROSA ACADÈMICA</b>
<b>INDIVIDU A</b>	Total s. preposicionals .....	13.6 %	17.5 %
<b>INDIVIDU B</b>	Total s. preposicionals .....	16.3 %	18.8 %
<b>INDIVIDU C</b>	Total s. preposicionals .....	13.3 %	15.8 %
<b>INDIVIDU D</b>	Total s. preposicionals .....	16.3 %	18.5 %
<b>INDIVIDU E</b>	Total s. preposicionals .....	14.3 %	14.6 %

(\* Percentatges sobre el total de mots)

<b>TAULA 55 * - S. PREPOSICIONALS (Dispersió de les dades)</b>		
19		<b>BANDA</b>
18	Escrit B Escrit D	<b>DE LA</b>
17	Escrit A	<b>PROSA ACADÈMICA</b>
16	Classe B Classe D	
15	Escrit C	<b>BANDA D'INTERSECCIÓ</b>
14	Classe E Escrit E	
13	Classe A Classe C	<b>BANDA DE LA CLASSE MAGISTRAL</b>
12		
11		
10	Conversa 5	
9	Conversa 4	<b>BANDA</b>
8	Conversa 1 Conversa 2	<b>DE LA</b>
7	Conversa 3	<b>CONVERSA INFORMAL</b>
6		

(\* Percentatges sobre el total de mots)

A continuació donaré uns exemples de textos per mostrar de manera pràctica el comportament dels sintagmes preposicionals. He seleccionat dos fragments, un de la conversa 3 i un altre de l'escrit de l'individu D, els més extrems de la mostra en aquest apartat. Els dos fragments són de 271 i 270 mots respectivament. Hi he subratllat les preposicions febles i hi he assenyalat amb blau el sintagma preposicional sencer que introdueixen (no he indicat amb color, en tres ocasions, la preposició *de* dins de la perífrasi d'obligació *haver de*):

“ALL: també\  
però no\  
va ser xulo eh\  
hem vist--

VIE: ha set interessant\  
sí\  
ALL: veure el paisatge a l'hivern  
és gust\  
bueno no sé\  
i sense- --

SEP: i neu n'heu vista\  
VIE: sí\  
ALL: sí\  
VIE: que passa que no ens ha nevat\_  
i no du- --  
i\_  
i dubto que hagi nevat\  
per allà ont estàvem no ha nevat cap dia\  
el que passa que al--  
que a les vores de la carretera sí que hi havia neu\  
ALL: clar\  
a les zones més altes\

VIE: però molt maco per això home\  
ALL: però bueno\  
no\  
ha estat bé\  
VIE: oh\  
n'hi han molts de castells eh\  
vull dir n'hem mi- --  
n'hem n'hem mirats\_  
ai\  
n'hem anat a visitar  
què n'hem visitat\  
potser deu/  
ALL: un dia un\_  
quatre\_  
cinc\_  
sis\_  
set\_

uns deu\  
 VIE: uns deu i n'hi han\_  
 n'hi n'hi han bastants eh\  
 que siguin\_  
 que siguin\_  
 m [dels](#)\_  
[d'aquests que](#)\_  
 ALL: càtars\  
 càtars\  
 VIE: càtars\  
 càtars\  
 n'hi ha uns--  
 veiam  
 un dos tres quatre\_  
 estan numerats\  
 cinc sis set vuit nou deu\_  
 onze dotze\  
 dotze càtars\  
 i tots els altres són badies\  
 són\_  
 ALL: no\  
 n'hi ha que també són castells\  
 VIE: bueno\  
 ALL: però que no són feudals\  
 VIE: també són castells\  
 ALL: aï\  
 no són\_  
 VIE: ja no són medievals\  
 ALL: aï\  
 RIA: però molt bé\  
 ???: gràcies\  
 MAG: però\_  
 aquí\_  
 RIA: està ben ben\  
 MAG: [per exemple](#)\_  
 nosaltres tots aquests castells que hi havien\_  
[per aquí](#)\_  
 ah bueno\  
 ja estan tots eliminats\_  
 tots\  
 i [en aquests llocs](#) encara\_  
 encara no sé què és això\  
 VIE: aquests dos són molt diferents\  
 aquests [d'aquí](#) eren\_  
 eren\_  
 eren z [de de](#)\_  
 suposo\  
 que eren feu- --  
[de feudals](#)\  
 devia ser [de](#)--  
 MAG: feudals\

sí\  
VIE: de  
MAG: Castellbisbal\_  
VIE: de te --  
de terratinents”

(Conversa 3, pàg. 74-75, lín. 3433-

3527)

“La definició d'uns temes, d'un estil i d'una moralitat va ser un procés que arrencà, com ja hem vist, de dins de la crisi romàntica dels anys trenta i quaranta. La consolidació de l'historicisme i de la mitificació ruralitzant, d'una banda, i l'abandonament dels temes clau per a qualsevol literatura adulta, de l'altra, són dos aspectes del mateix procés de definició. L'encarrilament de la literatura en català en aquesta direcció no va ser un procés tan simple com podríem suposar. Testimoniats diversos ens permeten d'intuir que, almenys en els primers anys, trobà unes certes resistències i provocà vacil·lacions i dubtes. En aquest sentit hem d'interpretar, per exemple, el picament de peus de Piferrer en els Jocs Florals de 1842, descontent pel caràcter tan marcadament historicista de les comunicacions premiades i per l'olor de resclosit de la cerimònia de lliurament dels premis, o l'explícit reconeixement del Gaiter del Llobregat que els seus versos eren acollits amb un escepticisme evident entre els seus mateixos amics, o el deambular d'altres personatges pels marges del renaixentisme emergent sense decidir-se a participar-hi, com Jaume Tió.

Implicà d'entrada una confrontació amb altres pràctiques literàries amb les quals havia de coexistir i respecte a les quals els homes de la Renaixença varen prendre, en ocasions, posicions d'un fort antagonisme. Com que el que de debò intentaven de fer no era purament construir una literatura i, menys encara, salvar una llengua, sinó resistir en el pla cultural els efectes disgregadors de la transformació per la qual passava el país, la pugna ideològica amb altres corrents, que més endavant explicitem, podia i havia de ser necessàriament de gran intensitat.”

(Escrit individu D, pàg. 276, lín. 12843-12866)

Com en els apartats anteriors, el contrast entre els dos fragments és ben remarcable. Així, el fragment de conversa conté un total de 19 preposicions, que han donat lloc a 13 sintagmes preposicionals (els sis restants corresponen a cinc construccions estroncades i a una repetició contigua de la preposició *de*). En el fragment de prosa acadèmica, en canvi, hi ha 57 preposicions, 54 de les quals encapçalen un sintagma preposicional (n'hi ha tres dins de la perífrasi verbal *haver de* + infinitiu).

Les funcions d'aquests sintagmes preposicionals s'han distribuït tal com indica la taula 56 (vg. pàgina següent), que mostra amb claredat les tendències en els dos fragments analitzats. D'una banda, la conversa té un nombre més reduït de sintagmes preposicionals, la gran majoria dels quals han exercit com a complements verbals. De

l'altra, la prosa acadèmica presenta un nombre molt més alt de sintagmes preposicionals, la majoria dels quals són complements nominals. El nombre de preposicions és tan alt en la prosa acadèmica que fins i tot la funció de complement del verb hi presenta més ocurrencies que en la conversa informal (15 en la prosa contra 9 en la conversa). La complementació nominal per mitjà de sintagma preposicional, i fins i tot la verbal, es mostra, per tant, com un procediment essencial en l'arquitectura del discurs escrit acadèmic, mentre que la conversa s'ha de basar necessàriament en altres recursos<sup>216</sup>.

<b>TAULA 56 *</b>	<b>Fragment de conversa informal</b>	<b>Fragment de prosa acadèmica</b>
<b>Complement nominal</b>	1	34
<b>Complement verbal</b>	9	15
<b>Altres funcions</b>	4	5
<b>TOTAL</b>	13	54

(\* Dades absolutes d'ocurrencies)

#### **4.7.4- CONCLUSIONS DE L'APARTAT**

Els sintagmes preposicionals han estat més freqüents en la prosa acadèmica que en la classe magistral i més en aquesta que en la conversa informal, tant en les dades absolutes com en el càlcul, amb unes desigualtats encara més accentuades, respecte del nombre de clàusules. El principal motiu d'aquest resultat ha estat l'evolució de la preposició *de*,

---

<sup>216</sup> Els verbs de la conversa, que en nombre són força més que en la prosa acadèmica, han de ser complementats per altres sistemes. A tall d'exemple, es pot esmentar que en aquests dos fragments el nombre de pronoms febles ha estat de 13 per a la conversa i de només 1 per a la prosa, xifres que per si mateixes ja contraresten amb escriure les de la complementació verbal per mitjà de sintagmes preposicionals.

que, especialment en la funció de complement del nom, ha augmentat espectacularment el nombre d'aparicions des de la conversa fins a l'escrit. Les altres preposicions febles han experimentat també un augment, però molt més moderat. En conjunt, per tant, és la complementació nominal la que augmenta d'una manera més remarcable.

D'altra banda, la preposició *en* ha presentat un comportament peculiar, amb un nombre molt baix d'ocurrències en la conversa, que es podria explicar en una anàlisi de la naturalesa oral o escrita d'aquesta preposició.

La distribució de resultats per textos ha estat ben satisfactòria entre la conversa i la classe, i ben poc entre la classe i la prosa, amb una superposició i una confusió notable de les dades.

Els resultats sobre l'ús dels sintagmes preposicionals ha aportat encara més informació per a la caracterització dels estils nominal i verbal en els gèneres que analitzem. Així, els sintagmes preposicionals, al costat dels adjectius, complementen més els noms en la llengua escrita i conformen uns sintagmes nominals més llargs i de més densitat informativa. En la llengua oral, en canvi, la major quantitat de verbs es correspon amb un nombre menor de noms i de mots que els complementin (sintagmes preposicionals i adjectius). En definitiva, aquest conjunt de quatre categories, noms, verbs, adjectius i preposicions, en nombres diferents i amb distribucions complementàries, es combinen d'una manera efectiva per a la construcció dels textos, més densa i compacta en el cas de la llengua escrita i més mòbil i fluida en la llengua oral. La classe magistral, amb el caràcter que té de discurs oral formal i especialitzat, es perfila amb claredat com a gènere intermedi entre la conversa i la prosa acadèmica.

## 4.8- LES NOMINALITZACIONS

### 4.8.1- FONAMENTACIÓ CONCEPTUAL I HIPÒTESIS

El paper de les nominalitzacions en el discurs oral i escrit ha estat destacat per D. Biber (1988), W. Chafe (1982), W. Chafe i J. Danielewicz (1987), G. Redeker (1984) i L.W. Smith (1994). El primer les relaciona amb el discurs no dependent de la situació, juntament amb diversos tipus de clàusules subordinades i en troba significativament més a *academic prose* que a *prepared speeches* i, sobretot, més que a *face-to-face conversation*<sup>217</sup>. A continuació podem veure'n les dades:

<b>TAULA 57 *</b>	<i>Face-to-face conversation</i>	<i>Prepared speeches</i>	<i>Academic prose</i>
<i>Nominalizations</i>	0.9 %	2.1 %	3.6 %

(\* Percentatges sobre el total de mots - Taula elaborada a partir de les dades de Biber, 1988: 255, 265 i 269)

W. Chafe (1982) i W. Chafe i J. Danielewicz (1987), per la seva part, les consideren essencials per a la densitat lèxica i per a l'organització sintàctica del text escrit:

"The most characteristic integrative device in our samples of written language was nominalization. (...) A nominalization allows a notion which is verbal in origin to be inserted into an idea unit as if it were a noun" (Chafe, 1982: 39)

---

<sup>217</sup> D. Biber (1988: 227) cerca automàticament els mots acabats en *-tion*, *-ment*, *-ness* i *-ity*, i els plurals corresponents.

“Nominalizations are a principal means whereby a single clause can be constructed from what might otherwise have been several clauses.” (Chafe i Danielewicz, 1987: 100)

W. Chafe (1982) també relaciona, com D. Biber, les nominalitzacions amb la independència d'una situació immediata i amb el discurs abstracte, al costat d'altres trets com els verbs en passiva:

"Nominalization seems to be another device associated with detachment, besides functioning as an integrative device as described above. Not only does nominalization allow predications to be integrated within larger sentences; it also supresses involvement in actions in favor of abstract reification." (Chafe, 1982: 46)

Les característiques d'integració sintàctica, de densitat lèxica, de no implicació amb la situació i d'abstracció en el discurs justifiquen, per a W. Chafe (1982), que només hi hagi un 0.5 % de nominalitzacions en l'oral i, en canvi, un 5.6 % en l'escrit<sup>218</sup>.

G. Redeker (1984), tot inspirant-se en les conclusions de W. Chafe, fa un control més estricte de la variable gènere i arriba a uns resultats diferents, amb uns percentatges bastant més baixos en la llengua escrita<sup>219</sup>:

<b>TAULA 58 *</b>	<i>Spoken story</i>	<i>Spoken explanation</i>	<i>Written story</i>	<i>Written explanation</i>
<i>Nominalizations</i>	0.3 %	0.5 %	0.8 %	1.0 %

(\* Percentatges sobre el total de mots - Taula elaborada a partir de les dades de Redeker, 1984: 50)

<sup>218</sup> W. Chafe (1982) inclou en el grup de les nominalitzacions els derivats postverbals (sense sufix) i, en canvi, només hi compta els noms derivats de noms i de verbs (descarta els noms derivats d'adjectius). Pel que fa a W. Chafe i J. Danielewicz (1987), no especifiquen els criteris aplicats en el càlcul.

<sup>219</sup> G. Redeker (1984) tampoc no informa del tipus de mots i de sufixos inclosos en el càlcul. El percentatge tan baix que obté en llengua escrita pot ser motivat per la naturalesa dels escrits que analitza, que estan obtinguts en condicions no reals i de poca elaboració, com ja he comentat en el punt 3.4.1.



En el seu treball de 1987, W. Chafe i J. Danielewicz primfilen més que en el de W. Chafe (1982) i distingeixen entre quatre gèneres de discurs diferents: converses, conferències, cartes i articles científics. Els resultats indiquen que en la llengua científica oral hi ha quasi tantes nominalitzacions com en les cartes, i només les converses i els articles científics hi ocupen posicions extremes:

<b>TAULA 59 *</b>	<i>Conversations</i>	<i>Lectures</i>	<i>Letters</i>	<i>Academic papers</i>
<b>Nominalizations</b>	2.7 %	5.6 %	5.5 %	9.2 %

(\* Percentatges sobre el total de mots - Taula elaborada a partir de les dades de Chafe i Danielewicz, 1987: 100)

W. Chafe i J. Danielewicz justifiquen aquestes dades de la manera següent<sup>220</sup>:

“Many words which originated as nominalizations have become standard items of the academic vocabulary, without which academic writers would be unable to say the kinds of things they like to say. Such words then become easy for speakers to use as well. (...) However, the current use of such nouns by speakers no longer shows a productive act of nominalizing, but only the use of a technical term. It may be for this reason that the occurrence of nominalizations is relatively high in spoken language, and specially high in lectures.” (Chafe i Danielewicz, 1987: 100)

Aquesta observació podria anar acompanyada d’una de més general —i, potser, més simple—, sobre la naturalesa intermèdia dels gèneres *lectures* i *letters* en el *continuum* oral-escrit, motivada per la confluència d’altres paràmetres contextuais (camp general o específic, mode espontani o preparat, tenor formal o informal, etc.) que acompanyen la

<sup>220</sup> Crida l’atenció que els resultats de W. Chafe (1982) i els de W. Chafe i J. Danielewicz (1987) difereixin bastant entre si, que difereixin encara més amb els de D. Biber (1988) i, sobretot, amb els de G. Redeker (1984). La clau s’ha de trobar en la disparitat de criteris lèxics en el còmput de cada autor o bé en les diferències entre els paràmetres contextuais que condicionen els textos de les mostres respectives.

distinció acústic / gràfic. Des d'aquest punt de vista, els resultats en aquests quatre gèneres estarien dotats de plena congruència i no necessitarien justificació suplementària<sup>221</sup>.

Sense que ofereixi dades quantitatives d'ús, M.J. Cuenca (2000b) assenyala també la relació de la nominalització amb la llengua escrita. Aquesta autora, a partir de M.T. Serafini (1992), estableix la diferència entre un *estil cohesionat*, exclusiu dels escriptors experts, i un *estil segmentat*, que és el propi de la llengua oral, dels escriptors principiants i dels escriptors experts que, per raons estilístiques, opten per utilitzar-lo. La referència a la nominalització —i, en general, als sintagmes nominals— apareix quan es parla de transformar un text d'estil segmentat a estil cohesionat:

“El segon tipus de regles de reducció consisteixen a convertir estructures oracionals simples en aposicions o en sintagmes nominals amb un substantiu d'origen verbal (nominalitzacions), i a simplificar clàusules subordinades substituint-les per sintagmes semànticament equivalents.” (Cuenca, 2000b: 107)<sup>222</sup>

Els estudis que acabo de ressenyar, per més que presentin disparitats entre si, coincideixen en la tendència general a una presència superior de nominalitzacions en els gèneres escrits. Això permet de fixar com a hipòtesi per a aquest apartat un nombre més elevat de derivats nominals en la prosa acadèmica que en la classe magistral i un nombre també més elevat en la classe que en la conversa informal<sup>223</sup>.

---

<sup>221</sup> Hi ha, encara, el treball de L.W. Smith (1994: 58) que ja he citat en apartats anteriors. Aquest autor ha trobat un 3.5 % de nominalitzacions a *informal conversation*, un 5.0 % a *formal interview* i un 9.6 % a *written essay*. Lamentablement, L.W. Smith tampoc no concreta què ha comptat com a *nominalization*.

<sup>222</sup> Com es pot suposar, l'afirmació de M.J. Cuenca té més conseqüències que les estrictament relatives a la nominalització, ja que la substitució de clàusules subordinades per sintagmes nominals implica, de fet, una reducció del nombre de subordinades i del nombre total de clàusules en la llengua escrita, qüestió que abordaré en l'apartat corresponent (vg. § 5.2.3).

<sup>223</sup> Quan ja tinc fets els càlculs d'acord amb els procediments emprats per la majoria d'autors, he descobert a P.J. Gillie (1957) i J.A. Devito (1967) una curiosa fórmula per determinar el nivell d'abstracció d'un text, que inclou la presència de sufixos nominalitzadors:

“The language elements and the directions for calculating the formula are as follows:

- a. the number of finite verbs per 200 words plus
- b. the number of definite articles and their nouns per 200 words plus
- c. the constant 36 minus
- d. the number of nouns of abstraction (words ending in the suffixes *-ness*, *-ment*, *-ship*, *-dom*, *-nce*, *-ion* and *-y*) times 2.

## 4.8.2- UNITATS, PROCEDIMENTS I CRITERIS DE CODIFICACIÓ

La nominalització és el procediment derivatiu a través del qual es poden formar noms a partir de radicals no nominals:

“1. Nominalització. — De radicals de característiques diferents podem obtenir noms si els combinem amb sufixos nominals.” (Cabré i Rigau, 1986: 49)

El terme *nominalització*, utilitzat per designar els noms obtinguts per aquest procediment, sembla típic de la bibliografia anglòfona. En la tradició catalana es troben les expressions —menys còmodes— *noms derivats d'altres noms*, *noms derivats d'adjectius* i *noms derivats de verbs* (Fabra, 1956: 116-138).

Als efectes del present càlcul, he considerat nominalitzacions els noms derivats d'adjectius i els noms derivats de verbs per mitjà d'un sufix i he exclòs, per tant, els derivats postverbals (obtinguts sense l'adjunció de cap sufix)<sup>224</sup>. P. Fabra (1956: 122-129) ofereix una llista considerable de sufixos nominalitzadors. D'aquesta llista m'han interessat els que signifiquen 'qualitat de' o 'acció i resultat de', que coincideixen a oferir al parlant la possibilitat de triar entre l'ús del nom derivat o del mot originari, adjectiu o verb, en un estil més nominal, més adjectival o més verbal de construcció dels enunciats. Vegeu-ne l'exemple següent:

---

The result is the abstraction score. According to Gillie the scores are to be interpreted as follows: 0-18=very abstract; 19-30=abstract; 31-42=fairly abstract; 43-54=standard; 55-66=fairly concrete; 67-78=concrete; 79-90=very concrete.” (Devito, 1967: 356)

La inclusió dels noms amb article definit com a factor que minva l'abstracció s'entén quan es fa la precisió (Devito, 1967: 358) que s'hi compten només els noms que no són complementats per cap adjectiu, sintagma o clàusula. Així, tant els verbs com els noms no complementats s'oposen als noms derivats i donen una mesura de l'abstracció del text. D'acord amb aquesta fórmula, J.A. Devito analitza els discursos orals i escrits (més aviat formals i especialitzats en tots dos casos) de deu professors universitaris i troba un índex d'abstracció de 50.8 (=standard) en la llengua oral formal i de 39.9 (=fairly abstract) en la llengua escrita dels mateixos individus sobre el mateix tema.

<sup>224</sup> L'exclusió que he fet dels postverbals ha tingut una raó pràctica prioritària: els noms derivats amb sufix es poden trobar amb una relativa facilitat a partir d'una recerca automàtica de les terminacions; els postverbals, en canvi, haurien exigut una anàlisi minuciosa de tots els noms del text.

Un correcte *enfocament* del tema no pot deixar de banda l'*aparició* d'una problemàtica personal ...

Si *enfoquem* bé el tema no podem deixar de banda que *apareix* una problemàtica personal...

(Exemple preparat *ad hoc*)<sup>225</sup>

Amb aquest criteri, els sufixos nominalitzadors seleccionats han estat els següents:

#### A/ Noms derivats d'adjectius - Significat 'qualitat de':

*-ància*: ignorant ® ignorància

*-ència*: negligent ® negligència

*-esa*: feble ® feblesa

*-etat*: breu ® brevetat

*-ïtat*: digne ® dignitat

*-tat*: igual ® igualtat

*-isme*: real ® realisme

#### B/ Noms derivats de verbs - Significat 'acció i resultat de':

*-atge*: muntar ® muntatge

*-(a)ció*: aprovar ® aprovació

*-(i)ció*: abolir ® abolicció

*-(altres)ció*: percebre ® percepció

*-ió*: comprendre ® comprensió

*-(a)ment*: abonar ® abonament

*-(e)ment*: conèixer ® coneixement

*-(i)ment*: abaratir ® abaratiment

He descartat altres sufixos que, perquè no tenen un caràcter tan abstracte, perquè no signifiquen *qualitat de* o *acció i resultat de* o, simplement, perquè no són tan productius o habituals, donen lloc a substantius que no se solen substituir pel verb o per l'adjectiu del qual provenen (p. ex.: *fregall*, de *fregar*; *observatori*, d'*observar*; *xisclat*, de *xisclar*; *xerradissa*, de *xerrar*; *assegurança*, d'*assegurar*; etc.).

---

<sup>225</sup> Aquest exemple confirmaria l'opinió de M.A.K. Halliday que l'elaboració nominal es contraposa a la complexitat sintàctica, ja que als dos noms derivats del primer exemple han correspost dos verbs subordinats en el segon.

També he descartat, dins del grup nombrós de noms derivats amb els sufixos seleccionats, aquells que difícilment s'intercanviarien amb el mot del qual provenen o amb el qual es relacionen morfològicament, adjectiu o verb, potser perquè el parlant ja no els hi associa. Vegem a continuació alguns exemples d'aquests mots exclosos:

- . *reportatge* (no admet l'intercanvi amb el verb *reportar*)<sup>226</sup>
- . *habitació* (no admet l'intercanvi amb el verb *habitar*)
- . *situació* (no admet l'intercanvi amb el verb *situar*)
- . *societat* (no admet l'intercanvi amb l'adjectiu *social*)
- . *ajuntament* (no admet l'intercanvi amb el verb *ajuntar*)

Sí que hi he comptat, en canvi, el gruix principal dels noms derivats, com *hostilitat*, *moralitat*, *frustració*, *estabilització*, *modernització*, *devaluació*, *transformació*, *mitificació*, *vacil·lació*, *confrontació*, *desenvolupament* i un llarg etcètera, que permeten clarament d'optar entre el nom derivat i l'adjectiu o el verb originari<sup>227</sup>.

D'acord amb el criteri ja seguit en altres apartats, he calculat el percentatge de nominalitzacions no només sobre el total de mots sinó també sobre una altra unitat que sigui més assimilable al fenomen analitzat. En aquest cas, ho he fet sobre el total de noms, que es poden considerar la unitat de referència per a les nominalitzacions. La recerca ha estat feta sobre fragments de 5000 mots, és a dir, sobre el total de la mostra<sup>228</sup>.

---

<sup>226</sup> De fet, segons el *Diccionari etimològic* de Jordi Bruguera (pàg. 738-739), el mot català *reportatge* no prové de *reportar*, sinó que és un neologisme a partir del francès *reportage*, format sobre l'anglès *reporter* ('periodista'), provinent al seu torn del francès antic *reporteur* ('que reporta') i, aquest darrer sí, format, com el català *reportar*, a partir del llatí PORTARE. Tampoc *societat* no prové de *social*, sinó que ambdós són derivats —ja ho eren en llatí— de *soci* < SOCIUS. Per això he especificat "el mot del qual provenen o amb el qual es relacionen morfològicament".

<sup>227</sup> Reconec que en algun cas la línia divisòria entre els substantius intercanviables i els no intercanviables pot ser difícil de traçar. De fet, aquest criteri l'he adoptat sobretot per excloure del còmput alguns mots en què era molt clar que no era possible l'intercanvi. D'altra banda, els resultats que presento a continuació mostren unes tendències tan accentuades que les possibles vacil·lacions no hi poden haver tingut cap repercussió significativa.

<sup>228</sup> El procediment seguit ha estat la recerca manual de les nominalitzacions a partir de la llista que *TACT* ofereix de tots els mots del text ordenats alfabèticament per les terminacions, és a dir, com en un diccionari invers. Les instruccions de *TACT* han estat *TactFreq - Reverse Alphabetical*.

### 4.8.3- RESULTATS I DISCUSSIÓ

Un cop més, la recerca ha confirmat les hipòtesis i ha ofert uns resultats fins i tot més rotunds que els de les investigacions prèvies. Vegem-ho a les taules 60 i 61:

<b>TAULA 60 *</b>	<b>CONVERSA INFORMAL</b> (27254 mots)	<b>CLASSE MAGISTRAL</b> (26374 mots)	<b>PROSA ACADÈMICA</b> (26468 mots)
<b>Nominalitzacions</b>	0.1 % (40)	2.0 % (544)	4.3 % (1148)

(\* Percentatges sobre el total de mots i, entre parèntesis, dades absolutes d'ocurrències – Test  $\chi^2$ : probabilitat d'atzar inferior al 0.0001 %)

<b>TAULA 61 *</b>	<b>CONVERSA INFORMAL</b> (3652 noms)	<b>CLASSE MAGISTRAL</b> (5248 noms)	<b>PROSA ACADÈMICA</b> (6485 noms)
<b>Nominalitzacions</b>	1.1 % (40)	10.1 % (544)	17.7 % (1148)

(\* Percentatges sobre el total de noms i, entre parèntesis, dades absolutes d'ocurrències – Test  $\chi^2$ : probabilitat d'atzar inferior al 0.0001 %)

La taula 60 reflecteix la presència minúscula de derivats nominals en la conversa informal. La proporció —més ben dit, la desproporció— amb la classe magistral i amb la prosa acadèmica és espectacular: d'1 a 20 i d'1 a 43, respectivament<sup>229</sup>.

---

<sup>229</sup> Si es té en compte que, com hem vist en el punt 4.2.3, els noms han passat del 13.4 % en la conversa informal al 24.5 % en la prosa acadèmica, amb un diferencial d'11.1 punts, les xifres de la taula 60 certifiquen que un 4.2 % —més o menys, la tercera part— de l'increment de noms ha estat a càrrec de les nominalitzacions.

Els càlculs de la taula 61, que es fan sobre el total de noms, suavitzen aquestes quantitats, ja que el nombre de noms és progressivament més alt cap a l'escrit i, en actuar com a denominador, dóna lloc a una xifra cada vegada més atenuada. Malgrat això, d'acord amb el conjunt de les dues taules, es pot afirmar categòricament que les nominalitzacions són molt més presents en la classe magistral i en la prosa acadèmica que en la conversa informal.

La major intervenció de la nominalització es relaciona, doncs, amb la utilització d'un lèxic més específic, propi dels temes tractats en les classes i en els escrits, i amb una estructuració sintàctica de caràcter nominal, que prefereix els noms derivats abans que els adjectius o els verbs corresponents. A més, els noms derivats tenen un caràcter sempre més abstracte que els mots originaris i augmenten la mitjana de longitud de mot, pel fet elemental que l'adjunció de sufixos allarga el nombre de sons, de síl·labes i de caràcters de les paraules. En conjunt, aquestes quatre característiques —especificitat, estructuració sintàctica, abstracció i longitud de mot— cooperen en l'augment de la densitat lèxica en els dos gèneres acadèmics objecte del present estudi.

En el desglossament dels resultats per sufixos es pot observar les particularitats en la freqüència de cadascun, encara que tots, sense excepció, coincideixen a aparèixer més en la prosa acadèmica que en la classe magistral i més en la classe magistral que en la conversa informal. La taula 62 (vg. pàgina següent) mostra els diferents graus d'utilització dels sufixos nominals en els textos<sup>230</sup>. El més usat és el sufix *-(a)ció*, amb un total de 562 ocurrències, a gran distància del menys usat, el sufix *-ància*, amb només 25. A continuació ofereixo l'ordenació dels sufixos segons el nombre d'aparicions, en dues llistes, una primera (llista A) tal com apareixen a la taula, amb la vocal temàtica, i una segona (llista B) amb el total de cada sufix:

---

<sup>230</sup> He exclòs d'aquesta taula els sufixos *-esa*, *-atge*, *-etat* i *-tat*, a causa del baixíssim índex d'aparició que tenen en les condicions requerides.

TAULA 62 *	CONVERSA INFORMAL	CLASSE MAGISTRAL	PROSA ACADÈMICA	TOTAL
<i>-ància</i>	1	9	15	25
<i>-ència</i>	2	56	86	144
<i>-ïtat</i>	2	94	178	274
<i>-isme</i>	0	40	66	106
<i>-(a)ció</i>	15	169	378	562
<i>-(i)ció</i>	3	33	46	82
<i>-(altres)ció</i>	2	57	176	235
<i>-ió</i>	1	34	79	114
<i>-(a)ment</i>	14	19	74	107
<i>-(e)ment</i>	0	27	33	60
<i>-(i)ment</i>	0	6	17	23
<b>TOTAL</b>	40	544	1148	1732

(\* Dades absolutes d'ocurrències)

**A/ Ordenació dels sufixos segons la freqüència (amb la vocal temàtica):**

- |                        |                    |                     |
|------------------------|--------------------|---------------------|
| 1. <i>-(a)ció</i>      | 5. <i>-ió</i>      | 9. <i>-(e)ment</i>  |
| 2. <i>-ïtat</i>        | 6. <i>-(a)ment</i> | 10. <i>-ància</i>   |
| 3. <i>-(altres)ció</i> | 7. <i>-isme</i>    | 11. <i>-(i)ment</i> |
| 4. <i>-ència</i>       | 8. <i>-(i)ció</i>  |                     |

**B/ Ordenació dels sufixos segons la freqüència (totals):**

- |                                    |                                   |
|------------------------------------|-----------------------------------|
| 1. <i>-ció</i> : 879 ocurrències   | 5. <i>-ió</i> : 114 ocurrències   |
| 2. <i>-ïtat</i> : 274 ocurrències  | 6. <i>-isme</i> : 106 ocurrències |
| 3. <i>-ment</i> : 190 ocurrències  | 7. <i>-ància</i> : 25 ocurrències |
| 4. <i>-ència</i> : 144 ocurrències |                                   |



La distribució de resultats, en el cas de les nominalitzacions, mostra unes bandes perfectament delimitades i sense encavalcaments. Com es pot observar a les taules 63 i 64, totes les converses s'han agrupat entre el 0.6 % de la conversa 5 i l'1.6 % de les converses 2 i 4; les classes, entre el 6.3 % de l'individu E i el 13.3 % de l'individu A, i els escrits entre el 14.1 % de l'individu E i el 20.7 % de l'individu D.

<b>TAULA 63 *</b>	<b>NOMINALITZACIONS</b>	<b>CONVERSA</b>
<b>CONVERSA 1</b>	Total nominalitzacions .....	0.7 %
<b>CONVERSA 2</b>	Total nominalitzacions .....	1.6 %
<b>CONVERSA 3</b>	Total nominalitzacions .....	1.1 %
<b>CONVERSA 4</b>	Total nominalitzacions .....	1.6 %
<b>CONVERSA 5</b>	Total nominalitzacions .....	0.6 %

(\* Percentatges sobre el total de noms)

<b>TAULA 64 *</b>	<b>NOMINALITZACIONS</b>	<b>CLASSE MAGISTRAL</b>	<b>PROSA ACADÈMICA</b>
<b>INDIVIDU A</b>	Total nominalitzacions ...	13.3 %	18.6 %
<b>INDIVIDU B</b>	Total nominalitzacions ...	11.8 %	18.3 %
<b>INDIVIDU C</b>	Total nominalitzacions ...	9.2 %	16.8 %
<b>INDIVIDU D</b>	Total nominalitzacions ...	10.0 %	20.7 %
<b>INDIVIDU E</b>	Total nominalitzacions ...	6.3 %	14.1 %

(\* Percentatges sobre el total de noms)

La dispersió de les classes (7 punts) i dels escrits (6.6 punts) no ha estat suficient per produir cap encavalcament. Ben al contrari, la distància entre la classe amb l'índex més alt i l'escrit que el té més baix és de 0.8 punts, i la que hi ha entre la conversa amb l'índex més alt i la classe amb l'índex més baix és de 4.7 punts<sup>231</sup>. Les bandes de dispersió són les que es reflecteixen a la taula següent:

<b>TAULA 65 * - NOMINALITZACIONS (Dispersió de les dades)</b>			
22			
20	Escrit D		
18	Escrit A	Escrit B	<b>BANDA</b>
16	Escrit C		<b>DE LA</b>
14	Escrit E		<b>PROSA ACADÈMICA</b>
12	Classe A Classe B		
10	Classe D		<b>BANDA</b>
8	Classe C		<b>DE LA</b>
6	Classe E		<b>CLASSE MAGISTRAL</b>
4			
2	Conv. 2	Conv. 3	Conv. 4
0	Conversa 1	Conversa 5	
	<b>BANDA DE LA CONVERSA INFORMAL</b>		

(\* Percentatges sobre el total de noms)

Pel que fa a la comparació de les produccions de cada individu, es dona una coincidència interessant: els individus E i C són els que tenen uns índexs més baixos tant en la prosa acadèmica (individu E: 14.1 % i individu C: 16.8 %), com en la classe magistral (individu E: 6.3 % i individu C: 9.2 %), de manera que coincideixen a usar menys derivats nominals en tots els seus textos acadèmics, tant en els orals com en els escrits.

Com en els apartats anteriors, presentaré alguns exemples per a una visió més clara del funcionament de la nominalització en els textos. Amb aquest objectiu he triat uns

<sup>231</sup> Aquestes distàncies són més simètriques en el càlcul sobre el nombre total de mots, que no he reflectit en aquestes taules per no donar un excés d'informació. En aquest cas, la distància entre la classe amb l'índex més alt (2.5 %) i l'escrit que el té més baix (3.6 %) és d'1.1 punts, i la que hi ha entre la conversa amb l'índex més alt (0.2 %) i la classe amb l'índex més baix (1.2 %) és d'1 punt.

fragments de l'escrit de l'individu A, que és el que presenta un índex més gran de nominalitzacions respecte del total de mots (4.8 %)<sup>232</sup>. Com es pot observar, he assenyalat amb vermell els derivats nominal formats amb els sufixos de la llista:

“Una probable **derivació** d'això pot ser, a més a més d'una natural **participació** creixent en el producte global, la **multiplicació** del nombre d'efectius —en **absència** d'un procés de **concentració** empresarial. Res d'això succeeix en el nostre cas. **L'aportació** dels promotors-constructors al producte del sector, mesurada en metres quadrats construïts, mostra una **tendència** a llarg termini a un augment lleuger, que, d'altra banda, resulta menys pronunciat que el dels altres grans promotors. I si ens fixem en el nombre de persones que cada dècada (natural) inicien o finalitzen llur tasca de **promoció**, s'observa una simptomàtica disparitat entre les dues categories de grans promotors.”(...)

“Com és sabut, la indústria de la **construcció**, sobretot el subsector de la **construcció** d'edificis, ha entrat amb gran retard en el procés de **desenvolupament** capitalista. En aquesta branca de la **producció**, la **revolució** industrial és un fenomen dels nostres dies. No sense raó, s'ha titllat la **construcció** de ser la ventafocs del procés **d'industrialització** (entès com a **transformació** de les tècniques i **l'organització** de la **producció** tendents a elevar fortament la **productivitat**). El segle passat, el grau de **concentració** de l'estructura empresarial era extremadament baix. Els mètodes, les tècniques i els materials de la **construcció** ben poc variaren.”(...)

“En un estadi ulterior de **desenvolupament**, tindrà lloc, d'una banda, una creixent **integració** del procés productiu en l'esfera d'acció empresarial del contractista general, limitant-se a subcontractar les feines molt especialitzades, en les quals li resulta antieconòmic mantenir permanentment ocupada mà d'obra; d'altra banda, en aquest tercer estadi apareixeran grans empreses constructores caracteritzades per operar a molt major escala que les cases constructores tradicionals, la qual cosa mena a una gradual **dependència** d'aquestes —els petits constructors—, relegades a fer obres de **reparació** i **manteniment** o a dedicar-se a treballs altament especialitzats, generalment per encàrrec de les grans societats constructores.”

(Escrit individu A, pàg. 246-250, lín. 11544-11552, 11698-11706 i 11740-11749)

En el primer fragment hi ha diversos exemples de construccions amb derivats nominal que podrien haver estat oracions amb un verb, en forma personal o no. Vegem la possible transformació de dos d'aquests exemples:

**Versió nominal:** Una probable **derivació** d'això pot ser, a més a més d'una natural **participació** creixent en el producte global, la **multiplicació** del nombre d'efectius...

**Versió verbal:** D'això, probablement, se'n **deriva** que, a més de **participar** cada vegada més en el producte global, es **multiplicarien** el nombre d'efectius...

---

<sup>232</sup> A diferència d'altres apartats, no dono cap fragment de conversa que serveixi de contrapunt, perquè la presència de derivats nominal en aquest gènere és tan minsa que només serviria per veure com no n'hi ha cap o gairebé cap. Tampoc no ofereixo cap fragment de les classes, en les quals les nominalitzacions tenen un comportament similar al dels escrits, només que hi apareixen en un nombre menor.

**Versió nominal:** L'**aportació** dels promotors-constructors al producte del sector, mesurada en metres quadrats construïts, mostra una **tendència** a llarg termini a un **augment** lleuger...

**Versió verbal:** Els promotors-constructors **aporten** al producte del sector un nombre de metres quadrats construïts que **tendeix** a llarg termini a **augmentar** lleugerament...

Cal destacar que en les oracions d'estil més verbal hi ha aparegut dues vegades la conjunció *que*, per la necessitat d'articular entre si els verbs, que conformen oracions sintàcticament dependents o subordinades. A més, hi han minvat les ocurrències de la preposició *de*, en consonància amb les dades del punt 4.7.3. En el segon exemple, també he transformat en infinitiu el substantiu derivat postverbal *augment*, ja que es tracta del mateix fenomen, encara que els postverbals no hagin format part del meu còmput.

En el segon fragment s'hi observa l'altra gran font d'aparició de derivats nominals: la terminologia. És a dir, hi ha diversos mots com ara *construcció*, *industrialització*, *producció*, *productivitat*, etc., que no semblen aparèixer-hi perquè substitueixin declaradament els verbs originals —tot i que aquesta substitució seria possible—, sinó que hi són perquè expressen conceptes complexos elaborats en el si del camp d'especialitat al qual pertany l'escrit que analitzem.

El tercer fragment, en darrer lloc, forneix una mostra de dues solucions diferents, una a favor de la derivació nominal i l'altra a favor de l'estil verbal. La construcció del fragment es basa en l'articulació de les dues expressions *d'una banda* i *d'altra banda*, separades per punt i coma. En el primer cas, el redactor ha optat per la nominalització i en el segon, en canvi, ha triat una construcció exclusivament verbal:

En un estadi ulterior de desenvolupament, **tindrà lloc**, d'una banda, una creixent **integració** **de**l procés productiu... (un verb, dos substantius i la preposició *de*)

...d'altra banda, en aquest tercer estadi **apareixeran** grans empreses constructores... (un verb)

Aquestes dues oracions podien haver estat com segueix:

En un estadi ulterior de desenvolupament, d'una banda, s'**integrarà** cada vegada més el procés productiu... (un verb)

...d'altra banda, en aquest tercer estadi **tindrà lloc** l'**aparició de** grans empreses constructores... (un verb, dos substantius i la preposició *de*)

Cal observar que en cada transformació ha aparegut o desaparegut la preposició *de*, que ha acompanyat el derivat nominal i l'ha complementat amb el sintagma que en la construcció verbal és el subjecte de l'oració.

Així, s'ha vist en aquests tres fragments com l'individu A ha triat les articulacions sintàctiques que han donat, generalment, un caràcter més nominal al conjunt del seu text.

#### **4.8.4- CONCLUSIONS DE L'APARTAT**

Les nominalitzacions han estat molt més freqüents en la prosa acadèmica i en la classe magistral que en la conversa informal. De fet, en aquest gènere, la presència de derivats nominals a partir dels sufixos seleccionats ha estat molt baixa o, en alguns casos, gairebé nul·la.

En el detall dels resultats per sufixos s'ha observat una aclaparadora majoria de *-ció*, seguit d'*-itat* i de *-ment*. Les bandes de dispersió han resultat bastant ben delimitades, més que en altres ocasions, sense superposicions i, fins i tot, amb una distància important entre la conversa i la classe.

En definitiva, s'han confirmat les hipòtesis inicials segons les quals la nominalització és un recurs més utilitzat pels escriptors i, a una certa distància, pels oradors, i no és gens freqüent en la conversa quotidiana. La presència de derivats nominals està, doncs, relacionada amb el grau d'especificitat i d'abstracció d'un text i contribueix decisivament, al costat dels altres elements que he comentat en els apartats anteriors, a la densitat lèxica i a la conformació d'un estil preferentment nominal.

## 4.9- LA VARIACIÓ LÈXICA, LA REPETICIÓ LÈXICA I LA LONGITUD DE MOT

### 4.9.1- FONAMENTACIÓ CONCEPTUAL I HIPÒTESIS

Els trets de variació lèxica, repetició lèxica i longitud de mot no són esmentats per M.A.K. Halliday i, en canvi, sí que han estat tractats per K. Perera (1984), W. Chafe i J. Danielewicz (1987) i D. Biber (1988). Sembla lògic que, al costat de la proporció entre mots lexicals i mots gramaticals, el grau de variació o de repetició dels mots hagi de tenir importància en la densitat informativa del discurs. En aquesta línia, ha estat afirmat per diversos autors que els discursos orals són, en general, més redundants informativament que els discursos escrits. Vegeu, per exemple, les paraules següents de K. Perera (1984) i de W. Chafe i J. Danielewicz (1987):

“The use of co-ordination, repetition and rephrasing causes spoken language to have a low level of **lexical density** and a high level of **redundancy**. This means that lexical words are spaced out, separated by grammatical words, and that a high number of words is used to convey a given amount of information. Obviously, these features aid not only the speaker’s production of language but also the listener’s reception. Such assistance is necessary, since the listener does not have the reader’s opportunity to backtrack in order to retrieve a forgotten word. (...) language produced in such a way —basically, oral language— tends to lack variety of vocabulary and sentence structure and (...) it is more prolix than writing need be. In speech, lexical repetition is both less avoidable and more tolerable than in writing. The writer can avoid it because, unlike the speaker, he has time to search his memory for the precise word he wants without losing track of his train of thought.” (Perera, 1984: 184-187)

“As a consequence of these differences, speakers tend to operate with a narrower range of lexical choices than writers. Producing language on the fly, they hardly have time to shift through all the possible choices they might make, and may typically settle on the first words that occur to them. The result is that the vocabulary of spoken language is more limited in variety.” (Chafe i Danielewicz, 1987: 88)

Ambdues citacions coincideixen en les raons per les quals es justifica el fenomen. Així, la major redundància de la llengua oral es deu al poc temps de què disposa el parlant per fer una tria lèxica precisa i acurada, i del temps també reduït que té l'oient per processar-la. I, inversament, la major varietat lèxica de la llengua escrita és motivada per la possibilitat de què gaudeix l'escriptor de cercar els mots i les estructures més adequades al pensament que vol expressar i per la possibilitat que té el lector de tornar sobre el text ja llegit per reinterpretar-lo<sup>233</sup>.

Si aquests autors enfoquen la repetició com una irregularitat o un defecte —“*less avoidable and more tolerable than in writing*”—, n’hi ha d’altres que la plantegen com una forma d’organització típica de la llengua oral (D. Tannen, 1987; C. Blanche-Benveniste, 1993; G. Herrero, 1995; A.M. Vigara, 1995; A. Briz, 1998 i J. Freixa, 1998). La repetició d’un mot pot ser causada per una indecisió del parlant i llavors forma part dels fenòmens associats a la producció *on-line*, o bé pot ser un recurs de cohesió textual al servei de la progressió informativa del text<sup>234</sup>. En aquesta línia, C. Blanche-

---

<sup>233</sup> Una explicació semblant és la que dona G. Brown (1978):

“Speakers, like tired writers, often get into a lexical or syntactic rut, especially in conversation. The speaker, to reiterate, is speaking under pressure of time. Often he has a choice of using a word or structure that he has just produced or searching around for a new or more satisfactory one, which will mean that he has to manufacture channel holding devices. If he is under stress, especially for instance in an interview situation, he is very likely to use again an already activated word or structure. This is a frequent feature in conversations where it seems to have a very positive solidary social function.” (Brown, 1978: 274)

<sup>234</sup> Crec que en moltes de les repeticions de la mostra es poden combinar les dues explicacions. Això sembla especialment clar en les classes magistrals, en què al costat de la funció necessàriament didàctica de la repetició, el parlant aprofita la reiteració dels mots que acaba de dir i que té més disponibles en memòria per facilitar-se la tasca de construcció del discurs i el manteniment d’un torn de parla llarg. Vegem-ne uns exemples:

“vol dir\_  
es es restableix un marc de **protecció** ee comercial normal\_  
diguem\_  
s'acaba amb tot el muntatge autàrquic\_  
i es passa es configura **una economia que és una economia**\_  
ee **proteccionista normal**\_  
diguem així\_  
**proteccionista normal**\_  
no autàrquica\_  
eh/  
de fet **el aranzel** que s'instaura a continuació\_  
**l'aranzel** del mil nou-cents seixanta **és un aranzel protector**\_  
**és un aranzel protector**\_

Benveniste (1993) analitza un tipus particular de repetició en francès que anomena “*répétition de lexique et glissement vers la gauche*”, que consisteix en operacions de repetició amb intercalació d’incisos, i G. Herrero (1995) parla de les *construcciones-eco*, que “*repiten de forma exacta o con ligerísimos cambios que reflejan las nuevas condiciones de la enunciación (...) enunciados previos o parte de los mismos, producidos en el turno de habla inmediatamente anterior.*” (1995: 126)<sup>235</sup>.

Per la seva banda, D. Tannen (1987), que és sens dubte qui ha parlat d’una manera més completa sobre els diversos valors de la repetició en la conversa, ha postulat que el major nombre de repeticions en la llengua oral no s’ha atribuir a una “pobresa lèxica”<sup>236</sup>, sinó que constitueix una manera particular de fer avançar la informació i de mantenir la cohesió en el discurs oral<sup>237</sup>.

---

que la protecció nominal mitjana se situa en el vint-i-cinc per cent\_  
que no és petita\_  
vint-i-cinc per cent\_  
i segons quins productes és molt més elevada\_  
pot ser del del del quaranta al cinquanta o més\_  
eh”

(Classe de l’individu A, pàg. 183, lín. 8703-8721)

<sup>235</sup> Uns exemples de *construcciones-eco* que dona la mateixa autora són els següents (1995: 126-127):

“(1) A: Juan no viene mañana a comer. B: ¿Juan no viene mañana a comer?”

“(2) A: Me duele un montón la cabeza. B: ¿Te duele qué? (...)”

“(4) A: Lo has hecho mal. B: ¿Lo has hecho mal, lo has hecho mal! ¡A la porra!”

“(5) A: Tráeme el desayuno, anda. B: ¡Que le traiga el desayuno! ¡Vamos! ¡Qué cara!”

<sup>236</sup> Aquesta concepció de la pobresa de vocabulari de la conversa ha portat a pensar que la parla col·loquial pot desenvolupar-se amb un repertori lèxic molt limitat. E.A. Nida (1967) avisa de l’error que pot suposar aquesta aproximació al tema:

“As the result of analyses of telephone conversations, some persons have concluded that if some five hundred words constituted about ninety per cent of the vocabulary of telephone conversations, certainly a thousand words should be enough to express almost anything that anyone wanted to say. However, this is a fatal mistake, for it is that extra ten per cent of the vocabulary which is so significant.” (Nida, 1967: 149)

<sup>237</sup> J. Freixa (1998) ha fet un resum excel·lent de l’aportació de D. Tannen, una part del qual ofereixo a continuació. També A.M. Vigara (1995), sense fer-hi referència explícita, assumeix les idees de D. Tannen i les sintetitza en un paràgraf especialment encertat:

“d) *Interacció*: la repetició té també diverses funcions interactives, com pot ser agafar o mantenir la paraula, demostrar interès en el missatge, evitar respostes, encadenar respostes, connectar les idees d’un parlant amb les d’un altre, ratificar la contribució de l’altre i incloure a la interacció una persona que no ha sentit l’expressió anterior.” (Freixa, 1998: 200)

“De algún modo, parece que es precisamente la oralidad, el modo específico de concebir y realizar la lengua hablada (no leída o recitada, ni escrita), la que propicia (y hasta precisa de) el *recurso*



Pel que fa a la longitud de mot, un tret que també afecta la densitat del text, ha estat tinguda en compte per pocs autors. Vegem, a continuació, el que n'han dit G.H.J. Drieman (1962) i D. Biber (1986a):

“From the quantitative results it appears that in written language as compared to spoken language (...) we find shorter texts, longer words, fewer words of one syllable, more words of more than one syllable, more attributive qualities (and) a more varied vocabulary.” (Drieman, 1962: 56)

“A more varied vocabulary reflects extensive use of words having very specific meanings. Longer words also convey more specific, specialized meanings than shorter ones; words become shorter as they are more frequently used and more general in meaning.” (Biber, 1986a: 394)

Com es pot veure, tots dos autors fan referència a la longitud de mot en relació amb la variació lèxica i el segon, a més, valora el paper que té en el grau d'especificitat del text<sup>238</sup>. Pel que fa a les dades concretes, D. Biber (1988) aporta les següents:

---

*espontáneo a la repetición (formal o conceptual)*, es decir, de la recurrencia, como mecanismo de organización y progresión discursiva por un lado, y de relación interlocutiva, por otro. La abundancia de los ejemplos y los múltiples valores significativos y de sentido del fenómeno sugieren, además, que la recurrencia no actúa normalmente como información sobrante o inútil en el coloquio, sino como estrategia esencial del modo en que las personas interactuamos durante nuestra actividad comunicativa.” (Vigara, 1995: 208)

En el discurs oral monologat de la classe magistral, la repetició també actua com a part d'una estratègia de reformulació, en la qual un enunciat es repeteix parcialment en un altre que, a parer del parlant, millora el primer.

<sup>238</sup> La longitud de mot és un tret que té molta significació en l'àmbit anglòfon, en el qual hi ha la tradició de referir-se als mots cultes com a *long words*. Efectivament, l'ús de mots derivats del llatí, que coincideixen a ser més llargs, per davant dels de procedència anglosaxona és una característica de l'anglès culte. F.N. Akinnaso (1982) ho assenyala de la manera següent:

“(...) through a series of experiments involving choice between synonym pairs, Levin, Long, and Schaffer (1981) have found that words whose etymology is Latin are favored over synonymous Anglo-Saxon words «under explicit instructions to be formal, or lacking such instructions, under conditions that are strongly coercive toward formal language» (p. 171)” (Akinnaso, 1982: 103)

En el cas del català la situació no és tan ostensible, ja que la immensa majoria dels mots són d'origen llatí, encara que els mots col·loquials tendeixen a ser evolutius i provenen del llatí vulgar, mentre que els cultismes solen ser manlevats directament al llatí clàssic, per via literària, i coincideixen amb l'anglès a ser paraules més llargues.

<b>TAULA 66 *</b>	<i>Face-to-face conversation</i>	<i>Prepared speeches</i>	<i>Academic prose</i>
<i>Word length</i>	4.1	4.4	4.8

(\* Nombre mitjà de caràcters per mot - Taula elaborada a partir de les dades de Biber, 1988: 255, 265 i 269)

Així, es percep un augment de la longitud de mot a *Prepared speeches* i a *Academic prose*, apreciable encara que no sigui espectacular<sup>239</sup>.

En conclusió, esperaré de trobar un índex més alt de variació lèxica en la llengua escrita i un índex més alt de repetició lèxica en la llengua oral. Igualment, la hipòtesi serà que la longitud de mot augmentarà en direcció a la llengua escrita.

#### **4.9.2- UNITATS, PROCEDIMENTS I CRITERIS DE CODIFICACIÓ**

La manera més habitual de calcular la variació lèxica és efectuar la suma del nombre total de mots d'un text, comptant-hi les repeticions (*tokens*), i fer un segon càlcul del nombre de mots sense comptar-hi les repeticions (*types*). La proporció resultant és la mesura de la variació (*type / token*) o de la repetició (*token / type*) en el text, que és explicada per G. Youmans (1991) de la manera següent:

“Earlier<sup>240</sup> quantitative studies (...) focus on the relation between the number of types and tokens in texts. Youmans 1990<sup>241</sup> expresses this relationship through type-token

<sup>239</sup> Les xifres de longitud de mot consignades per D. Biber no presenten pas fluctuacions gaire grans. Globalment, en els seus 23 gèneres, la longitud de mot varia només de 3.9 a 4.9 caràcters en les mitjanes per gènere (*Mean value*), i de 3.7 (*Minimum value*) a 5.1 (*Maximum value*) en les mitjanes de textos concrets (1988: 247-269).

<sup>240</sup> G. Youmans (1991) considera que els estudis sobre la relació *type / token* són primerencs (“*earlier*”) perquè ell planteja una eina d’anàlisi més sofisticada del repartiment i l’evolució de la densitat lèxica en cada porció del text. De totes maneres, per a l’abast i els objectius del present estudi, el càlcul de la relació *type / token* resulta més que suficient.

<sup>241</sup> G. Youmans (1990) mesura les corbes *type / token* de l’obra literària de diversos escriptors anglòfons i arriba a conclusions sobre l’estil i sobre la competència lèxica global d’aquests escriptors.

vocabulary curves, which are constructed in the following way: as a discourse unfolds, the total vocabulary (the number of types) is plotted against the total number of words (tokens) that have been used to that point in the text. For the first few words of a normal discourse, every new token is also a new vocabulary word; initially, then, the number of types equals the number of tokens. After the first repeated word, however, the number of tokens exceeds the number of types, and this difference increases with each repetition. Theoretically, if a discourse were long enough, the speaker's total vocabulary would be exhausted, and no new types could be added." (Youmans, 1991: 764-765)

Una altra descripció acurada de la relació entre l'extensió d'un text i l'extensió del vocabulari que conté és la que ofereix C. Müller (1968)<sup>242</sup>:

“VII - Extensión del vocabulario y riqueza lexical. Crecimiento de V.

Dado un texto (y una norma), se comenzará por contar las palabras que lo componen, y se obtendrá el valor numérico N, número de estas palabras, que es una medida de la extensión de este texto.

Se añadirá a continuación cada una de estas palabras a un vocablo, y se obtendrá un segundo valor numérico, V, número de los vocablos que tienen, por lo menos, una ocurrencia en el texto.

Está claro que V es función de N, es decir que para un texto dado, V crece con N. Pero, ¿según qué ley? No se sabe todavía con precisión. Ciertamente, es evidente que V crece menos deprisa que N, puesto que cada palabra que representa un vocablo ya utilizado en el texto infiere una unidad de retraso a V con relación a N. Al principio del texto, V es igual a N, hasta la primera repetición de un vocablo cualquiera; (...) V va a desengancharse de N y la curva se inflexionará hacia la horizontal. A continuación, las repeticiones se multiplicarán, el retraso sobre N se acentuará, llegará a ser cada vez más raro encontrar un vocablo que no haya tenido todavía ocurrencia en lo que precede.

Pero V no cesa de crecer, pues ningún texto agota el léxico de su autor (...) El resultado es que nuestra curva tendrá partes planas cada vez más largas, pero que de vez en cuando se elevará en una unidad en V; (...) se obtendrá una línea que tiende a llegar a ser paralela al eje de abscisas, pero sin que jamás llegue a serlo completamente.” (Müller, 1968: 267-268)

El càlcul de la variació ( $V / N$  o *type / token*) resulta en una xifra sempre inferior a 1, que en la mesura que s'hi aproxima mostra un nivell de variació més alt i en la mesura que

---

<sup>242</sup> C. Müller (1968) anomena *riqueza léxica* aquesta relació. En el meu treball, he preferit parlar de variació i de repetició, perquè són conceptes més objectius i perquè la riquesa lèxica d'un text hauria de ser una noció més matisada que depengués també d'altres factors com la raresa, l'adequació, la precisió, l'elaboració, etc. dels mots que hi apareixen.

s'acosta a zero mostra un nivell de variació més baix<sup>243</sup>. El càlcul de la repetició, en canvi, resulta en una xifra sempre superior a 1, que en la mesura que se n'allunya mostra un nivell de repetició més alt i en la mesura que s'hi apropa el mostra més baix<sup>244</sup>.

El càlcul d'aquestes proporcions sempre s'ha de fer amb un nombre fix de mots, ja que, per exemple, en els dos-cents primers mots d'un text és més fàcil trobar-hi repeticions que en els cent primers. Dit d'una altra manera, un cop ha aparegut un mot en un text té moltes més probabilitats de reaparèixer-hi que abans de la primera aparició. Per tant, tal com assenyala D. Biber (1988), una comparació amb textos de longituds molt diferents falsejaria els resultats:

“This feature is computed by counting the number of different lexical items that occur in the first 400 words of each text. (...) The relation between the number of types and the total number of words in a text is not linear. That is, a large number of the different words used in the first 100 words of a text will be repeated in each successive 100-word chunk of text. The result is that each additional 100 words of text adds fewer and fewer additional types. In a comparison of very short texts and very long texts, the type / token ratio computed over the entire text will thus appear to be much higher in the short texts than in the long texts.” (Biber, 1988: 238-239)

---

<sup>243</sup> Perquè un text fos de valor 1, hauria de tenir el mateix nombre de *types* que de *tokens*, de manera que no contindria cap mot repetit. Per a l'opció contrària, que un text tendís a zero, hauria de tenir un nombre progressivament menor de *types* —el zero absolut és impossible, perquè el text més redundent concebible hauria de contenir almenys un mot, per més que aquest mot hi estigués repetit centenars o milers de vegades.

<sup>244</sup> A J.M. Castellà (1996c: 178-183) havia proposat un sistema de càlcul que no es basés en tots els mots del text sinó només en els mots lexicals i, fins i tot, només en els radicals, evitant de comptar com a mots diferents una paraula i el plural, el femení, el diminutiu, el derivat, la forma verbal, etc. corresponents. Malgrat la bona intenció que hi vaig posar, al final vaig arribar a uns resultats equiparables als que inclouen tots els mots del text i que s'obtenen d'una manera automàtica, molt menys laboriosa, amb el programa *TACT*. Ambdues raons, els resultats i el menor esforç requerit, m'han fet decidir de calcular aquest índex de la manera més senzilla. D'altra banda, l'estudi de la redundància d'una manera global en el discurs seria, de segur, una qüestió molt més complexa del que s'està plantejant aquí. La repetició pot fer-se amb un mateix element lèxic o bé amb un sinònim o amb un mot hipònim o hiperònim, i pot fer-se, fins i tot, amb tota una paràfrasi. La repetició pot ser volguda —i llavors és un recurs retòric— o no volguda —i ser resultat de les restriccions de temps en la construcció del discurs o de la poca habilitat de l'emissor —aquesta segona opció, especialment en el text escrit. A més, un cert grau de redundància informativa forma part de la naturalesa mateixa de l'estructura discursiva (articulació tema-remà). Aleshores, la repetició pot formar part de l'estructura necessària d'un discurs o pot ser-ne un element superflu, etc. L'estudi, interessantíssim, de qualsevol d'aquestes qüestions requeriria una atenció detallada que no puc prestar en aquest treball i que, d'altra banda, tampoc no és la línia habitual de les investigacions que he consultat.

En el cas present, per assegurar l'equiparació dels resultats he procurat que els fragments utilitzats tinguin un nombre molt similar de mots —tots estan inclosos en una banda estreta que oscil·la entre els 5241 mots de la classe de l'individu D i els 5557 mots de la conversa 1— i he fet el còmput sobre el total de mots del text<sup>245</sup>.

Pel que fa a la longitud de mot, l'he calculada com D. Biber (1988), que ho ha fet a partir del nombre de caràcters per mot i, per tant, des d'una perspectiva ortogràfica<sup>246</sup>.

### 4.9.3- RESULTATS I DISCUSSIÓ

Els índexs de variació i de repetició lèxica que he obtingut estan reflectits a la taula 67 (vg. pàgina següent). Aquesta taula mostra amb claredat una tendència a una variació lèxica més gran en la prosa acadèmica i, complementàriament, a una repetició lèxica superior en la conversa informal. Una vegada més, la classe ha ocupat una posició intermèdia, aquesta vegada més decantada cap a la conversa que cap a l'escrit. A l'índex de variació, que com he dit es mou en una franja entre 0 i 1, la classe ha quedat a 0.04 punts de la conversa i a 0.09 de l'escrit. L'índex de repetició lèxica, complementàriament, indica que els mots s'han repetit, de mitjana, 5.7 vegades en la conversa informal, 4.7 vegades en la classe magistral i 3.3 vegades en els escrits.

---

<sup>245</sup> Empro el conjunt del text perquè, encara que els resultats no seran comparables amb el treball de D. Biber (1988), una xifra més alta ha de donar més fiabilitat als resultats. D. Biber no té altra opció que fixar un nombre baix de mots, perquè, pel que sembla, disposa d'alguns textos molt breus en el corpus. D'altra banda, el treball de W. Chafe i J. Danielewicz (1987) que he esmentat a l'inici tampoc no permet la comparació, ja que, tot i que hi presenten uns resultats sorprenents —no perquè siguin extrems sinó, tot al contrari, perquè els diversos gèneres gairebé no difereixen entre si—, no aclareixen amb quin nombre de mots han realitzat el càlcul, ni si han equiparat els fragments de textos per fer-los contrastables.

<sup>246</sup> Per raons pràctiques, és preferible de fer aquest càlcul amb caràcters i no amb síl·labes, ja que *TACT* reconeix fàcilment els caràcters però no sap què és una síl·laba (i, probablement, seria bastant complicat explicar-li-ho). És possible, però, que en la percepció subjectiva de la longitud d'una paraula hi intervingui més el nombre de síl·labes que el de caràcters. Intuïtivament, sembla més llarg a l'oïda el mot *cafetera* que el mots *destrenar* o, fins i tot, *destrempar*, que tenen menys síl·labes, però més caràcters i més sons. D'altra banda, aquest possible valor de la síl·laba com a unitat de longitud dels mots podria prevaler també en el grau de dificultat en el processament. Una pista significativa a favor d'això és que en l'experiment de F. Ferreira i altres (1996: 328), en què es calcula el temps de processament dels mots en la llengua oral segons la freqüència, es té en compte com a variable de control el nombre de síl·labes i no el de caràcters o el de sons.

<b>TAULA 67</b>	<b>CONVERSA INFORMAL</b>	<b>CLASSE MAGISTRAL</b>	<b>PROSA ACADÈMICA</b>
<b>Variació lèxica (<i>type / token</i>)</b>	0.17	0.21	0.30
<b>Repetició lèxica (<i>token / type</i>) *</b>	5.7	4.7	3.3

(\* Nombre mitjà de repeticions de cada mot en el total de mots)

La prosa acadèmica ha demostrat, per tant, ser més variada en l'ús del vocabulari, com es correspon amb la condició de producció en diferit, amb temps suficient per a una tria lèxica elaborada. La classe, en aquest tret, sembla més condicionada per la producció *on-line* o per l'ús d'estratègies específiques de progressió de la informació, que l'associen amb la conversa informal, i no pel caràcter formal i especialitzat que en altres ocasions l'ha aparellada amb la llengua escrita.

L'altre indicador d'aquest apartat, la longitud de mot, ha presentat els resultats següents:

<b>TAULA 68 *</b>	<b>CONVERSA INFORMAL</b>	<b>CLASSE MAGISTRAL</b>	<b>PROSA ACADÈMICA</b>
<b>Longitud de mot</b>	3.6	4.3	4.9

(\* Nombre mitjà de caràcters per mot)

Tal com reflecteix la taula 68, els mots de la conversa informal han estat més curts, amb una mitjana de 3.6 caràcters per mot; els de la classe magistral han resultat una mica més llargs, de 4.3 caràcters per mot, pràcticament equidistants de la conversa i de la prosa; finalment, els mots de la prosa acadèmica han presentat una longitud superior, amb una mitjana de 4.9 caràcters per mot. Aquesta diferència, encara que no sigui espectacular,

constitueix una contribució a la densitat del text escrit. De fet, si es compara amb els resultats de D. Biber (1988) que he presentat anteriorment, en el seu treball hi havia uns contrastos encara més reduïts.

La distribució de les dades per textos és, en aquest apartat, d'una gran coherència. Presento a les pàgines següents (taules 69, 70, 71, 72, 73 i 74) les xifres i les bandes de repetició lèxica i les de longitud de mot, i ometo les de variació perquè són, de fet, complementàries de les de repetició i donen una distribució gairebé idèntica.

Com he dit, la distribució de resultats és altament coherent i conforma unes bandes, especialment en la repetició lèxica, perfectament delimitades. Així, l'índex de repetició de les converses oscil·la entre 5.8 i 5.5, en les classes ho fa entre 4.9 i 4.3, i en els escrits, entre 3.5 i 3.0. La dispersió és de 0.3 per a la conversa, de 0.6 per a la classe i de 0.5 per a la prosa acadèmica, amb una distància de 0.6 entre la conversa i la classe i de 0.8 entre la classe i la prosa.

La longitud de mot, d'altra banda, té un comportament similar: en les converses se situa entre 3.5 i 3.7, en les classes entre 4.1 i 4.5, i en la prosa entre 4.6 i 5.0. En aquest cas, la dispersió dins de cada gènere és molt reduïda (0.2, 0.4 i 0.4, respectivament). Però, en canvi, les distàncies entre gèneres són curtes, de 0.4 entre la conversa i la classe, i de només 0.1 entre la classe i la prosa, de manera que la classe amb l'índex més alt (4.5) s'aproxima molt a l'escrit amb l'índex més baix (4.6), sense que les bandes arribin, però, a superposar-se en cap moment<sup>247</sup>.

---

<sup>247</sup> En aquest apartat he optat per no oferir cap exemple de text perquè, encara que es puguin trobar fragments de conversa amb una gran repetició, també n'hi ha que presenten més variació, i el mateix ocorre amb les classes i amb els escrits. De fet, els índexs que he presentat responen a un càlcul sobre la globalitat del text i han donat lloc a diferències numèriques relativament petites que només es poden obtenir a partir de mitjanes fetes amb tots els mots del text. En conseqüència, la tria de qualsevol fragment podria ser, o bé arbitrària, o bé subjectiva, i podria produir una falsa impressió sobre la variació i la repetició lèxiques.

<b>TAULA 69 *</b>	<b>REPETICIÓ LÈXICA</b>	<b>CONVERSA</b>
<b>CONVERSA 1</b>	Repetició lèxica .....	5.7
<b>CONVERSA 2</b>	Repetició lèxica .....	5.8
<b>CONVERSA 3</b>	Repetició lèxica .....	5.5
<b>CONVERSA 4</b>	Repetició lèxica .....	5.5
<b>CONVERSA 5</b>	Repetició lèxica .....	5.8

(\* Nombre mitjà de repeticions de cada mot en el total de mots)

<b>TAULA 70 *</b>	<b>REPETICIÓ LÈXICA</b>	<b>CLASSE MAGISTRAL</b>	<b>PROSA ACADÈMICA</b>
<b>INDIVIDU A</b>	Repetició lèxica .....	4.9	3.2
<b>INDIVIDU B</b>	Repetició lèxica .....	4.7	3.4
<b>INDIVIDU C</b>	Repetició lèxica .....	4.3	3.0
<b>INDIVIDU D</b>	Repetició lèxica .....	4.9	3.3
<b>INDIVIDU E</b>	Repetició lèxica .....	4.9	3.5

(\* Nombre mitjà de repeticions de cada mot en el total de mots)



<b>TAULA 71 *</b>	<b>LONGITUD DE MOT</b>	<b>CONVERSA</b>
<b>CONVERSA 1</b>	Longitud de mot .....	3.7
<b>CONVERSA 2</b>	Longitud de mot .....	3.5
<b>CONVERSA 3</b>	Longitud de mot .....	3.6
<b>CONVERSA 4</b>	Longitud de mot .....	3.6
<b>CONVERSA 5</b>	Longitud de mot .....	3.5

(\* Nombre mitjà de caràcters per mot)

<b>TAULA 72 *</b>	<b>LONGITUD DE MOT</b>	<b>CLASSE MAGISTRAL</b>	<b>PROSA ACADÈMICA</b>
<b>INDIVIDU A</b>	Longitud de mot .....	4.5	5.0
<b>INDIVIDU B</b>	Longitud de mot .....	4.2	5.0
<b>INDIVIDU C</b>	Longitud de mot .....	4.2	4.8
<b>INDIVIDU D</b>	Longitud de mot .....	4.5	4.9
<b>INDIVIDU E</b>	Longitud de mot .....	4.1	4.6

(\* Nombre mitjà de caràcters per mot)

<b>TAULA 73 * - REPETICIÓ LÈXICA (Dispersió de les dades)</b>		
2.8		
3.0	Escrit C	
3.2	Escrit A	<b>BANDA</b>
3.4	Escrit D Escrit B	<b>DE LA</b>
3.6	Escrit E	<b>PROSA ACADÈMICA</b>
3.8		
4.0		
4.2		
4.4	Classe C	<b>BANDA</b>
4.6		<b>DE LA</b>
4.8	Classe B	<b>CLASSE MAGISTRAL</b>
5.0	Classe A Classe D Classe E	
5.2		
5.4		
5.6	Conv. 3 Conv. 4	<b>BANDA DE LA</b>
5.8	Conv. 1 Conv. 2 Conv. 5	<b>CONVERSA INFORMAL</b>
6.0		

(\* Nombre mitjà de repeticions de cada mot en el total de mots)

<b>TAULA 74 - LONGITUD DE MOT (Dispersió de les dades)</b>		
5.1		
5.0	Escrit A Escrit B	
4.9	Escrit D	<b>BANDA</b>
4.8	Escrit C	<b>DE LA</b>
4.7		<b>PROSA ACADÈMICA</b>
4.6	Escrit E	
4.5	Classe A Classe D	<b>BANDA</b>
4.4		<b>DE LA</b>
4.3		<b>CLASSE MAGISTRAL</b>
4.2	Classe B Classe C	
4.1	Classe E	
4.0		
4.1		
3.8		
3.7	Conversa 1	<b>BANDA</b>
3.6	Conversa 3 Conversa 4	<b>DE LA</b>
3.5	Conversa 2 Conversa 5	<b>CONVERSA INFORMAL</b>
3.4		

(\* Nombre mitjà de caràcters per mot)

#### **4.9.4- CONCLUSIONS DE L'APARTAT**

Els resultats, tant en variació i repetició lèxica com en longitud de mot, han mostrat la ja habitual progressió de la conversa informal a la classe magistral i d'aquesta a la prosa acadèmica. La conversa ha estat menys variada lèxicament i, per tant, ha tingut un nombre mitjà de repetició de mots més alt, alhora que la longitud de mot hi ha estat més baixa. La prosa acadèmica, per contra, ha estat més variada, amb menys repetició i amb la longitud de mot més alta. I la classe ha ocupat la posició intermèdia en tots tres índexs.

Les xifres han mostrat com la variació lèxica, la repetició lèxica i la longitud de mot contribueixen a la densitat informativa del text, conjuntament amb els trets estudiats en els apartats anteriors, i com presenten uns comportaments ben diferenciats i, alhora, ben idiosincràtics, a cadascun dels tres gèneres objecte d'aquesta recerca.

## **4.10- LA DENSITAT LÈXICA, L'ESTIL NOMINAL I L'ESTIL VERBAL: VISIÓ DE CONJUNT**

Aquest apartat consisteix en el resum i la visió global de molts dels fenòmens que he abordat en punts anteriors d'aquest capítol. En algunes qüestions, en particular, faré una presentació conjunta de dades que abans han aparegut separadament, disseminades al llarg de les anàlisis de cadascun dels aspectes tractats. Aquesta visió panoràmica facilitarà diverses comparacions i contrastos, i permetrà de perfilar algunes línies de comportament dels trets lingüístics estudiats que, per separat, quedaven amagades en el garbuix de les dades i en la distància física de les pàgines que dificulta de relacionar-los entre si.

Els aspectes globalitzadors que tractaré en aquest apartat es concreten en els set punts següents:

- 4.10.1- Densitat lèxica: dades i interpretació globals**
- 4.10.2- Estil nominal i estil verbal: dades i interpretació globals**
- 4.10.3- Comparació català-anglès**
- 4.10.4- Resum dels trets característics de cada gènere**
- 4.10.5- Posició relativa de la classe magistral**
- 4.10.6- Distribució dels resultats: bandes i correspondències**
- 4.10.7- Estil nominal i estil verbal: exemples de textos**
- 4.10.8- Conclusions de l'apartat**

#### **4.10.1- DENSITAT LÈXICA: DADES I INTERPRETACIÓ GLOBALS**

La densitat lèxica ha estat un tret pres en consideració per diversos autors. Com ja he assenyalat, més enllà de la proporció entre mots lexicals i mots gramaticals, hi ha factors com la variació lèxica, la longitud de mot i d'altres que també hi incideixen. A continuació els presento subdividits en dos grups, segons si contribueixen a la densitat lèxica o si la fan minvar:

##### **A/ Trets que afavoreixen la densitat lèxica:**

- 1- Mots lexicals.
- 2- Variació lèxica.
- 3- Longitud de mot.
- 4- Nominalitzacions.

##### **B/ Trets que disminueixen la densitat lèxica:**

- 1- Noms, verbs i adjectius de freqüència alta.
- 2- Mots gramaticals.
- 3- Repetició lèxica.

A l'apartat A encara podrien afegir-s'hi els noms, els adjectius atributius, els sintagmes preposicionals complements del nom i de l'adjectiu i, fins i tot, els verbs predicatius (en contrast amb els copulatius). En el cas dels tres primers, en la mesura que conformen uns sintagmes nominals més llargs, són uns bons contribuïdors a la densitat informativa del text. En el quart, els verbs predicatius aporten més contingut que els copulatius, mers lligams, gairebé només sintàctics, entre el subjecte i el predicat. De totes maneres, com que aquests trets ja apareixen en el punt següent com a elements constitutius de l'estil nominal i de l'estil verbal, seria repetitiu d'incloure'ls també aquí. La distribució dels trets dels apartats A i B, es pot veure a la taula següent:

<b>TAULA 75 *</b>	<b>CONVERSA INFORMAL</b>	<b>CLASSE MAGISTRAL</b>	<b>PROSA ACADÈMICA</b>
<b>+ DENSITAT LÈXICA</b>			
<b>Mots lexicals</b>	- (34.1 %)	+/- (40.8 %)	+ (45.0 %)
<b>Variació lèxica</b>	- (0.17)	+/- (0.21)	+ (0.30)
<b>Longitud de mot</b>	- (3.6)	+/- (4.3)	+ (4.9)
<b>Nominalitzacions</b>	- (1.1 %)	+/- (10.1 %)	+ (17.7 %)
<b>- DENSITAT LÈXICA</b>			
<b>Noms, verbs i adjs. de freqüència alta</b>	+ (div.)	+/- (div.)	- (div.)
<b>Mots gramaticals</b>	+ (65.9 %)	+/- (59.2 %)	- (55.0 %)
<b>Repetició lèxica</b>	+ (5.7)	+/- (4.7)	- (3.3)

(\* Percentatges i índexs diversos, provinents de les taules 8, 18, 19, 30, 38, 61, 67 i 68)

Com es pot observar, la distribució dels signes + i - reflecteix absolutament les tendències dominants a totes tres columnes<sup>248</sup>. La conversa informal es caracteritza, com

<sup>248</sup> Sembla difícil d'obtenir un índex numèric únic que acumulés unes xifres tan dispars com les que han aparegut entre parèntesis a la taula 75. He fet alguns intents en aquesta direcció, que no crec que puguin passar de mers jocs aritmètics. Potser l'únic que podria resistir una mínima anàlisi amb sentit comú seria el que seleccionés els tres components principals de la densitat i n'obtingués el producte, és a dir: (percentatge de mots lexicals) x (índex de variació lèxica) x (longitud mitjana de mot). Aquesta combinació afegiria al tret bàsic del nombre de mots lexicals les característiques de ser més o menys variats i més o menys llargs i, per tant, amb un grau major o menor de dificultat de processament. Els resultats d'aquest producte són els que ofereixo a continuació:

. **Conversa informal:** 34.1 % (mots lexicals) x 0.17 (índex de variació lèxica) x 3.6 (longitud mitjana de mot) = 20.8

. **Classe magistral:** 40.8 % (mots lexicals) x 0.21 (índex de variació lèxica) x 4.3 (longitud mitjana de mot) = 36.8

. **Prosa acadèmica:** 45.0 % (mots lexicals) x 0.30 (índex de variació lèxica) x 4.9 (longitud mitjana de mot) = 66.2

correspon a un gènere oral prototípic, per la presència de signes negatius en els trets que afavoreixen la densitat i de signes positius en els que la fan minvar, la prosa acadèmica adopta la distribució totalment inversa i la classe magistral se situa en la posició central. Aquestes dades, doncs, conjuntament amb les de l'estil nominal i l'estil verbal del punt següent, donen un perfil ben definit dels tres gèneres en els trets de caràcter lexical.

#### **4.10.2- ESTIL NOMINAL I ESTIL VERBAL: DADES I INTERPRETACIÓ GLOBALS**

Els resultats dels apartats anteriors han fet evident, com ja he comentat àmpliament, que la llengua escrita es basa més en un estil nominal i la llengua oral en un estil verbal de construcció dels enunciats. De totes maneres, per obtenir una xifra que concreti numèricament aquesta afirmació, cal unificar els índexs obtinguts en diversos apartats. Per això, he seleccionat i he sumat les dades de les categories que contribueixen inequívocament a cadascun dels estils i que constitueixen, per tant, una mesura adequada del pes que tenen en els textos. Com es pot veure a les taules 76 A i 76 B, he relacionat els noms amb els complements corresponents, és a dir, amb els adjectius atributius i amb els sintagmes preposicionals complements del nom i de l'adjectiu, i he relacionat els verbs amb els adjectius predicatius i amb els sintagmes preposicionals complements del verb.

La taula 76 A mostra una evolució clarament ascendent dels trets de l'estil nominal des de la conversa informal (16.7 %) fins a la prosa acadèmica (44.5 %), amb una diferència entre ambdós índexs de 27.8 punts. La mateixa progressió, però en sentit contrari, és la que s'observa a la taula 76 B sobre l'estil verbal, amb la particularitat que aquí la corba evolutiva és més suau, parteix i arriba a uns índexs més baixos i presenta un recorregut total també més moderat (24.9 % per a la conversa i 15.1 % per a la prosa, amb una diferència de només 9.8 punts). En la conversa, l'índex de l'estil nominal queda a prop del de l'estil verbal (estil nominal 16.7 %, estil verbal 24.9 %), mentre que en la prosa queden molt allunyats (estil nominal 44.5 %, estil verbal 15.1 %).

<b>TAULA 76 A *</b>	<b>CONVERSA INFORMAL</b> (27254 mots)	<b>CLASSE MAGISTRAL</b> (26374 mots)	<b>PROSA ACADÈMICA</b> (26468 mots)
<b>Noms</b>	13.4 %	19.9 %	24.5 %
<b>Adjectius atributius</b>	0.8 %	4.7 %	9.2 %
<b>Sint. preposicional compl. nom i adjectiu</b>	2.5 %	7.3 %	10.8 %
<b>Total estil nominal</b>	16.7 % (4551)	31.9 % (8413)	44.5 % (11778)

(\* Percentatges sobre el total de mots i, entre parèntesis, dades absolutes d'ocurrències – Test  $\chi^2$ : probabilitat d'atzar inferior al 0.0001 % - Dades provinents de les taules 5, 29 i 48)

<b>TAULA 76 B *</b>	<b>CONVERSA INFORMAL</b> (27254 mots)	<b>CLASSE MAGISTRAL</b> (26374 mots)	<b>PROSA ACADÈMICA</b> (26468 mots)
<b>Verbs</b>	17.5 %	13.9 %	9.4 %
<b>Adjectius predicatius</b>	2.2 %	1.1 %	0.9 %
<b>Sint. preposicional complements del verb</b>	5.2 %	5.5 %	4.8 %
<b>Total estil verbal</b>	24.9 % (6786)	20.5 % (5407)	15.1 % (3997)

(\* Percentatges sobre el total de mots i, entre parèntesis, dades absolutes d'ocurrències – Test  $\chi^2$ : probabilitat d'atzar inferior al 0.0001 % - Dades provinents de les taules 16, 29 i 48)

Es comprova, doncs, que la conversa informal té un estil preferentment verbal, si bé d'una manera més moderada, i que la prosa acadèmica té un estil accentuadament nominal.



L'estil nominal i l'estil verbal (conformats pels noms, pels verbs, pels adjectius atributius i predicatius i pels sintagmes preposicionals en funció de complement del nom, de l'adjectiu i del verb), al costat de la densitat lèxica (conformada per la proporció de mots lexicals i mots gramaticals, pels mots de freqüència alta, per la variació i la repetició lèxiques, per la longitud de mot i per les nominalitzacions<sup>249</sup>) han resultat ser elements caracteritzadors bàsics del contrast entre els tres gèneres estudiats.

Així, les principals afirmacions de J. Ure i K. Perera sobre la densitat lèxica, de W. Chafe sobre la integració de la informació i de M.A.K. Halliday sobre els estils nominal i verbal, ampliades i matisades per altres autors com D. Biber i D.P. Hayes, han resultat àmpliament confirmades en la meva mostra en llengua catalana.

#### **4.10.3- COMPARACIÓ CATALÀ-ANGLÈS**

Al llarg dels apartats anteriors he anat aportant els resultats de recerques efectuades sobre la llengua anglesa, principalment de l'obra de D. Biber (1988) que, com he dit, és la més general sobre aquest tema. En moltes ocasions, aquests resultats no són comparables amb els meus, perquè han estat calculats en base a criteris diferents o perquè les unitats de l'anglès i del català divergeixen a causa de l'estructura pròpia de cada llengua. En altres casos, però, la comparació és més possible. Per això a continuació ofereixo, en el mateix ordre en què han aparegut, els resultats de les dues llengües, i comento les semblances i les dissemblances que presenten.

##### **A/ Els noms**

Les dades de D. Biber (1988) i les meves pel que fa als noms són les que consten a la taula següent:

---

<sup>249</sup> Les nominalitzacions incideixen en la densitat lèxica al mateix temps que afavoreixen un estil nominal i augmenten la longitud de mot. No les he comptat numèricament dins de l'índex de l'estil nominal perquè representarien una repetició respecte de la xifra general de noms, dins de la qual ja estan incloses, i en part respecte de l'índex de longitud de mot.

TAULA 77 *	A	B	C
Noms - Biber (1988)	13.7 %	18.9 %	18.8 %
Noms - Castellà (2001)	13.4 %	19.9 %	24.5 %

(\* Percentatges sobre el total de mots - A: *Face-to-face conversation* / Conversa informal, B: *Prepared Speeches* / Classe magistral, C: *Academic prose* / Prosa acadèmica - Dades provinents de les taules 4 i 5)

Per comparar aquestes dades cal tenir en compte que D. Biber no ha inclòs les nominalitzacions en el seu percentatge de noms. Al marge d'això, els resultats han estat molt similars en la conversa informal i en la classe magistral, però han estat bastant diferents en la prosa acadèmica. De fet, a la vista de les dades, necessita molta més explicació el fet que en anglès hi hagi hagut el mateix percentatge a *Prepared speeches* i a *Academic prose* que no pas la meva progressió, que es mostra molt més coherent. D'altra banda, si s'hi sumessin els seus resultats de nominalitzacions (vg. taula 57, § 4.8.1), tampoc no s'explicaria el perquè d'aquesta igualtat. Cal recordar, a més, que el percentatge que obtenien W.N. Francis i H. Kuèera (1982) per al conjunt de la llengua escrita en anglès era del 26.8 %, xifra que difereix en gran manera de la de D. Biber i, en canvi, s'acosta molt a la meva.

### **B/ Els pronoms personals**

Els meus resultats de la taula 12 (vg. § 4.2.3) no són comparables amb les de la taula 14 (vg. § íd.) de D. Biber, ja que aquest hi ha inclòs tots els pronoms personals, forts i febles, i nosaltres només els forts. Atesa la diferent funció (com a dítctics els pronoms forts, com a elements anafòrics els pronoms febles) d'aquests dos grups de mots, els resultats de cada taula responen a criteris heterogenis. L'única coincidència que es pot destacar és que, a totes dues llengües, s'han trobat més pronoms personals en la conversa informal que en els altres dos gèneres.

## C/ Els verbs

Els resultats dels verbs en la meva recerca i en la de D. Biber (1988) són els que s'indiquen a la taula 78. Tot i que en el còmput de D. Biber hi falti el temps de futur (vg. § 4.3.1), la baixa freqüència relativa d'aquest temps a la majoria de textos fa bastant comparables les dades que, d'altra banda, han estat tan pròximes que fins i tot han coincidit en un dels índexs. Per a les altres subdivisions, dins de la categoria verbal, que he efectuat en l'apartat corresponent (verbs de freqüència alta, verbs copulatiu, etc.), la comparació esdevé impossible, ja que D. Biber ha classificat les seves dades d'acord amb altres criteris<sup>250</sup>.

TAULA 78 *	A	B	C
Verbs - Biber (1988)	18.4 %	13.9 %	10.7 %
Verbs - Castellà (2001)	17.5 %	13.9 %	9.4 %

(\* Percentatges sobre el total de mots - A: *Face-to-face conversation* / Conversa informal, B: *Prepared Speeches* / Classe magistral, C: *Academic prose* / Prosa acadèmica - Dades provinents de les taules 15 i 16)

## D/ Els adjectius

En aquest apartat disposem dels resultats de W. Chafe, de D. Biber i de L. W. Smith per a l'anglès i dels meus per al català. Els he distribuït en dues taules, una per als adjectius atributius i una altra per als predicatius<sup>251</sup>:

---

<sup>250</sup> Aquest és també el problema, com ja he dit (vg. § 4.3.1), del percentatge de verbs de W.N. Francis i H. Kuèera (1982), que inclouen en el càlcul els verbs auxiliars i els modals i, per tant, els seus resultats no són comparables ni amb els de D. Biber ni amb els meus.

<sup>251</sup> L.W. Smith (1994) ofereix resultats per als adjectius atributius però, curiosament, no per als predicatius, ni tampoc no dóna una xifra global d'adjectius a partir de la qual es pogués induir el nombre de cada tipus. Només l'he inclòs, per tant, en la primera de les taules.

<b>TAULA 79 *</b>	<b>A</b>	<b>B</b>	<b>C</b>
<b>Adjectius atributius - Chafe (1982)</b>	3.4 %	–	13.5 %
<b>Adjectius atributius - Biber (1988)</b>	4.1 %	4.9 %	7.7 %
<b>Adjectius atributius - Smith (1994)</b>	1.8 %	2.2 %	6.9 %
<b>Adjectius atributius - Castellà (2001)</b>	0.8 %	4.7 %	9.2 %

(\* Percentatges sobre el total de mots - A: *Spoken / Face-to-face conversation / Informal conversation / Conversa informal*, B: *Prepared Speeches / Formal interview / Classe magistral*, C: *Written / Academic prose / Written essay / Prosa acadèmica* - Dades provinents de les taules 28 i 29)

<b>TAULA 80 *</b>	<b>A</b>	<b>B</b>	<b>C</b>
<b>Adjectius predicatius - Chafe (1982)</b>	1.6 %	–	2.0 %
<b>Adjectius predicatius - Biber (1988)</b>	0.4 %	0.4 %	0.5 %
<b>Adjectius predicatius - Castellà (2001)</b>	2.2 %	1.1 %	0.9 %

(\* Percentatges sobre el total de mots - A: *Spoken / Face-to-face conversation / Conversa informal*, B: *Prepared Speeches / Classe magistral*, C: *Written / Academic prose / Prosa acadèmica* - Dades provinents de les taules 28 i 29)

En aquest punt cal destacar, en primer lloc, la divergència entre els resultats de W. Chafe (1982), els de D. Biber (1988) i els de L.W. Smith (1994), que pot explicar-se, almenys,

perquè efectuen la seva recerca sobre gèneres diferents<sup>252</sup>. En segon lloc, es pot dir que les meves dades presenten, tant en els adjectius atributius com en els predicatius, unes tendències més definides i uns contrastos entre gèneres més accentuats. Per exemple, els adjectius atributius de l'estudi de D. Biber varien poc de *Face-to-face conversation* (4.1 %) a *Prepared speeches* (4.9 %), mentre que els meus mostren una diferència notable (0.8 % i 4.7 %, respectivament), i els adjectius predicatius de D. Biber gairebé no canvien en els tres gèneres (0.4 %, 0.4 % i 0.5 %), mentre que en el meu treball presenten una progressió decreixent clara i contrastada (2.2 %, 1.1 % i 0.9 %).

### **E/ Els adverbis**

En el cas dels adverbis no hi ha cap possibilitat de comparació, ja que els criteris de recerca de D. Biber i els meus han estat radicalment diferents: en el seu treball es distingeix entre diversos tipus d'adverbis sense que hi hagi cap apartat específic per als adverbis en *-ment*, que apareixen distribuïts en diverses agrupacions, mentre que en el meu m'he limitat precisament als adverbis en *-ment* i he prescindit de qualsevol altra forma.

### **F/ Les preposicions**

Els resultats pel que fa a les preposicions són només parcialment equiparables, perquè hi he comptat únicament les cinc preposicions febles del català, D. Biber hi ha inclòs les febles i les fortes de l'anglès, fins a un total de 41 preposicions i L. W. Smith n'ha pres en consideració un nombre no especificat<sup>253</sup>. Malgrat això, es pot fer un acarament de les dades per extreure'n les tendències generals:

---

<sup>252</sup> La discrepància encara augmenta més si es té en compte les dades de M.E. Poole i T.W. Field (1976) (vg. § 4.4.1), que no he incorporat aquí perquè aquests autors no distingeixen entre adjectius atributius i predicatius.

<sup>253</sup> L.W. Smith (1994) no concreta les preposicions que han format part del càlcul; per això podríem suposar que han estat diverses i de tota mena. De fet, la hipòtesi que n'hagi inclòs menys que D. Biber seria la més productiva, ja que explicaria les diferències considerables en els resultats entre ambdós investigadors.

TAULA 81 *	A	B	C
<b>Preposicions - Biber (1988)</b>	8.5 %	11.3 %	13.9 %
<b>Preposicions - Smith (1994)</b>	5.8 %	6.8 %	9.1 %
<b>Preposicions - Castellà (2001)</b>	8.8 %	14.7 %	17.0 %

(\* Percentatges sobre el total de mots - A: *Face-to-face conversation / Informal conversation / Conversa informal*, B: *Prepared Speeches / Formal interview / Classe magistral*, C: *Academic prose / Written essay / Prosa acadèmica* - Dades provinents de les taules 45 i 46)

En totes tres recerques s'ha observat una tendència creixent en l'ús de les preposicions des de la llengua oral cap a l'escrita, si bé en el meu cas apareixen unes xifres progressivament més altes. Cal destacar, a més, que aquesta diferència es produeix sense que hi hagi comptat les preposicions fortes i això sembla indicar una propensió especial del català, respecte de l'anglès, a l'ús de les preposicions *i*, per tant, a l'estructuració de la informació a través de sintagmes preposicionals.

### **G/ Les nominalitzacions**

Aquest és el tret lingüístic per al qual he disposat de més fonts, ja que hi ha, a més de les meves, dades procedents dels cinc treballs que especifico a la taula 82 (vg. pàgina següent). El que sorprèn en la nominalització, com ja havia remarcat en l'apartat corresponent, és la disparitat dels resultats sobre l'anglès entre els diversos autors. Per exemple, hi ha una gran distància entre el 5.6 % de l'escrit a W. Chafe (1982) i l'1.0 % de G. Redeker (1984). També divergeixen entre si les dades de W. Chafe (1982), les de W. Chafe i J. Danielewicz (1987) —que, curiosament, coincideixen prou amb les de L.W. Smith (1994)— i les de D. Biber (1988). Al final, les més pròximes a les meves són les d'aquest autor, potser pel fet que analitza uns gèneres més propers als del present estudi. Com ja comença a ser habitual, però, les meves dades són les més contrastades.

TAULA 82 *	A	B	C
<b>Nominalitzacions - Chafe (1982)</b>	0.5 %	-	5.6 %
<b>Nominalitzacions - Redeker (1984)</b>	0.3 %	0.5 %	1.0 %
<b>Nomin. – Chafe i Danielewicz (1987)</b>	2.7 %	5.6%	9.2 %
<b>Nominalitzacions - Biber (1988)</b>	0.9 %	2.1 %	3.6 %
<b>Nominalitzacions - Smith (1994)</b>	3.5 %	5.0 %	9.6 %
<b>Nominalitzacions - Castellà (2001)</b>	0.1 %	2.0 %	4.3 %

(\* Percentatges sobre el total de mots - A: *Dinnertable conversations / Spoken story / Conversations / Face-to-face conversation / Informal conversation / Conversa informal*, B: *Spoken explanation / Lectures / Prepared Speeches / Formal interview / Classe magistral*, C: *Academic papers / Written explanation / Academic papers / Academic prose / Written essay / Prosa acadèmica* - Dades provinents de les taules 57, 58, 59 i 60)<sup>254</sup>

## H/ La variació i la repetició lèxica

En aquest punt no hi ha comparació possible, ja que l'únic autor que ha pres en consideració aquests índexs (Biber, 1988) fa el càlcul sobre els primers 400 mots del text i aquí, en canvi, l'he fet sobre el total de mots, perquè he cregut que és una mesura més adequada. Com ja indicava en el seu moment, per a qualsevol maneig d'aquests índexs és imprescindible el treball amb un nombre de mots idèntic o molt similar.

## I/ La longitud de mot

A la taula següent hi apareixen les dades meves i les de D. Biber (1988) sobre la longitud mitjana dels mots del text:

<sup>254</sup> Per tal de poder encabir tots els autors que he citat en el punt 4.8.1 en aquest quadre de tres columnes he hagut de prescindir d'una dada en dos dels treballs que presentaven els resultats en quatre columnes. Les caselles sacrificades han estat la del gènere *written story* de G. Redeker (1984) i la del gènere *letters* de W. Chafe i J. Danielewicz (1987).

TAULA 83 *	A	B	C
<b>Longitud de mot - Biber (1988)</b>	4.1	4.4	4.8
<b>Longitud de mot - Castellà (2001)</b>	3.6	4.3	4.9

(\* Percentatges sobre el total de mots - A: *Face-to-face conversation* / Conversa informal, B: *Prepared Speeches* / Classe magistral, C: *Academic prose* / Prosa acadèmica - Dades provinents de les taules 66 i 68)

Com es pot veure a la taula 83, les meves dades són similars a les de D. Biber i, com en altres casos, una mica més contrastades.

Una mirada global a les taules comparatives que acabo de presentar per a cadascun dels trets permet de fer les observacions següents:

1/ Quan en anglès es disposa de diversos estudis, **les dades presenten una notable disparitat**, cosa que pot ser deguda a diferències en el corpus (gèneres i tipus de text seleccionats) o en la classificació i la codificació dels trets lingüístics estudiats.

2/ Les meves dades són en la majoria de casos similars a les de l'anglès, però **amb una progressió més clara i amb un contrast més accentuat entre els gèneres**<sup>255</sup>.

Les diferències en els resultats entre la llengua catalana i l'anglesa que acabo d'exposar podrien explicar-se per problemes accidentals d'equiparació dels criteris aplicats en cada recerca o bé per aspectes substantius de la naturalesa de cada llengua<sup>256</sup>. En qualsevol cas, es constaten unes línies i unes tendències ben definides per al conjunt dels resultats de la present recerca.

---

<sup>255</sup> Una diferència important entre els resultats de D. Biber (1988) i els meus és en la distribució i en les bandes de dispersió de les dades, fet que analitzaré en el punt 4.10.6.

<sup>256</sup> Una anàlisi aprofundida del perquè d'aquestes diferències escapa a l'abast d'aquesta tesi i hauria de ser abordada en un treball més específic sobre aquesta qüestió, amb la participació d'investigadors d'ambdues llengües. Aquí m'he limitat a comparar les dades que he trobat en una recerca que s'enfronta a altres objectius.



#### 4.10.4- RESUM DELS TRETS CARACTERÍSTICS DE CADA GÈNERE

Els resultats obtinguts en els apartats anteriors es poden agrupar i així es pot obtenir una caracterització detallada de cadascun dels gèneres pel que fa als principals trets estudiats en aquest capítol. Ofereixo en primer lloc, a la taula 84 (vg. pàgina següent), un contrast entre la conversa informal i la prosa acadèmica, perquè s'han mostrat com els gèneres més extrems o polaritzats en el present treball i perquè tenen un caràcter representatiu dins del conjunt de la llengua oral i de la llengua escrita<sup>257</sup>. I, a continuació, a la taula 85, dono una llista de la posició relativa de tots tres gèneres que fa patent d'una manera especial la naturalesa intermèdia de la classe magistral.

D'acord amb el resum contrastiu de la taula 84, hi ha dotze trets que es donen en major mesura en la conversa informal i tretze en la prosa acadèmica. Tots plegats conformen el conjunt dels 25 trets que han caracteritzat aquests gèneres al llarg d'aquest capítol. La taula 85 mostra les mateixes tendències però en tres columnes, amb dos gèneres en posició extrema i un en posició intermèdia. Segons aquesta taula, la conversa informal té el signe + en 11 ocasions, el signe – en 13 i la combinació +/- en 1. La prosa acadèmica té + en 11 casos, – en 10 i +/- en 2. I la classe magistral ha obtingut dues vegades el signe +, una vegada el signe – i 22 vegades els dos signes +/-.

Es confirma, per tant, de manera concloent el caràcter intermedi de la classe magistral enmig dels altres dos gèneres. Així, en la mesura que la conversa informal i la prosa acadèmica siguin representatives de la llengua oral i de la llengua escrita, es pot considerar que la classe magistral s'allunya del que és prototípic del mode oral, adquireix una posició perifèrica dins d'aquest mode i se situa en la zona de transició cap a la llengua escrita.

---

<sup>257</sup> Ja he comentat en el capítol de metodologia (vg. § 3.2.5) el caràcter clarament prototípic de la conversa i la representativitat menys definida de la prosa acadèmica. Pel que fa als trets lingüístics que he estudiat, caldria una recerca de gran abast, amb la consideració d'un nombre important de gèneres, per confirmar que aquests dos gèneres tenen efectivament el caràcter representatiu que aquí els he atorgat. Per tant, les llistes que presento a les taules 84 i 85 s'ha de considerar només indicatives de les tendències de la llengua oral i escrita.

**TAULA 84**

<p><b>CONVERSA INFORMAL</b> (Llengua oral)</p>	<p><b>PROSA ACADÈMICA</b> (Llengua escrita)</p>
<p>Més noms de freqüència alta  Més pronomes personals forts  Més demostratiu neutre <i>això</i>  Més verbs  Més verbs de freqüència alta  Més verbs copulatius  Més adjectius de freq. alta  Més adjectius predicatius    Més mots gramaticals    Més sint. prep. compl. del SV    Més repetició lèxica    Més estil verbal</p>	<p>Més noms        Més mots per clàusula  Més adjectius  Més adjectius atributius  Més adverbis en <i>-ment</i>  Més mots lexicals  Més preposicions febles  Més sint. prep. compl. del SN  Més sint. prep. per clàusula  Més nominalitzacions  Més variació lèxica  Més longitud de mot  Més estil nominal</p>

TAULA 85	CONVERSA	CLASSE	PROSA
1. Noms .....	-	+ / -	+
2. Noms de freqüència alta ...	+	+ / -	-
3. Pronoms personals forts ...	+	+ / -	-
4. Demostratiu neutre <i>això</i> ....	+	+ / -	-
5. Verbs .....	+	+ / -	-
6. Verbs de freqüència alta ...	+	+ / -	-
7. Verbs copulatius .....	+ / -	+	-
8. Mots per clàusula .....	-	+ / -	+
9. Adjectius .....	-	+ / -	+
10. Adjectius de freq. alta .....	+	-	+ / -
11. Adjectius atributius .....	-	+ / -	+
12. Adjectius predicatius .....	+	+ / -	-
13. Adverbis en <i>-ment</i> .....	-	+	+ / -
14. Mots lexicals .....	-	+ / -	+
15. Mots gramaticals .....	+	+ / -	-
16. Preposicions febles .....	-	+ / -	+
17. Sint. prep. compl. del SN ...	-	+ / -	+
18. Sint. prep. compl. del SV ...	+	+ / -	-
19. Sint. prep. per clàusula .....	-	+ / -	+
20. Nominalitzacions .....	-	+ / -	+
21. Variació lèxica .....	-	+ / -	+
22. Repetició lèxica .....	+	+ / -	-
23. Longitud de mot .....	-	+ / -	+
24. Estil nominal .....	-	+ / -	+
25. Estil verbal .....	+	+ / -	-

#### 4.10.5- POSICIÓ RELATIVA DE LA CLASSE MAGISTRAL

Acabo de concloure que la classe magistral ocupa un lloc inequívocament intermedi entre la conversa informal i la prosa acadèmica, però cal també demanar-se si aquest lloc és ben central o bé si es decanta cap a algun dels altres dos gèneres. Per a això he pensat un índex que reflecteixi la inclinació que mostra la classe en els principals trets analitzats. Aquest índex s'expressa per una xifra entre 0 i 1 que si és menor de 0.5 indica que la classe s'ha aproximat més a la conversa i si, per contra, és superior a 0.5, significa que s'ha inclinat més cap a la prosa acadèmica. A la taula de la pàgina següent podem veure aquest índex<sup>258</sup>.

Les dades de la taula 86 són bastant ajustades i deixen la classe magistral en una zona central dins de la posició intermèdia que li és pròpia, si bé lleugerament decantada cap a la prosa acadèmica. En 11 ocasions la classe s'ha acostat més a la conversa informal i en 14 a la prosa. L'índex global obtingut ha estat el 0.54, que s'inclina cap a la prosa acadèmica en la mesura mínima del 0.04 sobre 1. La majoria dels índexs es mouen en una franja central, decantada també cap a la prosa, per sobre de 0.31 i per sota de 0.85, excepte dos que han resultat molt excèntrics: els pronoms personals forts, en què amb un 0.96 la classe pràcticament s'ha igualat amb la prosa acadèmica, i els verbs de freqüència alta, en què la classe pràcticament s'ha igualat amb la conversa informal. Els 11 trets en què la classe magistral ha estat més a prop de la conversa i en els quals ha prevalgut més un estil verbal lligat a la implicació amb la situació i a la producció *online*, com correspon a un discurs oral, han estat els següents:

---

<sup>258</sup> L'índex que comento es calcula dividint la diferència entre el percentatge més alt dels tres gèneres i el percentatge intermedi per la diferència entre el percentatge més alt i el més baix (en la majoria de casos això consisteix a dividir la diferència entre el percentatge de la conversa i el de la classe per la diferència entre el percentatge de la conversa i el de la prosa). Per a aquest càlcul he utilitzat els percentatges obtinguts amb el denominador que en cada cas he argumentat com el més adequat (p. ex., per als noms de freqüència alta, utilitzo directament els percentatges sobre el total de noms i no sobre el total de mots). Pel que fa a la mitjana total, des del punt de vista estadístic podria discutir-se'n la validesa, ja que la mitjana aritmètica dóna el mateix pes a trets d'importància diversa. De totes maneres, com que no sabria determinar el valor numèric del pes relatiu de cada tret en el posicionament de la classe magistral respecte dels altres dos gèneres, he preferit de fer-ho igualitàriament, encara que tingui un caràcter merament indicatiu.

<b>TAULA 86</b>	<b>ÍNDEX</b>	<b>GÈNERE MÉS PROPER</b>
1. Noms .....	0.59	Prosa
2. Noms de freqüència alta ..	0.55	Prosa
3. Pronoms personals forts ..	0.96	Prosa
4. Demostratiu neutre <i>això</i> ...	0.40	Conversa
5. Verbs .....	0.44	Conversa
6. Verbs de freqüència alta ..	0.06	Conversa
7. Verbs copulatius .....	0.32	Conversa
8. Mots per clàusula .....	0.32	Conversa
9. Adjectius .....	0.39	Conversa
10. Adjectius de freq. alta ....	0.98	Prosa
11. Adjectius atributius .....	0.46	Conversa
12. Adjectius predicatius .....	0.85	Prosa
13. Adverbis en <i>-ment</i> .....	0.83	Prosa
14. Mots lexicals .....	0.57	Prosa
15. Mots gramaticals .....	0.61	Prosa
16. Preposicions febles .....	0.72	Prosa
17. Sint. prep. compl. SN .....	0.67	Prosa
18. Sint. prep. compl. SV .....	0.67	Prosa
19. Sint. prep. per clàusula ....	0.43	Conversa
20. Nominalitzacions .....	0.44	Conversa
21. Variació lèxica .....	0.31	Conversa
22. Repetició lèxica .....	0.42	Conversa
23. Longitud de mot .....	0.54	Prosa
24. Estil nominal .....	0.55	Prosa
25. Estil verbal .....	0.53	Prosa
<b>MITJANA .....</b>	<b>0.54</b>	<b>PROSA</b>

- Demonstratiu neutre *això*
- Verbs
- Verbs de freqüència alta
- Verbs copulatius
- Mots per clàusula
- Adjectius
- Adjectius atributius
- Sintagmes preposicionals per clàusula
- Nominalitzacions
- Variació lèxica
- Repetició lèxica

En canvi, els 14 trets en què ha prevalgut un estil nominal motivat per l'especificitat en el tema i per la formalitat de la situació, i que han aparellat la classe magistral amb la prosa acadèmica malgrat que es tracti de llengua oral, han estat els següents:

- Noms
- Noms de freqüència alta
- Pronoms forts
- Adjectius de freqüència alta
- Adjectius predicatius
- Adverbis en *-ment*
- Mots lexicals
- Mots gramaticals
- Preposicions
- Sintagmes preposicionals complement del sintagma nominal
- Sintagmes preposicionals complement del sintagma verbal
- Longitud de mot
- Estil nominal
- Estil verbal

Com es pot observar en aquestes llistes, hi ha una certa agrupació dels trets (p. ex. en els noms la classe s'ha decantat cap a la prosa i en els verbs cap a la conversa) que no es dóna en tots els casos (p. ex. en els adjectius, en les preposicions i en els trets de densitat lèxica s'ha decantat ara cap a la prosa ara cap a la conversa sense una direcció determinada).

#### 4.10.6- DISTRIBUCIÓ DELS RESULTATS: BANDES I CORRESPONDÈNCIES

Les taules que he anat oferint en els apartats anteriors amb les bandes de dispersió dels resultats de cada text permeten de fer-ne una agrupació per cercar-hi afinitats i correspondències. La taula-resum de totes les bandes és la que presento a continuació:

TAULA 87	Noms	Verbs	Adjectius	Adverbis	Mots lexic.	Preposicions	Nominalitz.	Repetició	Long. mot
<b>Banda de la prosa</b>	A B E	B	D	A	A E	B D	D	C	A B
		D E	B E	B		A	A B	A	D
	C D	A C	A C	C D	B C D	C	C	D B	C
<b>Banda de la classe</b>	D	D	A D	D	D	B D	A B	C	A D
	B	B		A			D		
	C	A	B C	C	A C	E	C	B	B C
	A E	C E	E	E	B E	A C	E	A D E	E
<b>Banda de la conv.</b>	5	15	14	15	145	5	234	34	1
	1					4			
	4	23	35	34	2	12			34
	2								
	3	4	2	2	3	3	15	125	25

La primera impressió que provoca la taula 87 és d'heterogeneïtat i de garbuix. Una observació detallada, però, permet de veure-hi unes tendències, dibuixades amb més o menys claredat, que esbosso a continuació:

– Els escrits dels individus A i B se situen generalment a la zona alta de la banda, són els més allunyats dels altres dos gèneres i, per tant, els més representatius de les tendències de la prosa acadèmica.

– Les classes dels individus B i C tendeixen a ser les més centrals a la banda de la classe magistral i, per tant, les més representatives del gènere.

– Les converses 2 i 3 se situen en general a la zona més baixa de la banda de la conversa informal i són, doncs, les més representatives de les tendències del gènere.

– Els escrits dels individus C i E tendeixen a aparèixer a la zona baixa de la banda i són, per tant, els més pròxims a la classe magistral.

– La classe de l'individu D és sistemàticament a la zona alta de la banda i és, doncs, la més pròxima a la prosa acadèmica<sup>259</sup>.

– La classe de l'individu E apareix sistemàticament a la zona baixa de la banda de la classe magistral i és, per tant, la més pròxima a la conversa.

– Les converses 1 i 5 apareixen en general a la zona alta de la banda i són, doncs, les més pròximes a la classe magistral.

– L'individu E, tant en la prosa acadèmica com en la classe magistral, s'ha situat en les posicions més baixes, de manera que mostra una tendència personal a un estil menys nominal i a una densitat lèxica menor en tots dos gèneres. Els altres individus no presenten cap correspondència d'aquest tipus.

– L'escrit de l'individu D, la classe de l'individu A i la conversa 4 són molt oscil·lants i no mostren cap tendència clara.

---

<sup>259</sup> L'únic apartat en què la classe de l'individu D ha quedat en una posició discrepant és en el de la repetició, fet que resulta altament congruent amb el comentari que ja havia consignat al punt 4.2.3 sobre el gran nombre de vacil·lacions, trencaments i reformulacions que presenta aquest individu respecte dels altres quatre de l'estudi. Així, el caràcter vacil·lant del seu discurs té com a conseqüència lògica un alt índex de repetició, ben pròxim al de les converses, encara que, simultàniament, sigui en la majoria de trets la classe més acostada a la prosa acadèmica. La disparitat entre el grup majoritari de trets i aquest de la repetició i la vacil·lació és un fenomen que, com ja he dit anteriorment, no m'he sabut explicar.



Pel que fa a la delimitació de les bandes, en conjunt han aparegut bastant ben definides, amb només tres excepcions: la conversa informal i la classe magistral s'han superposat parcialment en els adjectius (vg. taula 33, § 4.4.3), i la classe i la prosa acadèmica s'han barrejat notablement en els sintagmes preposicionals (vg. taula 55, § 4.7.3) i s'han confós fins a esdevenir indistingibles en els adverbis (vg. taula 37, § 4.5.3, on, en nota, he destacat el valor explicatiu d'aquestes interseccions). En els noms, els verbs, els mots lexicals, les nominalitzacions, la repetició lèxica i la longitud de mot les bandes de dispersió s'han format poc o molt distants entre si, sense cap superposició. Així, de les 27 bandes presentades, 21 han aparegut impecables i només 6 han estat superposades o confuses. Els resultats són, per tant, d'una gran coherència, cosa que en reforça la validesa interna i externa<sup>260</sup>.

La consistència d'aquestes bandes contrasta amb la indefinició de les de D. Biber (1988: 255, 264 i 269). En els quadres generals de resultats per gèneres, aquest autor dóna una mesura de la dispersió de cada tret (*Range*), que resulta ser gran o molt gran en la majoria de casos. Només a tall d'exemple, ofereixo les bandes de dispersió en quatre dels trets que hi apareixen:

TAULA 88 *	<i>Face-to-face conversation</i>		<i>Prepared speeches</i>		<i>Academic prose</i>	
	M	R	M	R	M	R
<i>Nouns</i>	13.7 %	5.4 %	18.9 %	6.8 %	18.8%	15.8%
<i>Attributive adjectives</i>	4.1 %	3.8 %	4.9 %	4.4 %	7.7%	8.3%
<i>Prepositions</i>	8.5 %	4.8 %	11.2 %	4.5 %	13.9%	9.0%
<i>Word length</i>	4.1	0.5	4.4	0.7	4.8	1.3

(\* Percentatges sobre el total de mots - *M*: *Mean value* [Percentatge mitjà], *R*: *Range* [Banda de dispersió] - Taula elaborada a partir de les dades de Biber, 1988: 255, 265 i 269)

<sup>260</sup> Tant la validesa interna com l'externa (vg. § 3.4.1) recolzen sobretot en una bona tria del corpus i dels trets a analitzar, però també es demostren amb uns resultats que no siguin contradictoris. A parer meu, uns resultats cadòtics, amb un gran nombre de superposicions, farien pensar que hi ha factors altres que el gènere que no s'han tingut en compte i que estan alterant les dades.

Com es pot veure, les bandes de dispersió (*Range*) són tan grans que en alguns casos gairebé igualen el valor del percentatge mitjà (*Mean value*). I quan la diferència entre el percentatge d'un gènere i el d'un altre és petita, les bandes resulten totalment superposades. Així, en el cas dels adjectius atributius, la dispersió a *Face-to-face conversation* és de 3.8 punts i abraça des del 3.0 % al 6.8 % (Biber, 1988: 264) i la dispersió a *Prepared speeches* és de 4.4 punts i abraça des del 2.8 % al 7.2 % (Biber, 1988: 269), de manera que la banda de dispersió del primer gènere inclou totalment la del segon. Una cosa semblant passa amb els noms, que a *Academic prose* tenen una dispersió de 15.8 punts, que abraça des del 8.4 % al 24.2 % (Biber, 1988: 255) i també inclou totalment les bandes dels altres dos gèneres.

A la vista d'aquests resultats, els de la present recerca són d'una gran coherència i les poques superposicions que hi he hagut s'hi poden considerar anecdòtiques, mentre que en el treball de D. Biber semblen ser la norma<sup>261</sup>.

#### **4.10.7- ESTIL NOMINAL I ESTIL VERBAL: EXEMPLES DE TEXTOS**

Com he anunciat a l'inici de l'apartat, donaré uns exemples de textos com a mostra del conjunt de trets lingüístics tractats fins ara. Els textos escollits, d'acord amb l'anàlisi efectuada de les bandes de dispersió, se situen entre els més representatius de cada gènere: la conversa 2, la classe de l'individu B i l'escrit de l'individu A (en presento fragments amb 245, 244 i 241 mots, respectivament). Per visualitzar amb més facilitat el contrast entre l'estil nominal i l'estil verbal, hi he assenyalat amb vermell els components de l'estil nominal comptabilitzats en aquest capítol (noms, adjectius

---

<sup>261</sup> Una justificació d'aquest comportament de les bandes de D. Biber podria ser en el nombre de textos que analitza, que és molt superior al meu, si bé amb fragments més curts. Segons la informació que aporta el mateix autor, ha treballat amb 44 fragments de *Face-to-face conversation*, 14 de *Prepared speeches* i 80 d'*Academic prose* (Biber, 1988: 67). Però, encara que el nombre de fragments pugui afavorir la dispersió, caldria explicar-ne les causes d'una manera més consistent, atenent als criteris de constitució del corpus i a les característiques de cadascun dels textos que s'han agrupat sota el rètol de cada gènere. Lògicament, no estic en condicions de tractar un aspecte que no és abordat ni pel mateix D. Biber en la seva obra, que sembla tenir-ne prou amb els resultats en valors mitjans que mostren les tendències més globals en els gèneres que estudia.

atributius i preposicions febles introductores de sintagmes en funció de complement del nom i de l'adjectiu) i els adverbis en *-ment*, i amb verd els components de l'estil verbal (verbs, adjectius predicatius i preposicions febles introductores de sintagmes en funció de complement verbal):

“PPP: ves\  
m'ha despertat quan baixava la la la a el llit\  
brum\  
MMM: xuc\  
sí\  
jo també ho he sentit\  
PPP: cony\  
què és això\  
CME: doncs jo he dormit\_  
PPP: llavons sí que--  
CME: oh\  
PPP: jo sí que m'he despertat llavons\  
amb això del llit\  
i llavons no podia dormir\  
cagondena\  
i llavons tinc els peus que em cremen\  
CME: vagi a mirar l'arròs\  
perquè em sembla que\_  
PPP: tenia que treure els peus a fora\_  
i llavòrens tenia tos\_  
MMM: ja\  
marxa l'aigua de l'arròs\  
PPP: haig de treure el peus a fora sempre\  
cada nit\  
fora del llit\  
cagon coi\  
sí\  
no puc dormir de tan calents que són\  
és de tant tantes hores dret home\  
ahir vaig estar set hores seguides dret\  
CME: ja\  
PPP: fent neules\  
ara jo dic\_  
hi hi anireu a futbol o no/  
CME: ah\  
no sé\  
PPP: si hi aneu bueno\  
si no el portaré en el Vicenç que és que és--  
CME: quin dia juguen\  
PPP: que és de Lleida\  
e- --  
CME: el diumenge hem d'anar--  
PPP: el dia vint\

CME: el diumenge hem d'anar a bateig\  
 PPP: diumenge\  
 no\  
 juguen\  
 CME: però\_  
 llavors\_  
 si el diumenge hem d'anar a bateig no marxarem el cap de setmana\  
 PPP: és un dissabte\  
 nena\  
 CME: per això\  
 que com que el diumenge hem d'anar al bateig\_  
 el dissabte ja no marxarem fora el dissabte i el diumenge\  
 PPP: sí\  
 ah\  
 i llavors l'altre sí\_  
 CME: ai\  
 jo què sé\  
 però vull dir que llavors jo puc a- --  
 PPP: bueno\  
 home\  
 CME: vull dir que llavors jo puc anar-hi segurament a futbol”  
 (Conversa 2, pàg. 56-58, lín. 2624-2687)

“ja la digressió de la classe d'ahir va servir per presentar els orígens\_  
 de les diguem-ne dues maneres principals\_  
 o almenys diguem-ne estadísticament més freqüents\_  
 de ser musulmà en aquest món\  
 es pot viure l'islam essent sunnita\_  
 o essent xiïta\_  
 i d'això parlarem\_  
 de com es concreta això\_  
 i com es concreta això\_  
 aquesta forma de viure l'islam\_  
 des del punt de vista d'aquesta part del programa de islam i política\  
 les implicacions polítiques que té una cosa i l'altra\  
 els orígens de la paraula ja sabeu quins són\  
 sunnita ve de la pretensió que tenen de ser els que millor segueixen la sunna\_  
 sunna vol dir el costum la tradició del profeta per descomptat\_  
 de ser els que so segueixen el costum del profeta\  
 xiïta és una paraula que ve de la paraula xia que vol dir bàndol\_  
 bàndol de alí\_  
 xia alí\_  
 partidaris de alí el cosí i gendre del profeta\  
 quines implicacions té des del punt de vista\_  
 del que hem de tractar en aquesta part del programa\_  
 viure l'islam a través del xiisme o viure'l a través del sunnisme/  
 mmh bàsicament la les diferències entre el sunnisme i el xin xiisme són\_  
 rauen precisament en això\_  
 en les implicacions polítiques\_  
 des del punt de vista diguem-ne dogmàtic\_  
 des del punt de vista dogmàtic\_

les creences bàsiques són idèntics\  
els punts fonamentals allò que defineix un musulmà\_  
les seves principals obligacions\_  
ee les quatre columnes de l'islam\_  
com en diuen ells\_”

(Classe de l'individu B, pàg. 187-188, lín. 8874-8904)

“Sorprèn comprovar que en el segment del mercat corresponent als habitatges destinats als grups socials de baixos ingressos, l'estructura de l'oferta fou similar, a grans trets, a la dels altres segments. Això marca una dissemblança fonamental amb altres indrets, com ara Anglaterra, on, com he indicat, els promotors professionals sobresortien en la producció d'habitatges de baixa qualitat. A casa nostra, els terratinents del Pla s'acontentaren amb els guanys obtinguts en l'especulació sobre la terra i en les operacions de producció del sol urbà (parcel·lació, urbanització), i, de forma més secundària, en el finançament concedit als compradors de ls solars a l'hora de construir. Indubtablement, en aquests negocis s'amassaven majors fortunes —i més ràpidament i amb menys riscos— que construint i venent cases. Però resulta ben curiós que l'activitat d'aquests agents no derivés cap a la producció en sèrie d'habitatges senzills —i ¿qui estava en tan bones condicions com ells per aprofitar aquestes oportunitats?—, o bé que no sortissin altres actors interessats a escometre aquesta tasca. Es fa difícil imaginar que l'explicació pugui ser que aquests propietaris no disposaven del capital suficient per desenvolupar la promoció a gran escala. Al cap i a la fi, des de mitjan segle s'accelera el procés de concentració de la propietat de la terra en el Pla de Barcelona, alhora que té lloc una intensa mobilitat de la propietat, clars símptomes que grups socials econòmicament poderosos «prenien posicions» amb vista al creixement urbanístic que venia.”

(Escrit individu A, pàg. 248-249, lín. 11677-11696)

El fet d'haver assenyalat els mots amb dos colors diferents ja permet de fer-se, a simple vista, una idea de la tendència nominal o verbal de cada fragment. La conversa informal té pocs mots en vermell, menys que la classe magistral, i aquesta menys que la prosa acadèmica, que en concentra molts en poc espai. El color verd, al contrari, domina més en la conversa que en la classe i en la prosa. La proporció entre els tres gèneres segueix la progressió esdevinguda ja arquetípica en aquest treball: corresponen a l'estil nominal el 17.7 % dels mots del fragment de conversa, el 40.3 % dels mots del fragment de classe i el 53.1 % dels mots del fragment de prosa acadèmica.

La distribució dels adjectius atributius (0 en la conversa, 9 en la classe i 23 en la prosa) i dels sintagmes preposicionals en funció de complement del nom o de l'adjectiu (2 en la conversa, 16 en la classe i 28 en la prosa) corrobora la mateixa progressió cap a un creixent estil nominal. També hi ha més nominalitzacions en l'escrit (16 en total, amb mots com *producció*, *concentració*, *explicació*, *dissemblança*, *finançament*, etc.) que en

la classe (només 4) i en la conversa, on no n'hi ha cap. Els verbs, com és lògic, segueixen el camí invers i són gairebé el doble en la conversa (47 verbs) que en la prosa acadèmica (25 verbs). I, finalment, els adverbis en *-ment* tenen dues ocurrences en la prosa acadèmica, tres en la classe i una en la conversa —curiosament, ja que ha coincidit a aparèixer en aquest exemple un dels escassos adverbis en *-ment* que hi ha en tota la conversa 2.

#### **4.10.8- CONCLUSIONS DE L'APARTAT**

En aquest apartat, que presenta de manera panoràmica els principals aspectes del capítol, s'ha vist com les dades sobre l'estil nominal i l'estil verbal, agrupades d'una nova manera, han accentuat les tendències derivades dels resultats dels apartats anteriors i han confirmat el predomini de l'estil verbal en la conversa informal, de l'estil nominal en la prosa acadèmica, i d'un grau intermedi entre els dos estils en la classe magistral. Així, s'ha comprovat que són certes tant la visió de M.A.K. Halliday sobre el paper dels noms i dels verbs en la distinció del mode del registre, com la tesi de W. Chafe sobre la construcció més densa, integrada i compacta de la llengua escrita.

La comparació entre les dades de la llengua catalana i l'anglesa ha palesat, a més d'unes divergències importants entre els diversos estudis fets sobre l'anglès, una tendència dels meus resultats a ser més rotunds i a oferir una progressió en general més accentuada.

Un resum dels trets lingüístics característics de cada gènere ha situat la classe magistral en una posició intermèdia indiscutible, que encara s'ha vist més reforçada quan s'ha calculat la mesura en què s'acosta cap a cadascun dels altres dos gèneres: la classe magistral ocupa un lloc molt central, mínimament decantada cap a la prosa acadèmica. Per tant, l'especificitat dels temes que s'hi tracten i la formalitat de la situació contrasten d'una manera bastant equilibrada amb les condicions físiques de producció *on-line* pròpies dels discursos orals.

Finalment, una anàlisi de la distribució de les dades per textos ha mostrat unes bandes de dispersió clarament delimitades en la majoria dels trets, amb algunes excepcions que també han estat explicitades. La comparació amb la dispersió, molt més gran, del treball de D. Biber sembla que permet atorgar encara més coherència i més validesa als resultats de la present recerca.

