



**Arxiu històric** FUNDACIÓ JAUME BOFILL

# Manuscrits, grafies i ordinadors

Joan Torruella Casañas

FEBRER 1994

FUNDACIÓ  
*Fundació*  
**JAUME**  
*Jaume*  
**BOFILL**  
*Bofill*

Joan Torruella i Casañas  
 Department of Spanish and Portuguese  
 University of Manchester  
 GRAN BRETANYA M13 9PL

## MANUSCRITS, GRAFIES I ORDINADORS<sup>1</sup>

0.- El tractament i anàlisi de textos mitjançant l'ajut de tècniques computacionals és, avui, un fet inevitable per a qualsevol estudi lingüístic o literari que es vulgui realitzar. Per això, un pas previ és sempre la introducció del text que es vulgui treballar a l'ordinador. En aquest sentit, els textos medievals, degut, bàsicament, a les seves irregularitats normatives i a les varietats gràfiques que presenten, plantegen una casuística pròpia que cal considerar. La seva resolució, però, ha de respectar sempre els criteris filològics de l'edició del text.

En aquest article pretenc presentar els principals problemes que això pot comportar i unes possibles solucions, no les úniques, a aquests.

Hi ha dos aspectes que cal considerar a l'hora de valorar els suggeriments que es fan en aquestes pàgines.

La primera és que les solucions que aquí es proposen per a resoldre i codificar les grafies i les seves variants dels manuscrits medievals estan pensades per ser aplicades amb el mínim d'ingerència possible en la normativa filològica i sense trair-la mai. Sobretot, pensant que no s'han de fer dues edicions, una apta per als treballs computacionals i una altra per editar, sinó que es pretén que les edicions, filològica i tipogràficament aptes per ser editades, puguin servir per als treballs computacionals. Seria molt diferent crear un sistema de codificació que facilités molt més la posterior lematització dels textos, tal com fan, per exemple, els que estan treballant amb el 'corpus' del Dictionary of the Old Spanish Language, però aquest sistema també té les seves febleses i només seria possible si poguéssim comptar amb mitjans econòmics i amb

<sup>1</sup> Aquest treball ha sigut possible gràcies a l'ajuda de la FUNDACIO JAUME BOFILL.

personal suficient com per a poder entrar tots els textos de cap i de nou, pensant només amb la confecció d'un 'corpus' informatitzat. Tocant de peus a terra, això no és així; cal pensar, doncs, amb la possibilitat d'aprofitar els textos que els investigadors, per motius ben diferents cada un d'ells, estan entrant als seus ordinadors, i amb la possibilitat d'entrar textos ja editats i impresos sobre el paper, mitjançant l'ajuda d'un lector òptic o 'scanner'.

La segona consideració és que, avui per avui, encara és impensable concebre, sobretot per a textos medievals, un programa de lematització morfològica automàtic fiable. En textos ideals i normalitzats s'arriba a aconseguir percentatges bastant alts d'eficàcia, però en textos reals i tan poc normalitzats, com els que presenten els manuscrits medievals, els resultats són tan baixos i caòtics que la feina de revisió és més feixuga que la de fer-la de cap i de nou.

Per tant, la lematització que es proposa en aquestes pàgines és una lematització gràfica. És a dir, les paraules s'ordenen no segons criteris morfològics sinó segons criteris gràfics, intentant, però, destriar, al màxim dintre de cada grafia, els diferents significats semàntics.

El que és important, però, és que el percentatge de lematització semàntica que s'aconsegueixi tingui un cent per cent de fiabilitat, de manera que només calgui recórrer al percentatge restant per a resoldre les vacil·lacions. Per exemple, *nes* representa un sol lema gràfic però pot tenir diversos lemes semàntics. D'aquests diferents lemes semàntics, el més extès, i amb escreix, és el quantitatiu. Doncs bé, si en el text ja tenim marcats amb un caràcter diacrític (vegi's 4.1 i 11.2) tots els *nes* quantitatius, ens estalviarem haver de recórrer a moltes línies de concordança d'aquest lema per a destriar-lo d'entre els altres. Així, encara que no puguem lematitzar semànticament al cent per cent, sí que podem reduir el percentatge de vacil·lació en gran nombre. El més important és delimitar i tenir controlats tots els casos confusos.

Una tercera consideració que caldria afegir és la responsabilitat de l'editor a l'hora d'interpretar les grafies. Si *p* representa per o par és ell qui ho ha de decidir; amb tot, sempre es pot ajudar de les llistes i de les concordances del document per veure tots els casos on apareix aquesta abreviatura i determinar sobre la seva resolució.

Finalment cal dir que, a l'Edat Mitjana, cada document és un món, i que, per tant, cada un d'ells presenta la seva casuística pròpia. Aquesta caldrà anar-la resolent sobre la marxa, però sempre tenint clars els criteris a aplicar, tant

filològics com computacionals, ja que les incoherències en aquest darrer camp són fatídiques.<sup>2</sup>

1.- Pel que fa a la confecció d'Índices verborum, concordances i lematització de textos, la paraula és sempre la unitat bàsica en què ens movem. Ara bé, el concepte 'paraula' no és tan simple com pot semblar Jiri Krámsky, al seu llibre The word as a Linguistic Unit,<sup>3</sup> en dóna una bona mostra. De moment, i a l'espera d'afrontar més concretament aquesta qüestió pel que fa als problemes de verbs compostos locucions adverbials, frases fetes, etc., prenc com a 'paraula' el conjunt de caràcters amb valor fònic delimitats entre espais.<sup>4</sup>

Una premissa que cal observar, per tal de no distorsionar el document original, és la de respectar totes les seves particularitats gràfiques. Esmenem, però, tot el que són errors evidents, la majoria de vegades deguts al copista.

D'entrada, això es pot resoldre de dues maneres. L'una, mitjançant la codificació de les diferències gràfiques del manuscrit i, l'altra, mitjançant la reordenació de l'alfabet del programa a l'hora d'ordenar les paraules (vegi's 11 i 12).

2.- Bàsicament hi ha tres sistemes de transcripció de manuscrits. El paleogràfic, que retrata fidelment totes les particularitats del manuscrit sense questionar-les, (ex: Dictionary of the Old Spanish Language)<sup>5</sup>; l'interpretatiu, que representa una edició crítica però partint d'un sol manuscrit, (ex: Tesoro della Lingua Italiana delle

<sup>2</sup> Robert L. OAKMAN, Computer Methods for Literary Research, (Athens, 1984), 2ª edició, University of Georgia Press. Susan HOCKEY, A Guide to Computer Applications in the Humanities, (London, 1980).

<sup>3</sup> Jiri KRÁMSKY, The word as a Linguistic Unit, (Paris, 1969), Mouton & Co Publishers.

<sup>4</sup> Per a més informació sobre la qüestió de les 'formes fixes' veure Maurice GROSS, Une classification des phrases 'figées' du français, a Revue Québécoise du Linguistique, 11.2, 1982. ---- Les noms traceurs, a Cahiers de Lexicologie, 44, 1984-1. Gaston GROSS, Jacques CHAURAND, i altres, Typologie des noms composés, Section du comité national du C.N.R.S., 42, Univ. Paris, XIII. Florence MONTREYNAUD, Agnès PIERRON i François SUZZONI, Dictionnaire de proverbes et dictons, Les Usuels de Robert, Paris, 1984. I el treball i bibliografia d'Ana de AGUILAR-AMAT Aproximación al estudio de las formas fijas, Barcelona 1989 (no publicat).

<sup>5</sup> David MACKENZIE, A Manual of Manuscript Transcription for the Dictionary of the Old Spanish Language, (Madison, 1984), Hispanic Seminary of Medieval Studies.

Origini)<sup>6</sup>; i el crític, que, partint, moltes vegades d'un stemma codicum, vol reconstruir el text ideal.<sup>7</sup>

Pel que fa al problema de les variants gràfiques, el primer sistema queda descartat, ja que l'estricta representació paleogràfica comporta, inevitablement, la inclusió d'errors i capricis d'autors i copistes i, per tant, en un tractament automàtic tot això hi quedaria inclòs.

La major o menor aptitud en la utilització dels altres dos sistemes depèn, principalment, de la finalitat amb què es faci l'edició. D'entrada, però, si la finalitat és lexicogràfica, el camí més útil és el de l'edició interpretativa. Usar aquest sistema, que es basa en la interpretació d'un sol ms., no vol dir que l'editor no es pugui ajudar d'altres mss. o que no hi pugui operar esmenes. Això es pot fer sempre i quan es reflecteixin en el text totes les addicions i supressions per tal de no atribuir a un document concret lliçons d'altres documents o correccions del transcriptor.

2.1.- En aquest sentit, caldrà, doncs, marcar totes les intervencions que es fan en el document, provinents de la col·lació amb altres documents o no, per tal de no ser preses com a paraules documentades en el document. De la mateixa manera, igualment, caldrà tractar a part les cites i glosses posteriors que el document pugui contenir (veg. 10).

Amb les facilitats que els scanner o lectors òptics posen avui al servei de l'editor, és una gran temptació utilitzar-lo per tal de posar a suport magnètic edicions modernes de textos medievals. Això és una operació viable sempre que es compleixin dos criteris: primer, que l'edició sigui fiable i, segon, que sigui possible i mínimament fàcil, a partir de l'aparell crític, transformar el text en una edició que reflecteixi només el ms.

3.- Com ja he dit abans, la unitat bàsica en el treball amb llistes i concordances i en qualsevol altre tipus de treball lexicogràfic és la paraula. Si prenem en principi la paraula com la sèrie de caràcters que es troben delimitats entre dos espais o signes de puntuació, ens trobarem ja amb els primers problemes: les aglutinacions de paraules que apareixen en els documents.

---

<sup>6</sup> D'Arco Silvio AVALLE, *Al Servizio del Vocabolario della Lingua Italiana*, (Firenze, 1979), Accademia della Crusca.

<sup>7</sup> Alberto BLECUA, *Manual de crítica textual*, (Madrid, 1983), ed Castalia. Franca BRAMBILLA, *L'Edizione Critica dei Testi Volgari*, (Padova, 1984), Editrice Antenore.

Aquestes aglutinacions s'han de codificar de tal manera que es delimitin les paraules sense distorsionar-ne les grafies. Normalment els filòlegs dins l'àmbit de la Romània usen tres caràcters, signes convencionals, que els ajuden a marcar les cotes de cada mot.

Aquests tres caràcters són: l'apòstrof, el punt volat i el guió. (Aquest darrer, com veurem més endavant, també té funcions accentuals. vegi's 4.3).

3.1.- S'usa l'apòstrof per a indicar les aglutinacions de dos mots amb pèrdua del darrer element del primer mot (l'anima, qu'ell, mi'amor).

3.1.1.- Queda per resoldre el cas de futurs separats on la *a* que converteix el segon verb en futur, també s'ha d'aplicar al primer (ex: menjar e beur'a). Bonner proposa separar la *a* de futur del segon verb mitjançant l'apòstrof per a indicar que aquesta *a* també és marca de futur del primer verb. La solució no és gaire operativa, ja que en una lematització automàtica apareixaria una *a* indeterminada fàcilment identificada amb preposició o verb, que no seria real.

3.1.2.- Una qüestió per a decidir és si s'apostrofa un pronom o un article davant d'una paraula començada per *s* líquida (qui n'scolta - qui·a scolta, ffins a l'stat - ffins al stat). Penso que, seguint un criteri fònic, és més adient apostrofar davant de *s* líquida, la qual cosa no comporta cap problema a l'hora de lematitzar automàticament.

3.1.3.- Les solucions dels articles contractes davant de paraules començades per vocal són opcionals, podent-se mantenir contractes o no. De totes maneres, per respecte a la forma del ms., si aquests articles hi són en la seva configuració contracta, és sempre millor no separar-los per unir el segon dels seus elements amb la paraula següent. A. Bonner, en les seves Normes per a la Transcripció i Edició de les Obres Catalanes de Ramon Llull, addueix al respecte tres criteris, que a mi em semblen molt convincents, a favor de no separar els articles contractes:

Crec que en aquests casos la separació a l'estil modern té tres caires negatius. El primer és la falsificació de la imatge mental de divisions fonètiques medievals (presumint, és clar, que estam treballant amb un Ms. suficientment conseqüent). Em sembla que l'escrivà medieval sentia del, per exemple, com una unitat lingüística tant davant consonant com davant vocal. El segon caire negatiu és forçar una simetria que no existeix al Ms., deformant frases com e·l coneixen, e·l honoren, e·l serveixen en e·l coneixen, e l'honoren, e·l serveixen, que té un cert aspecte de pedanteria innecessària i modernitzadora. El tercer caire negatiu és que ens amaga una mica una preferència

medieval catalana a formular articles i pronoms dèbils com a enclitics i no proclitics, amb gran freqüència de formes com e·l, que·l, e·n, per que·n, no·ls, que·o, si·s, fins i tot davant de vocal, donant e·l aucis, del home, etc.<sup>2</sup>

També Bonner crida l'atenció sobre alguns casos sense elisió de vocals quan la consonant inicial de la segona paraula ha estat modificada segons una nova posició intervocàlica, com per exemple la grafia freqüent d'assaber, o la grafia italianitzant d'alle (= a la)'.<sup>3</sup>

En aquests casos, Bonner creu que és millor no separar els dos elements ja que, de fer-ho, tindriem formes visuals estranyes (a ssaber) i en alguns casos fonèticament desorientadores (a lle). Crec, sobretot pensant amb la ordenació alfabètica automàtica, que és millor separar (amb espai, ja que no es perd cap element de les dues paraules que desunim) els dos elements. El problema resultant és el de tenir una paraula començada per doble consonant, però aquest fet normalment ja el trobem en els mss. sense ser aquest cas i, per tant, és una qüestió que caldrà resoldre en el seu apartat. (vegi's 7.4). Respecte a la creació d'unes formes fonèticament desorientadores, no crec que hi hagi gaire casos; l'únic que se m'acut és el presentat per Bonner (a lle), i en qualsevol cas, cal assumir-lo. Representa menys desconcertant tenir aquesta forma fonèticament enganyadora que no una aglutinació de dues paraules sense marcar-se; si fos així, la segona paraula no constaria mai en les llistes i la primera, fora de context, podria semblar una sola forma i resultar una paraula fantasma. Si es confirmés que alle és l'únic cas en què això es presenta, es podria considerar com un article contracte i no hi hauria cap problema per mantenir l'aglutinació. El guió de paraula composta també podria ser una solució per a aquest cas.

3.2.- S'usa el punt volat per a indicar les aglutinacions de dos mots amb pèrdua del primer element del segon mot (parla·b, e·ls, dona·m)

Cal fer notar aquí que les edicions de Els Nostres Clàssics tenen un sistema d'ús propi quan a la separació d'aglutinacions. En realitat segueixen estrictament les normes fabrianes i, a tot arreu on la situació escapa a aquestes normes, apliquen el punt volat. Però hi fan una distinció: si un dels dos elements desaglutinats és pronom feble o article determinat no deixen espai entre aquests i la paraula a la qual s'aglutinen perquè són àtons (d'aquesta·a té lo setge); mentre que si cap dels dos elements desaglutinats és

<sup>2</sup> Anthony BONNER, Normes per a la Transcripció i Edició de les Obres Catalanes de Ramon Llull, Comissió Editora de la Nova Edició de les Obres de Ramon Llull, Patronat Ramon Llull, inèdit. nota 19, pp. 7-8.

àton llavors deixen un espai entre el punt volat i la paraula sencera (d'est· ofença, un· arma ·n pot ser tolta).

Sobretot en poesia, també hi ha el costum d'indicar amb el punt volat les partícules us, ho (=u), hi (=y), en les seves diferents formes gràfiques, quan perden el seu caràcter sil·làbic en aglutinar-se amb la paraula precedent (ja·us veuré, si·us voleu, no·u sabré, tots ells de tres síl·labes). No crec que s'hagi de marcar aquest fet gràficament, donat que les aglutinacions fòniques en altres casos on no hi ha pèrdua de vocal no les marquem, i el punt volat, aquí, donaria la impressió de pèrdua d'algun element que en realitat no s'ha perdut.

Així, doncs, per a no distorsionar les representacions gràfiques de les paraules, aquestes tres partícules caldrà separar-les de la paraula precedent amb un espai en blanc, a no ser que la paraula precedent sigui un verb, amb la qual cosa, llavors, com a pronoms febles que són, caldrà separar-les amb un guió, (vegi's 3.3).

Moltes vegades, l'elecció entre l'ús de l'apòstrof o del punt volat és opcional. Quan en aquests casos una de les dues paraules, o les dues, és buida de valor semàntic i pot tenir més d'una categoria gramatical, l'elecció d'un o altre caràcter ens la pot indicar (Que·n (=prep.) que·n tot el dia, qu'en (=pron). qu'en vols?).

Tot i així queden situacions difícils de resoldre i que potser la millor solució és posar-hi una marca convencional i treura-la un cop fets els treballs ordinatorials. Aquest podria ser el cas de ·s que quasi sempre representa el pronom es però algunes vegades representa la 3ª persona del singular del present d'indicatiu del verb 'ser', és. Llavors podríem utilitzar, per exemple, ·s per a representar el pronom i ^·s per a representar el verb i, un cop fets els llistats i concordances, suprimir el caràcter ^. Això és, però, enterament opcional de l'editor i, evidentment en l'edició impresa no hauria d'aparèixer aquest signe; però fent-ho així, en les concordances ens estalviariem haver de resseguir moltes entrades del lema ·s per a separar les que són pronom de les que són verb.

En el cas que les dues paraules aglutinades tinguin la mateixa o semblant categoria gramatical i no presentin conflictivitat d'interpretació (nom-adjectiu, verb-nom, etc.) sempre és preferible deixar sencera la segona de les dues paraules i retallar la primera que no pas a l'inrevés, perquè de cara a l'ordenació alfabètica automàtica una paraula mancada de començament sempre queda situada molt més lluny del seu lloc real que no pas una paraula mancada de final, que a efectes d'ordenació pràcticament no es desplaça.



Així doncs, en un text, la paraula està sotmesa a tres possibilitats i és convenient marcar-les diferentment per tal que l'ordinador pugui identificar i classificar les paraules. Aquestes possibilitats són: que la paraula no sofreixi cap aglutinació, cosa que marquem amb espai o qualsevol caràcter de puntuació (o accentual com en el cas del guió); que la paraula estigui aglutinada amb la següent i perdi la vocal final, cosa que marquem amb un apòstrof; i, finalment, que la paraula estigui aglutinada amb la precedent i perdi la vocal inicial, cosa que marquem amb un punt volat.

Aquí cal fer un aclariment sobre el punt volat: aquest ha de ser representat per un caràcter del codi ASCII (ex. 249) i no pas per un caràcter formatejat, o sigui pel punt normal en format elevat (superescrit), com passa algunes vegades. Pensem que els formats que qualsevol processador de textos dona del text a la pantalla o a la impressora desapareixen quan el text és passat a ASCII; i per a treballar un text amb programes de concordances o altres processos per a l'estil, això és imprescindible.

3.3.- El guió s'usa normalment en quatre casos: el primer, per a indicar un mot partit a final de línia; el segon, per a enllaçar paraules compostes (aquí també cal posar-hi els números (vegi's 8.1)); el tercer, fent funció de parèntesi; i el quart, per a enllaçar el pronom feble darrera el verb.

Normalment els processadors de textos ja ofereixen tres tipus de guió; altrament, els codi ASCII (American Standart Character) n'ofereix també tres de diferents: núm 45 -, núm 95 \_ i núm 196 -.

3.3.1.- El núm 45 parteix les paraules a final de línia i, per tant, caldria usar-lo només per al primer cas. D'aquesta manera, li és fàcil, a l'ordinador a l'hora de treballar les paraules, no tenir en compte el guió de final de línia i considerar una sola paraula les dues parts separades per aquest guió. És cert que tots els processadors de textos ofereixen la possibilitat de justificar el text i, per tant, és innecessari aquest guió, però podria donar-se el cas que l'editor volgués respectar les línies del ms. i, llavors, aquest guió seria imprescindible.

3.3.2.- El núm 95 manté sempre unides les dues parts que enllaça i, per tant, es convenient usar-lo en el segon cas. Aquest guió sí que es tindria en compte a l'hora de córrer els programes i, si es volgués, podria considerar les dues parts que enllaça com a una unitat. El fet que aquest guió sigui baix (\_) estèticament dona una impressió una mica estranya, però és fàcil, un cop fets els treballs

computacionals, en un moment, canviar-los pel guió normal, que no usem aquí perquè el reservem per al quart cas.

3.3.3.- En el cas de guió equivalent al parèntesi, per tal d'evitar confusions i problemes a l'ordinador, és recomanable utilitzar el núm 196 però doble. D'aquesta manera es podria incloure en el programa el doble guió com a caràcter de puntuació -- que en realitat és el que és -- i així no hi hauria interferències amb els altres casos.

3.3.4.- En el quart cas, posem el guió, núm 196, en els pronoms febles quan van darrera el verb. Aquí, però, el guió no marca cap aglutinació ni cap pèrdua vocàlica sinó la tonalitat del pronom, per la qual cosa és més adient tractar aquesta qüestió a l'apartat d'accentuació (vegi's 4.3.1).

D'aquesta manera tenim un tipus de guió diferent per a cada cas i no hi haurà interferències a l'hora de córrer els programes.

3.3.5.- Darrerament la Dra. Lola Badia, en el seminari de medievalistes de la Universitat de Barcelona, ha propugnat, amb la finalitat de simplificar al màxim la qüestió de separacions de paraules, l'abolició del punt volat i del guió tot substituint-los per l'apòstrof i l'espai, respectivament. Em sembla que, donat que ja estem habituats a veure aquests caràcters en les edicions de textos medievals, és una simplificació innecessària i que fins i tot pot complicar la lectura del text. Per altra part, cal pensar que, a l'hora de lematitzar, la utilització d'aquests dos caràcters simplifiquen força el problema. Si només utilitzéssim l'apòstrof, l'ordinador no podria saber mai si es tracta d'una aglutinació amb pèrdua del darrer caràcter de la primera paraula o d'una aglutinació amb pèrdua del primer caràcter de la segona paraula. Caldria llavors construir un programa adicional que tingués en compte la posició de la vocal respecte de l'apostrofof i que en funció d'això, separés l'aglutinació. En qualsevol cas, si el punt volat i el guió molesten l'editor, un cop fetes les llistes i concordances, i s'hagi d'imprimir el text definitiu, qualsevol processador de textos pot canviar, en pocs segons, tots els punts volats per apòstrofs i els guions per espais.

No crec, com ja he dit abans, que s'hagi de pensar en dues edicions diferents, però sí que es pot pensar en una edició apta per a l'ordinador i fàcilment transformable en una edició per editar.

4.- Entrem ara a l'apartat de l'accentuació. Aquí cal partir d'un gest d'humilitat per part del filòleg i acceptar que encara hi ha, almenys en la llengua catalana medieval,

un gran desconeixement de parts de la seva fonètica. Això pel que fa a la posició de la síl·laba tònica de la paraula, però, encara més, pel que fa a l'obertura de les vocals.

Quant a la posició de la síl·laba tònica en paraules comuns, els problemes no són excessius (encara que hi són), però allà on la qüestió esdevé més complicada és en els noms propis. Possiblement mai no sabrem com pronunciaven a l'Edat Mitjana Ypocras, Macrobius o Pictagoras.

Quant a l'obertura de les vocals, el desconeixement és encara considerable. Els problemes que té un editor a l'hora de discernir rimes falses en paraules que avui tenen obertura vocàlica diferents són pràcticament insalvables.

Totes aquestes dificultats s'accentuen quan tractem textos que contenen paraules provençals o franceses, ja que aquestes tenien fonètiques ben diferents, o quan tractem textos escrits en català que no és del principat. Encara que l'escriptura de la Corona d'Aragó fos de les més uniformes, hi ha textos mallorquins (ex. R. Lluï) que presenten clarament obertures vocàliques oposades a les del Principat.

Totes aquestes llacunes seran més salvables a mida que puguem disposar de bastants rimaris entrats a l'ordinador. De moment, però, ens cal partir de la premissa que desconeixem bona part de la fonètica de les paraules. Per tot això, podem convenir que ens cal utilitzar l'accentuació només com a signes diacrítics convencionals.

De totes maneres, per què aquests signes diacrítics convencionals siguin eficients cal que es regeixin per unes normes fixes i concretes. Altrament, podria ser més dolorós el remei que la malaltia i la seva utilització amb ordinadors seria caòtica.

Disposem en principi de tres caràcters per a representar l'accentuació: accent, dièresi i guió precedint pronoms.

4.1.- L'accent, en català té dues possibilitats: agut i greu. Aquestes dues possibilitats depenen del tancament o obertura de la vocal que el porta. Però, com que ja hem dit que aquest valor encara ens és força desconegut, serà bo no donar una falsa impressió del seu coneixement i limitar-nos a l'ús d'un sol accent: l'agut sembla ser el més convencional.

Donada, doncs, la funció només diacrítica que atribuïm a l'accent, d'acord amb això podem reformular quan és convenient el seu ús. En aquest cas la proposta feta el 1926 per Mario Roques i postulada l'any següent a Catalunya per Jordi Rubió sembla força encertada:

La Commission s'est prononcée pour l'emploi exclusif de l'accent aigu: elle ne pense pas qu'il y ait lieu de chercher, comme on l'a fait souvent, très imparfaitement d'ailleurs, à distinguer les voyelles ouvertes et fermées; l'accent aigu doit servir seulement à marquer la place de l'accent tonique, quand celle-ci peut, suivant les habitudes de la lecture moderne, rester douteuse pour le lecteur; ... (ROQUES).<sup>9</sup>

En la qüestió dels accents, hom defuig de prejudicar la valor oberta o tancada de les vocals, decidint-se solament per l'accent tònic en els aguts, ... (RUBIO).<sup>10</sup>

Cal considerar, però, que aquestes normes estan pensades per editar textos medievals a base de llapis i fitxes i on totes les paraules s'observen en el seu context. Ara bé, avui l'editor i molt més el lexicògraf treballa amb ordinadors i sobre llistes alfabètiques de freqüències que aïllen la paraula del seu context. Això fa que, sense traïr el bon criteri filològic de l'escola d'en Roques, en ares d'una major operativitat i eficàcia a l'hora de lematitzar, es pugui estendre el criteri de l'accentuació tònica en els mots aguts fins on el transcriptor, editor o lexicògraf cregui convenient.

Un primer pas, del tot necessari, seria la de l'accentuació dels monosíl·labs homònims, ja que aquests són els més nombrosos i conflictius. El verb *haver* moltíssimes vegades escrit sense h, n'és el principal exponent (a prep. i verb).

També seria necessari qüestionar si cal o no accentuar els homònims polisíl·labs (*dona-dóna*) i fins i tot els casos de distinció entre nom i verb que es produeix en algunes paraules esdrúixoles (*idola-idola, sàvies-savies*).

Una altra possibilitat que es presenta a l'editor i que cal valorar, és, donat el fet que molts monosíl·labs homònims presenten més de dues possibilitats (*aes = posat, 30 dies, però, quantitat*), utilitzar altres formes d'accents a més de l'agut i, un cop fets els estudis ordinatorials necessaris, transformar-los en accents aguts o treure'ls, segons es cregui convenient.

4.2.- També la dièresi pressuposa moltes vegades un coneixement de la pronúncia de les paraules que no tenim.

<sup>9</sup> Mario ROQUES, 'établissement de règles pratiques pour l'édition des anciens textes français et provençaux', *Romania* 52, pp. 243-9. Jordi RUBIO, 'Regles pràctiques per a l'edició d'antics textos francesos i provençals', *Estudis Universitaris Catalans*, 12, pp. 309-12 [cf. Roques]. Alfred FOULET, i Mary ELAKEL SPEER, *On Editing Old French Texts*, (Kansas, 1979), The Regents Press.

<sup>10</sup> RUBIO, op.cit. p. 311.

La dièresi en català té una doble funció. Una, la d'indicar la pronunciació de la u en els diftongs creixents, i l'altra, la de desfer un diftong quan el segon element és una i (=y).

4.2.1.- En el primer cas, la dièresi sobre la u en els diftongs creixents, no crec que la seva abolicció portés cap problema, ja que no fa variar pas ni el significat ni el nombre de síl·labes de la paraula (qüestió-questió). A Mallorca i altres llocs pronuncien *questió*.

4.2.2.- El segon cas ja és més complicat.

Encara que l'ús en molts manuscrits de una h separant la i (=y) de la vocal anterior podria fer sospitar que a l'època medieval resolien la dièresi d'aquesta manera i que en els altres casos feien diftong, això no és així. Si ho fos, massa versos quedarien amb una síl·laba de menys, la qual cosa vol dir que moltes vegades una i (=y) precedida d'una altra vocal no formaven diftong i que per tant, el fet de la dièresi existia. El problema és saber quan existia i codificar-ho.

J. M. Pujol sintetitza molt bé la qüestió de la dièresi de la següent manera:

En efecte, les vacil·lacions gràfiques i la sobreposició de diversos estrats cronològics en el procés de còpia plantegen autèntiques apories filològiques: un fehya en un determinat text del tercer quart del s. XIII que ens ha arribat en còpia de mitjan s. XIV és faïa encara o ja feia? En els textos en vers, d'altra banda, si bé és innegable que la mètrica pot oferir més possibilitats d'orientació, també és cert que molt sovint hi pot haver implicada una concurrència de casos, una tria entre elisió i sinalefa o pot no estar clara la validesa d'alguns hiats.<sup>11</sup>

El que sí sembla clar és que, sobretot en poesia, d'alguna manera l'afecte de la dièresi és necessari.

La dièresi com a caràcter, moltes vegades també sembla un signe innecessari, ja que donat l'ús només tònic que hem atribuït a l'accent, aquesta fa la mateixa funció i és absurd usar dos signes per a una mateixa missió (*disinuit* = *disinuit*). L'elecció, en darrer terme, d'utilitzar un o altre caràcter dependrà de l'editor.

L'abolicció del caràcter en els casos on avui posem dièresi no tindria gaire importància en la prosa, ja que els casos on fa una funció diacrítica són mínims (*beneit-beneit*), però en poesia ajuda força a marcar el ritme i a justificar el

<sup>11</sup> J.M. PUJOL, *L'edició crítica (Modificacions a la proposta inicial)*, (Barcelona, 1989), p. 2. (no publicat).

número de síl·labes. No és estrany trobar en un mateix text poètic la paraula *juj* (=judici) que per raons sil·làbiques a vegades necessita desfer el diftong i a vegades no (*juj* = *juÿ*).

Pel que sí que no hi ha cap raó, és per fer les excepcions que fa la normativa moderna d'excloure l'ús del caràcter de la dièresi en els infinitius, gerundis, futurs, condicional, sufixos en *-isse*, *-ista*, etc. Això no aportaria cap avantatge i sí molts enrenous. (*beneÿr*, *salayr*, *hoÿreu*, etc.).

#### 4.3.- El guió.

4.3.1.- El guió davant dels pronoms també podem considerar-lo com un signe d'accentuació, ja que, en realitat, el que fa és marcar el pronom com a pronom feble, amb les conseqüències tonals i interpretatives que aquest fet pugui comportar. Nosaltres, que desconeixem les conseqüències tonals, altra vegada ens hem de limitar a usar el guió en un sentit diacrític de reconeixement del pronom com a feble, cosa que en alguns casos seria impossible sense alguna marca i que en altres portaria confusions en les frases (*stant-vos ab companyia* = estan jo amb companyia de vós. / *stant vos ab companyia* = estan vós amb companyia d'altres). En aquest cas, doncs, el guió ens ajuda a identificar la paraula que el precedeix com a pronom (*plangue-me mort* (=em planyi un cop mort) - *plangue me mort* (=planyi la meua mort)) i com a feble (*donar vos* (= vós sou donada) - *donar-vos* (=vós rebeu alguna cosa)), darrera d'un verb.

Caldrà decidir també quin paper ha de tenir el guió en les llistes i concordances. Podem optar per mantenir el pronom junt amb el verb que el precedeix; que el pronom precedit de guió aparegui a les llistes i concordances immediatament després del pronom no precedit de guió, que en les llistes les dues formes siguin una sola entrada però que en la concordança sota l'entrada sense guió apareguin tant les formes sense guió com les formes amb guió; i, per últim, que es prescindixi totalment del guió i tots s'agrupin dins la mateixa entrada. (Vegi's punt 11.).

Opcionalment, sobretot en poesia, el guió també pot marcar la separació d'un diftong format entre la darrera vocal del verb i la primera del pronom (*placia-us* (=4 síl·labes) - *placia us* (=3 síl·labes), *sovenque us la jornada* (7 síl·labes)). Però no crec que aquest fet sigui raó suficient per utilitzar un guió per a marcar-ho i, doncs, considero desaconsellable el seu ús.

4.3.2.- Un punt en el qual no hi ha acord en la seva resolució gràfica és el de la /z/, d'origen provençal, abans de la consonant i representada pels escribes indiferentment amb *-z*, *-t*, *-d*, *-s* o *-n* que s'usava, sobretot en poesia, per evitar

el hiatus de que, qui, ne o e amb una vocal següent<sup>12</sup> (que-n altrament, que-s ab plazer prenda, E-z el pillart). Aquesta /z/ és purament fònica i està totalment buida de contingut semàntic i morfològic. En aquest cas crec que és interessant separar aquesta lletra de la paraula a la qual està unida, ja que, com ja he esmentat, es tracta d'un caràcter purament fònic, extern a la paraula i la seva inclusió la distorsionaria. La millor manera per codificar aquest fet és separant la lletra de la paraula amb el caràcter guió. Si hi deixem espai quedarà una consonant sola que no sabrem quina funció fa en el text. Si hi posem apòstrof o punt volat semblarà que hi ha hagut una reducció vocàlica que en realitat no s'ha produït, (els editors provençals, que no utilitzen el guionet, solen posar punt volat), mentre que si posem guió no es pot confondre amb cap pronom i només pot representar aquest caràcter fònic. Hi ha qui emmarca aquest caràcter entre parèntesi; això tampoc em sembla gaire adient perquè els editors utilitzen normalment aquests caràcters per altres usos.

5.- Un altre problema a tenir en compte és el de les abreviatures. En aquest cas, el transcriptor allò que fa és donar uns signes gràfics a uns valors fonètics representats convencionalment al document.

5.1.- Aquí es poden presentar dos problemes: l'un, que no coneguem exactament quin és aquest valor fonètic (con-com), i l'altre, que coneguem quin és aquest valor fonètic i que no sapiguem com representar-lo (ny-yn).

En aquestes ocasions cal recórrer a l'estadística d'aquests casos en el document i mirar de seguir l'hàbit de l'autor o del copista. Si al document la nasal abans d'una oclusiva bilabial normalment s'escriu n (ex. impossible) sembla ser més respectuós transcriure també amb n aquest so quan desfem un abreujament; o, si la nasal palatal sonora apareix normalment escrita amb yn, també per respecte al document, és millor, quan aquesta està abreujada, mantenir la mateixa grafia que no pas escriure-la amb la grafia moderna (ny).

En principi, doncs, el més convenient és resoldre les abreviatures paleogràfiques segons l'hàbit del copista en els casos que aquests sons no estan abreujats en el document. El que passa és que també el copista moltes vegades fluctua entre dues solucions. Llavors, potser el criteri més vàlid a seguir serà l'etimològic, sempre i quan el resultat no atempti contra l'hàbit gràfic del ms.

Així, per exemple, en el cas de p-aula on normalment la barra indica er i l'etimologia demana ar, si a tot el

<sup>12</sup> Nathaniel B. SMITH, i Thomas G. BERGIN, *An Old Provençal Primer*, (New York-London, 1984), Garland Publishing, p. 22.

manuscrit hem trobat paraula o bé sempre hem codificat la barra com a er, és millor mantenir l'hàbit i transcriure paraula; però si sempre hem trobat escrit paraula o la codificació d'aquest signe ha anat fluctuant entre er i ar, és millor seguir la solució etimològica i escriure paraula.

Queden però encara quatre casos per resoldre: el dels signes & (unperson) i r (tironià); el dels números; el de abreviacions com ara rch o jhs. (dintre d'aquest darrer grup, hi podem incloure el particular cas de Jesuchrist); i les abreviatures que indiquen parts d'un text (cap.).

5.2.- Respecte a & i r tenim, en principi dues possibilitats: deixar-los igual o transcriure'ls. El primer cas, donat que són signes prou coneguts, no comportaria cap problema d'interpretació i fins i tot, tal com apunta J.M.Pujol, 'estalviaria de donar al text una superficial llatinització que fa més nosa que servei'.<sup>13</sup> Portaria, però, uns certs entrebancs a l'hora d'ordenar les paraules alfabèticament, ja que no sempre aquest signe representen una lletra sola i, per tant, no es podrien incloure a l'alfabet en un lloc concret. Si aquests signes es transcriuen, es poden resoldre, com en l'apartat anterior, segons l'hàbit del ms. o, en cas que no n'hi hagi, segons l'etimologia.

5.3.- Pel que fa a les abreviatures als numerals, aquí només tractarem de les terminacions llatines (.xv.º) o vulgars (.xv.º<sup>n</sup>) que porten els ordinals escrits en xifres romanes, normalment elevades (superescrit). (Per als altres problemes que comporten les representacions gràfiques dels numerals vegi's 8.).

Aquí, el signe 'º' no comporta cap problema, ja que és un caràcter ASCII, però la forma 'º<sup>n</sup>', donat que el superescrit és un format i no un caràcter, si; com ja he esmentat anteriorment, els formats al passar el text a ASCII es perden i, per tant, en resulta una forma idèntica a la preposició en.

La solució és que, com que aquests dos caràcters sempre aniran seguits d'un punt (vegi's 8.2.), al final de l'alfabet que regirà el programa de destriar i ordenar les paraules, s'entri com a una sola lletra la forma en. de la mateixa manera que també s'haurà d'entrar la forma º.

5.4.- Les abreviacions de casos com ara jhs (=Jesús) són una mica especials, ja que en lloc d'una abreviatura, es tracta d'unes inicials. En aquests casos crec que el més respectuós, donat que són inicials prou conegudes com perquè no portin a confusions, és mantenir-ho com en l'original, ja

<sup>13</sup> J.M. PUJOL, L'edició crítica (Barcelona, 1989), p. 2. (no publicat).



que aquests caràcters en una llista tothom sabrà a què es refereixen.

5.4.1.- Jesucrist és una paraula molt utilitzada en els textos de l'Edad Mitjana i apareix abreujada de moltes maneres (Jh'u xist, J-huxst). En cas que al ms. no aparegui escrita d'una manera regular, pot usar-se la forma més estesa: Jesuchrist. Si es considerés la solució massa atrevida, sempre es pot justificar en nota a peu de pàgina o a la introducció.

5.5.- Un altre cas és el de les abreviatures de paraules que indiquen les parts d'un text. Aquestes normalment es refereixen als capítols i van representades per cap, c o q. En el cas que l'editor decideixi desfer l'abreviatura no hi ha cap problema; però, en el cas que l'editor vulgui mantenir-la, llavors en les llistes, hi haurà una interferència entre cap (=part superior del cos/negació) i cap (=capítol), o l'aparició d'una c o una q soles, que potser en aquell moment no s'assimilaran a la paraula capítol a la qual representen.

Donat que normalment aquestes abreviatures sempre van seguides de punt, la solució en aquest cas és entrar-les com una lletra a l'alfabet del programa (cap., c., q.).

5.6.- Les sigles són un altre cas. Aquestes són lletres inicials que, per conveni, es fan servir com a abreviatures de la paraula a què pertanyen. El seu tractament, però, no pot ser el mateix que el de les abreviatures ja que les sigles tenen entitat per elles mateixes mentre que les abreviatures no. Per exemple, les sigles I.N.R.I. són més conegudes com a tals que no pas com a les paraules que representen (Jesús de Natzaret, Rei dels Jueus). La manera d'instruir el programa perquè identifiqui les sigles com a una unitat és que tracti el punt + Majúscula com a un caràcter alfabètic; així en les llistes i concordances les sigles es mantindran unides, com una sola unitat. Les possibles coincidències amb els números romans que pugui haver-hi al text (vegi's 8.2.) queden descartades pel fet que en els números romans no hi va punt entre els diferents caràcters i en les sigles sí. Per tant, doncs, és convenient posar entre cada lletra de les sigles un punt sense espai; ex, V.I. (= Vostre Il·lustríssima) diferent a .VI. (= 6).

5.7.- Per acabar aquest apartat de les abreviatures cal fer un aclariment. Hi ha qui les resol escrivint la part abreujada en itàliques. Aquí s'ha de tornar a dir el que ja s'ha dit en el cas dels caràcters en superescrit: les itàliques són un format i per tant a l'hora de passar el text a ASCII es perd i els caràcters tenen els mateixos valors que els caràcters sense itàliques.

Podria ésser interessant en alguns casos deixar constància del desenvolupament de les abreviatures, ja sigui per què constin en el text editat, ja sigui per estudiar-ne l'ús que en fa el copista, etc. Llavors el que es pot fer és emmarcar els caràcters abreujats de la paraula entre dos signes convencionals, els quals incorporats convenientment al programa d'ordenació, no afectaran l'ordre de les llistes. (ex: esp(er)anGa, majorme(n)t, etc.).

6.- Vocalisme. Entrem ara en l'apartat de les interferències que ens trobem entre les vocals en els documents medievals.

Aquí tenim dos tipus d'interferències: unes merament tipogràfiques comunes a les llengües medievals (i-j), i unes altres fonètiques pertanyents a cada llengua en concret (cat. a-e).

6.1.- En el primer cas, les interferències són entre els binomis i-j i u-v. 'Seria desitjable normalitzar aquestes últimes segons el so que tenien antigament, però per fer això cal imposar una diferenciació tipogràfica inventada per Gian Giorgio Trissino l'any 1524.<sup>14</sup> Durant l'Edat Mitjana, els signes semblants a les nostres j i v eren només variants gràfiques de la i i la u; és a dir, aquestes variants s'utilitzaven, no per distingir els fonemes sinó només per a facilitar-ne la lectura, per exemple quan la i i la u anaven precedides o seguides de certes lletres que en la escriptura gòtica podien comportar una confusió: ni, in, iii, i ui eren quasi indistingibles de la n, per tant s'escrivia nj, jn, ij, vi'.<sup>15</sup> Per exemple, en l'alfabet que dona Berenguer de Noya en la seva preceptiva Mirall de Trobar, del segle XIV, encara hi falten les lletres j i v.<sup>16</sup>

Donat, doncs, que els escribes no tenien grafies especials per a /dz/, ni per a /v/, sovint es presenten casos fonèticament dubtosos o impossibles de resoldre avui. Per exemple, coijat-comiat, que fins i tot, pel seu número de síl·labes, veiem que en un mateix text poètic alternen les dues solucions; altres casos poden ser neua-neva, jo-io, abreujar-

<sup>14</sup> CORRIGAN (ed.), *Ariosto/Trissino, Two Renaissance Plays*, Manchester University Press. R.B. MCKERROW, 'Some Notes on the Letters i, j, u and v in Sixteenth Century Printing', *The Library*, 3<sup>a</sup> sèrie, vol. I, (Londres, 1910), pp. 239-259. Edward JOHNSTON, *Writing & Illuminating & Lettering*, (London-New York, 1906), reprinted 1987, A&C Black Publishers. p. 247, nota 1. J.C. EGBERT, *Introduction to the Study of Latin Inscriptions*.

<sup>15</sup> Joan TORRUELLA, i Jeremy LAWRENCE, 'Un Projecte d'Arxiu Informatitzat de Textos Catalans Medievals : Algunes Normes', *Llengua & Literatura* 3.

<sup>16</sup> P. PALUMBO, (ed.), *Berenguer de Noya, Mirall de Trobar*, (Palermo, 1955), Manfredi Editore.

abreviar, etc.<sup>17</sup> Aquí, però, la decisió pertany a l'editor que haurà de resoldre el dilema de la manera filològica més convincent. La decisió pot tenir una certa influència en l'estadística en els casos que, degut a aquest problema, tinguem una mateixa paraula representada per caràcters diferents; però pel que fa a l'ordre en les llistes, el problema no és gaire greu ja que, donat el fet que en l'alfabet els dos binomis estan formats per caràcters contigus, les dues paraules aniran ordenades en llocs molt pròxims, probablement l'una després de l'altra.

6.2.- En el segon cas, el de les interferències fonètiques, la qüestió ja és més pròpia de cada llengua.

6.2.1.- En català, encara que no en tots els seus dialectes, la problemàtica principal és la de la representació de la vocal neutra que moltes vegades l'escriba o el copista feien amb una a o amb una e sense cap criteri ni regularitat i bastant a l'atzar (enemichs-enamichs-aneimichs-anaimichs).

D'això se'n deriva el problema que una mateixa paraula que té les seves vocals neutres representades per diferents lletres, en una llista automàtica no solament apareix repetida dues o més vegades sinó que en les estadístiques es comptabilitza com a dues o tres paraules diferents.

6.2.2.- El mateix passa amb la o en posició àtona que en alguns dialectes catalans es pronuncia /u/ i que per tant crea també incoherents usos de la o o de la u en la seva representació (engoxós-enguxós).

Aquí em sembla que no hi ha solució o, si més no, una solució àgil, ja que marcar en el text totes les desviacions que una paraula sofreix respecte a la seva representació etimològica podria ser una feina massa feixuga. Llavors el repàs i reajustament de les llistes a mà, de moment, em sembla ser l'única solució.

7.- Consonantisme. Aquí tractarem els diferents problemes que les irregularitats en l'ús de les consonants provoquen a l'hora de lematitzar i ordenar les paraules. La majoria d'aquests problemes, de moment, no tenen solució però el fet de tenir-ne consciència ja és un primer pas.

7.1.- S líquida. Normalment els grups *S*oclusiva llatins (SP, ST, SC) en català experimenten una prótesis de e abans de la s (scuma-escuma). En molts casos, però, aquest fenomen no es produeix, a vegades com a fet natural o a vegades perquè el copista volia llatinitzar el text. Com a resultat d'això

<sup>17</sup> BONNER, op. cit. en la nota 17 dona més detalls de la qüestió.

tenim que una mateixa paraula de tant en tant està escrita amb s líquida i de tant en tant no. Llavors es torna a produir el problema de duplicitat d'una mateixa paraula i que en l'ordre alfabètic les seves variants gràfiques apareixen a llocs molt distants. Una possible solució, que no resoluria la duplicitat de la paraula però que almenys situaria les dues versions en llocs pròxims, fóra la d'entrar a l'alfabet del programa les combinacions SP, ST, i SC en el mateix apartat de la e, però llavors ens trobaríem que, quan aquestes combinacions apareguessin en mig de paraula, serien aquestes paraules les que quedarien desordenades en les llistes. Per a evitar això, caldrà igualar aquestes combinacions a la e només quan es presentin a principi de paraula, o sigui, perquè l'ordinador ho entengui, quan vagin precedides d'espai (e=E=#sc=#Sc=#sp=#Sp=#st=#St).<sup>18</sup>

7.2.- Una altra interferència que sovint trobem en els mss. és la de la ç. En realitat, el segon caràcter vindria a ser una variant del primer. En els documents de l'època medieval trobem la ç utilitzada amb força confusió, normalment alternant capritxosament amb la c, amb valor sord.<sup>19</sup> Per exemple, en els tractats medievals Mirall de trobar de Berenguer de Noya de principis del segle XIV i en el Torcimany de Luis d'Averçó de finals del segle XIV, la ç no apareix pas a l'alfabet que donen però sí en el text. Aquesta irregularitat ens porta de nou a un problema de duplicitat de paraules. Però, si en l'alfabet del programa posem contigües les dues lletres, les diferents formes de la mateixa paraula ens apareixeran una després de l'altra.

Hi ha casos molt evidents, pel sentit, que l'escriba ha oblidat el traç a la ç; llavors és competència de l'editor el posar-la o no.

7.3.- Els binomis oclusius (P-B, T-D, C-G) també presenten a vegades interferències que provoquen desajustaments a les llistes i duplica paraules (sobtament-soptament). En aquest cas, però, la solució és feixuga: caldrà igualar a l'alfabet del programa tots els binomis seguits de consonant (ex. bt=pt ... ), altrament, la col·locació adient de les paraules s'haurà de fer manualment.

7.4.- Una altra qüestió és la de les lletres dobles (deffendre, sser). Aquí cal distingir entre si canvia de sonoritat al duplicar la lletra o no, perquè al canviar de sonoritat, hi ha la possibilitat que també canviï de significat (rosa-rossa, para-parra, cala-calla), mentre que si no canvia de sonoritat, tampoc canvia de significat. En català, les lletres que al duplicar-se canvien de sonoritat són r-rr i s-ss

<sup>18</sup> Prenc el signe # per a representar l'espai.

<sup>19</sup> Per informació sobre l'origen de la ç veure: R. MENENDEZ FIDAL, *Cantar de Mio Cid*, Vol. I, (Madrid, 1964), Espasa Calpe, pp. 213-221.

al mig de paraula i l·l tant al mig com a principi de paraula.

Quan al duplicar la lletra canvia la sonoritat, no hi ha cap problema, ja que a les l·listes cal que siguin paraules diferents. La qüestió està en els casos que al duplicar la lletra no canvia la sonoritat i que, per tant, interessa que les paraules de les dues variants caiguin sota del mateix lema. La solució està en igualar la doble lletra a la lletra simple a l'alfabet del programa. Cal tenir present, però, que per tal que el programa reconegui quan la rr o la ss són a principi de paraula, com fèiem amb la s líquida (veg. 7.1) caldrà marcar-les sempre precedides d'espai (#rr, #ss).

7.5.- Lligat al problema de les dobles lletres, hi ha la qüestió de la l geminada (l·l).

Aquí, tal com ens passava amb l'accentuació, hem d'acceptar que encara desconeixem bastant el seu funcionament, el qual, ja en l'època medieval, era complex. Tal com explica Bonner, 'En la llengua medieval hi ha cinc sons laterals: 1) la l simple; 2) la l doble (avui escrit l·l, o ll) per a mots com batll(·)e, o ametll(·)er del dialecte balear; 3) una l palatal procedent del llatí ll o s'l (gallina, illa), que en provençal donava una l "suau" ... i en castellà donava ll; 4) una segona l palatal procedent del llatí (l)ly (palla, all) o d'un grup d'oclusiu + vocal postònica que desapareix + l (auric(u)la > orella, pop(u)lu > poli) que en provençal donava la l palatal corrent, grafiada lh, i que en castellà donava j (als Mss. medievals catalans que distingeixen aquestes dues l palatals, la primera s'escriu ll i la segona ly o yl; ... 5) la l palatal doble de la majoria dels dialectes centrals (avui escrit ll(·)). Ara bé, el sistema ortogràfic del català medieval, per admirable i coherent que arribàs a ser en mans dels millors escriptors, no tenia sempre capacitat per a fer front a una situació tan complicada'.<sup>20</sup>

La l geminada és mínimament segura quan correspon a un cultisme, però en els altres casos, la seva utilització és insegura i, citant J.M.Pujol, 'és particularment blasmable ... en grafies no etimològiques (com al·legre, per exemple) on tenen un fals valor de regularització ortogràfica que no fa altra cosa que fomentar una imatge fonètica més absurda que la que es pretén evitar'.<sup>21</sup>

Totes aquestes incerteses i vacil·lacions fan que l'actitud més honesta sigui acceptar la inseguretats i no

<sup>20</sup> BONNER, op. cit. p. 10, nota 28.

<sup>21</sup> J.M. FUJOL, L'edició crítica (Barcelona, 1989), p. 2. (no publicat).

utilitzar la l geminada.<sup>22</sup> Casos com per exemple la no distinció entre les dues entrades que presenta el DCVB nul·la (=no té cap valor. cast. 'nulo') i nulla (=cap. cast. 'ningún') no crec que portin massa problemes.

7.6.- La H, de procedència etimològica llatina, té un ús molt irregular en els textos medievals. Degut a això, ni li podem atribuir mai amb seguretat una funció morfològica ni semàntica i, per tant, roman com una lletra buida de qualsevol contingut (per exemple, és absolutament normal trobar en un mateix ms. haver-aver). La seva aparició al mig de paraula també és força inconstant (plaher, loharia, crehensa). Considerant això, tenim tres possibilitats: una, que vulguem que la H no tingui cap influència i que, per tant, en les llistes i concordances, una paraula amb H es comptabilitzi dins de la entrada de la mateixa paraula sense la H; una altra, que vulguem que la H no influeixi en les llistes, però que en les concordances, sota la entrada sense H hi apareguin les diferents variants (amb H o sense), d'una mateixa paraula; i, tercera, que considerem la H com un caràcter prou significatiu com per què en les llistes i concordances apareguin com entrades diferents les variants d'una mateixa paraula (amb H o sense). (Vegi's 11.).

L'elecció del tractament que cal donar a la H, a l'igual que passa amb altres caràcters, dependrà, en bona mesura, de la importància que l'investigador doni a les variants gràfiques.

7.7.- La Y moltes vegades apareix, indiscriminadament, en funció de semiconsonant (ex. servey, veyá, etc.) igual que la i. La millor solució serà entrar els dígraf ny i yn a l'alfabet del programa després de la n i igualar les dues is (i=y) en el lloc de la i en l'esmentat alfabet. D'aquesta manera veyá i veia apareixeran contigus en els llistats.

7.8.- Les irregularitats en l'ús de la r muda no sol implicar problemes, ja que aquestes es presenten a final de paraula (coló= color) i, conseqüentment en les llistes una forma apareix després de l'altra.

8.- Pel que fa als números cal dividir la qüestió en dos apartats: els cardinals escrits en lletres i els números escrit en xifres romanes. Els números escrits en xifres aràbigues que es puguin trobar en tractats científics, astronòmics, etc. no presenten cap inconvenient, ja que aquestes s'entren dins l'alfabet del programa.

<sup>22</sup> Respecte a la problemàtica de la l geminada veure: J. ANGLADE, *Grammaire de l'Ancien Provençal*, (Paris, 1921), p.191. C. APPEL, *Provenzalische Lautlehre*, (Leipzig, 1918), p. 68. J. COROMINES, *Entre dos llenguatges*, (Barcelona, 1976), vol.I pp. 13-51.

8.1.- El cardinals escrits en lletres, en la normativa actual porten un guió entre les desenes i les unitats i entre les unitats i les centenes. Aquest guió, però, no té cap valor ni cap funció, per la qual cosa el més operatiu és suprimir-lo i decidir si al seu lloc, s'hi posa espai o s'escriu tot el número seguit (set cents - setcents, trenta dos - trentadós). La forma més clara és la de deixar espai ja que escriure-ho tot seguit pot donar unes concentracions consonàntiques una mica estranyes (cinccents). Tampoc l'argument que el guió manté unida l'estructura del número és vàlida, ja que aquest només es posa en certs casos (entre centenes i desenes el guió no és posa). Si tan interessant es cregués mantenir l'estructura del número unida, llavors el guió s'hauria de posar unint totes les seves parts. Llavors caldria utilitzar el guió que usem per a les paraules compostes (veg. 3.3).

8.2.- Les xifres romanes tenen el problema que a vegades es poden confondre amb paraules (vi (=6) - vi (=beguda)). La solució pot ser, seguint la grafia de molts mss., posar un punt abans i un altre després del número (.vi.) i entrar totes les combinacions possibles de punt + primera xifra romana a l'alfabet del programa (.i, .v, .x, .l, .c, .d, .m).

En algunes edicions, aquests números van en itàliques, però ja s'ha esmentat anteriorment que les itàliques són un format i els formats desapareixen al passar el text a ASCII.

(Respecte a les abreviatures volades que moltes vegades porten aquests números vegi's 5.3).

9.- Normalment, l'ús de les majúscules en la transcripció de mss. es regularitza segons el costum modern. Per això les majúscules perden qualsevol funció distintiva dintre del text, amb la qual cosa el més prudent és igualar sempre la majúscula i la minúscula a l'alfabet del programa.

10.- En una edició preparada per a ésser treballada amb ordinador, hi podem trobar tres tipus d'afegits que d'alguna manera cal tractar. Aquests són: afegits deguts al propi autor; afegits deguts al copista o a altres mans; i afegits deguts a l'editor.

10.1.- Els afegits deguts al mateix autor es poden incloure directament al text o, si es vol, posar-los entre dos signes convencionals (ex. {}) per tal de tractar-los a part.

10.2.- Els afegits del copista o d'altres mans, en principi no s'han d'incloure en el text. Llavors es poden

tractar a part, per la qual cosa cal emmarcar aquests afegits entre els dos signes convencionals abans esmentats, o bé es poden desestimar i no ser tractats. En aquest cas, aquests afegits s'han d'emmarcar entre dos signes convencionals diferents dels anteriors (ex. {}), per tal d'excloure'ls dels tractaments ordinationals.

10.3.- Els afegits deguts a l'editor normalment mai no s'han d'incloure en el text. Aquí hi poden haver dos tipus d'afegits: reconstruccions del text, i ordres per a l'ordinador.

Les reconstruccions del text, a no ser que siguin evidències claríssimes, i tot i així seria discutible, cal excloure-les de la base del text; vull dir que, per exemple en una concordança, podrien entrar com a context d'un lema però no com a lema.

Les ordres per a l'ordinador són totalment externes al text i, per tant, han de ser emmarcades clarament com a tals. Aquestes normes normalment marquen divisions internes en el text, reals o no, per tal de fornir unes pautes al programa per a donar les referències de cada lema. Per exemple, el lema X es troba en el capítol N<sup>o</sup>, a la pàgina M<sup>o</sup>, i a la línia P<sup>o</sup>, etc. Cada programa especifica les peculiaritats que han de tenir aquestes ordres i els caràcters que les han d'emmarcar. Cal tenir present que les línies plenes per 'ordres per a l'ordinador' no entrin en el recompte de línies que fa el programa de cada pàgina per a donar les referències de la situació de cada lema en el text.

10.4.- També moltes vegades és convenient tractar a part les paraules d'altres idiomes que el ms. pugui contenir (llatí, normalment, però també provençal, francès, italià, aragonès, castellà, català, àrab, galaico-portuguès, etc.), ja que aquestes podrien distorsionar la llista de paraules o la concordança, sobretot en paraules que en dos idiomes s'escriguin igual però tinguin significats diferents (pau = pau en català - pau = pavo en provençal) (veg. 2.1). Així, doncs, caldrà marcar convenientment aquestes paraules, frases o fragments del text.

De la mateixa manera, també les glosses que moltes vegades hi ha al marge d'un ms. poden ser motiu d'estudi a part, per la qual cosa caldrà marcar-les igualment.

10.5.- Resumint, doncs, en un text poden haver quatre tipus d'aportacions que, pensant en llistes i concordances, caldrà marcar distintament:

- parts de text que volem com a lema i com a context (no-res de senyal). Aquí hi ha la major part del text.



- parts de text que volem que entrin com a context d'un lema però no com a lema ([ ]); ex: reconstruccions de l'editor, paraules estrangeres, etc.
- parts de text que no ens interessin per res i que, per tant, no volem ni com a lema ni com a context o que volem tractar a part ({}); ex: glosses, cites d'altres autors, etc.
- ordres per a l'ordinador (<>); ex: principi de capítol, inici de nova numeració, etc.

El cas de les paraules estrangeres i glosses podria entrar en el segon o tercer grup o marcar-les amb altres caràcters (ex: %) per a fer-ne un tractament a part.

11.- De la mateixa manera, dins un text preparat per a ser tractat amb ordinador, podem trobar cinc tipus de caràcters que convindrà tractar de diferent forma:

11.1.- Caràcters normals: són tots aquells que intervenen en l'ordenació de les paraules.

11.2.- Caràcters diacrítics: són aquells que distingeixen entre homònims, i ordenen, com a lemes diferents, les paraules que els porten, immediatament després de les mateixes paraules que no els porten. Normalment, ho és qualsevol tipus d'accent, però també pot ser-ho el guió o qualsevol altra caràcter.

11.3.- Caràcters coixí: són aquells que apareixen amb la paraula però que no afecten en res la seva ordenació. De tal manera que una paraula amb caràcters coixí, apareix inclosa en el mateix lema que la paraula sense aquests caràcters. Aquí podríem incloure-hi, per exemple, els caràcters { i } quan aquests fossin utilitzats per senyalar caràcters d'abreviatures desfetes i volguéssim que no fossin una entrada nova però que es reflectissin en els lemes que hi ha sota de cada entrada.

11.4.- Caràcters ignorats: són aquells que no afecten l'ordenació i que ni tan sols apareixen a les llistes. Aquí, hi podríem incloure els dos caràcters anteriors, { i }, en el cas que volguéssim que les paraules amb aquests caràcters no fossin ni entrades noves ni que aquests caràcters es reflectissin en els lemes que apareixen a sota de cada entrada.

11.5.- Caràcters de puntuació: són aquells que només marquen els límits de la paraula. D'entre els quals hi ha l'espai.