

Ma. Antònia Martí i Antonín.

Tesi Doctoral

"Processament informàtic del llenguatge natural: un
sistema d'anàlisi morfològica per ordinador"

Departament de Filologia Romànica

Facultat de Filologia de la Universitat de Barcelona

Barcelona 1988

Directora:

Dra. Teresa Cabré i Castellví

Tutor:

Dr. Jesús Tusón

IV.- OBJECTIUS I BASES METODOLÒGIQUES

Atès que la majoria d'investigacions en el terreny de la Lingüística Computacional s'han desenvolupat sobre la llengua anglesa i que, com se sap, aquesta llengua presenta molt poques variants flexives per a cada arrel, i a la vegada moltes irregularitats des del punt de vista de la morfologia, la majoria de les investigacions en aquest terreny es basen en un diccionari que inclou una a una totes les variants flexives d'un mot. Això explica la manca d'interès per un problema que en d'altres llengües és bàsic.

La Morfologia Computacional s'ha desenvolupat fonamentalment quan s'han iniciat investigacions sobre llengües com ara l'alemany, el francès, el català, el castellà, l'italià, l'àrab, el finlandès, el rus etc. amb una forta capacitat flexiva i en les quals el reconeixement de les formes no es pot resoldre mitjançant el simple recurs a un diccionari.

IV.1.- Objectius de l'analitzador morfològic

L'objectiu d'aquesta tesi és la construcció d'un analitzador morfològic de la llengua catalana que reconegui qualsevol variant flexiva o derivativa d'un mot, mitjançant la combinatòria de segments de formes. A grans trets, es tracta de construir un analitzador basat en un autòmat les regles del qual han de permetre les combinacions possibles d'arrels i sufixos, tant flexius com derivatius, de la llengua. Com ja hem indicat en el capítol anterior, els altres components de l'analitzador són els models, els diccionaris d'arrels i sufixos i els atributs morfològics. L'objectiu de l'analitzador és disposar del màxim de vocabulari virtual en un espai de memòria mínim.

Els avantatges d'aquest analitzador respecte als altres sistemes que hem vist a II.3 són:

- que ocupa un espai de memòria relativament petit respecte al volum de vocabulari que permet reconèixer;

- que el manteniment del diccionari és una tasca senzilla: donar d'alta o de baixa tota una família de paraules es resol en pocs segons;

- que pot donar tota mena d'informació, lèxica i morfològica, a les unitats de l'anàlisi;

- que disposa d'uns programes de manteniment que simplifiquen el treball del lingüista.

- que disposa d'un mecanisme generador que facilita el control del vocabulari que s'analitza.

IV.2.- Bases metodològiques

IV.2.1- Supòsits i característiques generals

L'analitzador, com a instrument d'anàlisi lingüística, es va dissenyar des del supòsit que l'anàlisi dels mots, o si es vol el reconeixement de les formes, es produeix en el sentit d'esquerra-dreta. És per això que vam considerar adient la construcció d'un analitzador basat en un autòmat markovià, que permetia el reconeixement/anàlisi de les formes en aquest sentit.

El segon supòsit en què es basa és la possibilitat d'establir relacions de distribució entre als elements integrants del mot, que en lingüística s'anomenen morfemes, i que aquí, per evitar confusions, anomenem 'segments' (1).

(1) Per al tema de la terminologia, vegeu IV.2.2.1.

Una de les limitacions més fortes amb què es troba aquest model, altrament com la majoria de sistemes d'anàlisi morfològica automatitzada, és que l'entrada de textos és gràfica, la qual cosa emmascara en certa manera la realitat fònica del llenguatge.

Aquest ha estat el motiu fonamental pel qual hem creat una nomenclatura pròpia que permet designar les unitats definides per a la construcció de l'analitzador, atès que no coincideixen amb les unitats amb què es treballa en lingüística en un sentit estricte.

Una altra de les característiques remarcables de l'analitzador és que permet diverses estratègies per a l'anàlisi de les formes a tots els nivells:

- en l'assignació d'atributs,
- en la segmentació de les formes,
- en la definició dels models i;
- en el disseny de l'autòmat.

L'elecció d'una determinada solució d'entre les diverses possibles l'hem feta diverses vegades en funció dels avantatges que hem anat descobrint en la construcció de l'analitzador.

La creació de l'analitzador ha suposat les fases de treball següents:

a)- Definir els criteris de segmentació de les formes flexives i donar-los una nomenclatura.

b)- Agrupar els segments en models.

c)- Definir els criteris de classificació dels models.

d)- Assignar atributs a les formes resultants de la segmentació i als models.

e)- Determinar quin tipus d'informació s'ha de donar i el nivell d'assignació de la informació: als segments (nivell específic) o als models (nivell genèric).

f)- Construir l'autòmat:

- definir les estratègies de combinatòria dels models,
- definir els estats de l'autòmat
- definir les regles.

L'analitzador duu associat un mecanisme, el generador, que permet de generar, a partir de cada arrel entrada al diccionari, totes les formes que es puguin analitzar a partir d'aquesta arrel. La capacitat de generació ve, doncs, predeterminada per l'àmbit de vocabulari que l'analitzador resol.

Aquest generador té una doble funció:

a)- Permet controlar l'abast de l'analitzador;

b)- pot utilitzar-se en determinades aplicacions de la Lingüística Computacional, com ara la traducció automàtica.

Atès que hem concebut el generador només com un mecanisme associat a l'analitzador morfològic, presentarem el tema des de la perspectiva de l'anàlisi, i no de la generació.

A continuació exposem de manera detallada la metodologia que hem emprat en la segmentació de les formes, la definició dels models i la construcció de l'autòmat.

IV.2.2.- Segmentació de les Formes

Ateses les característiques del sistema informàtic que utilitzem, l'estratègia que hem seguit per a la segmentació de les formes de la llengua és la següent:

a)- Segmentació, com hem dit, d'esquerra a dreta, seguint criteris objectius formulats explícitament.

b)- Selecció dels segments rellevants per a l'analitzador d'entre tots aquells que es podrien definir.

Els diferents tipus d'unitats resultants d'aquesta segmentació no coincideixen amb les unitats definides en lingüística:

1)- En primer lloc perquè tractem el llenguatge des d'un punt de vista gràfic, mentre que des de la perspectiva lingüística es treballa a partir de les representacions fonològiques, essent la seva ortografia una qüestió accidental. Així, amb el tractament que realitzem, 'avanç-' i 'avanc-' són dues entrades de diccionari necessàries per a l'anàlisi de les formes del verb 'avançar'; en canvi lingüísticament representen una sola unitat, amb una sola representació. El mateix passa amb 'menj-/meng-', 'roj-/rog-', etc.

2)- En segon lloc, perquè pels objectius definits i atès que es treballa amb una representació gràfica de les formes, no ens ha semblat adequat reflectir mitjançant diversos passos de l'analitzador els morfemes de què es compon un mot; primerament lloc perquè una representació ortogràfica no sempre ho permet i, segonament, perquè complicaria exageradament l'autòmat.

La forma 'estimaven' permet exemplificar ambdós tractaments. Des d'una perspectiva lingüística està

constituïda pels següents elements (1):

estim- (arrel)
-a- (vocal temàtica)
-v- (característica temporal)
-n (flexió de nombre i persona)

En canvi, aquest analitzador la pot resoldre segmentant-la en 'estim-' i '-aven', de manera que l'anàlisi requereix tan sols dues regles, enfront de la solució anterior que n'exigia quatre.

Amb tot, a grans trets es poden establir paral·lelismes entre els segments que definim i els morfemes de la llengua, però, per tot el que acabem d'exposar, considerem necessari definir les unitats amb què treballem segons uns criteris que imposa el medi en què ens movem i els donem una nomenclatura pròpia per tal de no produir confusió.

IV.2.2.1.- Terminologia

La L.C. sol utilitzar sense cap mena de reticència la terminologia pròpia de la lingüística per referir-se a les seves pròpies unitats d'anàlisi. Les unitats de la L.C., però, no coincideixen necessàriament amb les de la lingüística. Vist que aquesta utilització equívoca

(1) J. Mascaró Morfologia, 1986

podria provocar una certa confusió i, també, la disconformitat per part dels lingüistes en veure emprada impròpiament la seva terminologia per d'altres disciplines, ens ha semblat útil d'emprar una terminologia específica, la qual definim a continuació:

- Segment inicial (SI): Part invariable de la paraula, dotada de categoria morfològica. Aquesta categoria s'expressa amb l'atribut CAT i els valors NOM, VERB, ADJ, etc.:

menj- CAT= VERB
 genoll- CAT= NOM
 bonic- CAT= ADJ etc.

Els models als quals pertanyen són els únics que surten de l'estat inicial de l'autòmat.

La categoria morfològica dels SI pot canviar si es combinen amb un segment intermedi (SM):

| | | |
|----------|---------|---------------|
| CAT=VERB | CAT=NOM | CAT=NOM |
| menj- | -ador- | -o = menjador |
| (SI) | (SM) | (SF) |

Quan un SI es combina amb un segment final (SF), la forma resultant és un mot de la llengua:

| | | | |
|-------|------|--------|------|
| menj- | -o | bonic- | -a |
| (SI) | (SF) | (SI) | (SF) |

Segment Inicial Bàsic: Forma bàsica d'un SI.

Segment Inicial Germà: Cada una de les variants que pot adoptar un SI bàsic:

| | | |
|----------|------------|----------|
| SI | | |
| passèig- | (SI bàsic) | CAT=NOM |
| passèj- | (SI germà) | CAT=VERB |
| passèg- | (SI germà) | CAT=VERB |

Els SI germans es poden vincular internament al SI definit com a bàsic i així, de la mateixa manera que a partir de 'salt-' es poden generar totes les formes del verb 'saltar', a partir de 'caç-', de 'preg-' o de 'menj-' es generaran totes les formes flexives dels verbs 'caçar', 'pregar' i 'menjar'.

A més de la flexió, la definició de SI germans permet també mantenir la cohesió entre els diferents SI integrants de famílies de paraules (1).

Segment Inicial de Base o simplement Bases (SIB):

Són SI que no poden combinar-se directament amb SF: fil- ('filòleg', 'filòloga', 'filologia') o bé 'abs-' ('absent', 'absència'). Cal tractar-los com a SI perquè es combinen amb SM i també amb grups de SM:

(1) Vegeu-ne l'explicació a IV.2.5.4.

| BASES -----> | GRUP DE SM | BASES---- | GRUP DE SM |
|--------------|------------|-----------|------------|
| fil- | -òleg | clau- | -ent |
| espele- | -òlog-a | pres- | -ència-a |
| crim- | -olog-ia | dem- | |
| antrop- | | etc. | |
| crimin- | | | |
| etc. | | | |

La definició de SIB permet la resolució computacional de mots manllevats del grec i del llatí, formats per composició, i de pseudo-derivats el SI dels quals no correspon a un SI en català actual.

Els models als quals pertanyen surten de l'estat inicial de l'autòmat com la resta dels SI.

- Segment Intermedi (SM): Segment que es combina amb un SI al seu davant i un segment final (SF) o un altre SM al seu darrera. Serveix per formar una nova paraula. Sol provocar sovint un canvi de categoria del SI, raó per la qual acostuma a ser caracteritzat per l'atribut CAT amb el valor categorial del canvi que produeix:

menj- ador(SM)-0 blanc- or(SM)-0
racional- itz(SM)-ar inform- ació(SM)-0

Corresponen, a grans trets, als elements que en lingüística s'anomena 'sufixos'.

- Segment Quasi-Finals (SQF): Segment que s'uneix directament a un SI amb el qual forma un conjunt que

es seguit d'un SF o d'un altre SQF. Aquests segments no provoquen cap canvi de categoria i només es donen a la conjugació verbal:

resol-gu(SQF)-i estudi-ar(SQF)-é
apare-ix(SQF)-er(SQF)-é
na-squ(SQF)-i corr-egu(SQF)-és
pon-dr(SQF)-à

- Segment Infix (SIN): Segment situat darrera d'un SI que no es pot unir a un SF si no és a través d'un SM:

got-eil(SIN)-in(SM)-ar
cafe-t(SIN)-er(SM)-a
tomb-ass(SIN)-ej(SM)-ar

Aquests segments, incorporats als SI, constitueixen noves entrades al diccionari de SI:

gotellin- , tombass- , cafet-

- Segment Final (SF): Segment que pot combinar-se al seu davant amb qualsevol altre dels segments definits excepte els SIN, i amb el qual es tanca l'anàlisi del mot.

De vegades es comporten com a SM perquè produeixen canvis de categoria. En aquest cas, han de portar l'atribut CAT amb el valor corresponent. És el cas dels verbs post-nominals: per ex., 'tomb-' és un SI qualificat com a nominal, però que en combinar-se amb els SF de la flexió verbal canvia de categoria morfològica i es converteix en verb, per tant els SF de la flexió porten l'atribut CAT=VERB:

<VINF> (1)
 CAT=VERB
 FNP=INF (1)

-ar

<IMPF> (1)
 CAT=VERB
 TEMP=IMPF (1)
 MODE=IND (1)

-ava PERS=1/3 (1)
 NBRE=SG (1)

<NMBS>
 CAT=NOM
 GEN=MASC

tomb-

Anàlisi de les formes:

tombar CAT=VERB
 FNP=PART

tombava CAT=VERB
 TEMP=IMPF
 MODE=IND
 PERS=1/3
 NBRE=SG

(1) Els models <VINF> i <IMPF> corresponen als models dels SF de l'infinitiu i de l'imperfet. L'atribut FNP correspon a les formes no personals del verb i el valor INF, a l'infinitiu. Els atributs TEMP i MODE corresponen a 'temps' i 'mode' i els valors IMPF i IND a 'imperfet' i 'indicatiu' respectivament. PERS i NBRE corresponen a 'persona' i 'nombre'. El valor 1/3 indica que es tracta d'una primera o tercera persones i SG indica 'singular'.

S'assumeix l'existència del SF '-0', el caràcter del blanc. Aquest SF pot portar, com els altres, atributs morfològics.

Els SF no poden anar seguits de cap altre tipus de segment i no són acumulatius.

IV.2.2.2.- Criteris per a la segmentació de les formes.

Proposem el seguiment d'un exemple i de les seves possibles solucions per tal de justificar els criteris que hem seguit en la segmentació de les formes.

Hem triat algunes formes verbals de la segona conjugació perquè, a causa de la seva complexitat, permeten mostrar les diferents possibilitats que planteja la seva segmentació. Prendrem en consideració el futur, l'imperfet de l'indicatiu i el present de subjuntiu dels verbs 'perdre' i 'absoldre'.

A)- Una primera segmentació possible seria la següent:

| | PERDRE | ABSOLDRE | |
|-----------|----------|------------|-------|
| FUTUR | perd-r-é | absol-dr-é | |
| PRES.SUB. | perd-i | absol-gu-i | Fig.1 |
| IMPERF | perd-ia | absol-ia | |

Observem que el primer segment delimitat es manté igual al llarg de tota la flexió (SI). Observem també que els últims segments '-é', '-i', '-ia' (SF) es mantenen invariables en totes les formes i que, a més, són els mateixos per a gairebé tots els verbs. El que canvia és la presència/absència de segments quasi-finals (SQF): '-r-', '-dr-', '-gu-', etc.

Hi ha d'altres verbs que tenen els mateixos o d'altres SQF però que mantenen idèntics els SF:

| | FUTUR | PRES.SUBJ. | IMPERF. |
|-----------------|---------------|------------|-------------|
| Moure | mou-r-é | mo-gu-i | mov-ia |
| Córrer | corr-er-é | corr-i | corr-ia |
| Aparèixer | apare-ix-er-é | apare-gu-i | apare-ix-ia |
| Prendre etc. | pren-dr-é | pren-gu-i | pren-ia |

Fig.2

Sobre la base d'aquesta segmentació, s'agrupen els SI en models segons els SQF i SF que accepten. Per a l'exemple de la Fig.1 serien:

| | |
|----------|------------|
| <PERDRE> | <ABSOLDRE> |
| TV=PERD | TV=ABSOL |
| perd- | absol- |

Es creen models per a cada un dels SQF:

<R>
-r-

<DR>
-dr-

<GU>
-gu-

! es defineixen els models de SF que, en el cas de la conjugació, s'organitzen en funció dels temps verbals:

| <FUT> | <PRSUB> | <IMPF> |
|----------------------|-----------------------|-----------------------|
| TEMP=FUT MODE=IND | TEMP=PRES MODE=SUB | TEMP=IMPF MODE=IND |
| -é -às ... | -i -is ... | -ia -ies ... |

A continuació es crea una estructura d'autòmat en que tots els models de SI van a parar a un mateix estat, RAR2. D'aquest estat en surten les regles dels models de SQF, validades segons el model de SI, així com els models de SF que no es combinen amb SQF. Totes les regles de SQF van a parar a un estat, RFUT o RSUB, d'on surten els models de SF. D'aquesta manera les regles dels SF són vàlides per a tots els SI, sigui quin sigui el seu SQF:

| Regles | Segments reconeguts |
|--|---------------------------|
| R1 START <PERDRE> RAR2 | 'perd-' |
| R2 START <ABSOLDRE> RAR2 | 'absol-' |
| R3 RAR2 <R> RFUT condició: TV=PERD | 'perdr-' |
| R4 RAR2 <DR> RFUT condició: TV=ABSOL | 'absoldr-' |
| R5 RAR2 <GU> RSUB condició: TV=ABSOL | 'absolgu-' |
| R6 RAR2 <IMPF> RV condició: TV=PERD, TV=ABSOL | 'perdia', 'absolia'... |

| | | | | |
|----|------|-------------------|----|----------------------------|
| R7 | RAR2 | <PRESUB> | RV | 'perdi'... |
| | | condició: TV=PERD | | |
| R8 | RFUT | <FUT> | RV | 'perdré', 'absoldré'... |
| R9 | RGUB | <PRESUB> | RV | 'absolgui'... |

B)- Una segona possibilitat de segmentar les formes verbals presentades és la de considerar cada variant com a un SI específic:

| | |
|--------|----------|
| PERDRE | ABSOLDRE |
| perd- | absol- |
| perdr- | absoldr- |
| | absolgu- |

Aquesta solució té l'inconvenient d'incrementar de manera considerable el nombre d'entrades del diccionari de SI, però té l'avantatge de simplificar l'estructura de l'analitzador, ja que només són necessàries dues regles per a l'anàlisi de cada forma.

Per avaluar les dues solucions proposades s'ha de tenir en compte que:

a)- La part més costosa en despesa de temps és l'accés al diccionari;

b)- Les regles que resolen casos semblants al que hem presentat a l'exemple de la Fig.1 no impliquen una complicació excessiva de l'analitzador i tenen un

rendiment força bo (1).

D'aquestes consideracions se'n desprèn que seria millor la primera solució. Així, seguint aquesta filosofia, hem resolt tots els canvis dels SI que consisteixen en l'afegiment de nous caràcters, definint aquests nous caràcters com a SQF, assignant-los un model i donant les regles de l'autòmat que els permeten combinar-se amb els SI que els accepten. L'única exigència que s'ha de tenir present és que aquests SQF han de tenir un mínim rendiment distribucional.

El rendiment distribucional d'un segment és determinat per dos factors:

1)- La combinatòria efectiva amb d'altres segments. En aquest cas cal decidir arbitràriament el nombre que es considera acceptable.

2)- La possibilitat de combinar-se amb segments ja existents per formar noves paraules, és a dir, la seva vitalitat actual.

Els supòsits que hem aplicat a la flexió, són els mateixos que hem aplicat als derivats: només hem

(1) El rendiment d'una regla ve determinat pel nombre de casos que resol.

definit aquells SM que tenen un cert rendiment distribucional, i els SM que presenten variants gràfiques, que normalment només servixen per resoldre un nombre reduït de formes, no els hem considerat com a SM i s'ha definit un nou SI que inclou el SM.

IV.2.3.- Els Models

Un cop definides les estratègies de segmentació de les formes, cal agrupar-les en models per tal que puguin ser reconegudes mitjançant les regles de l'autòmat.

En general, hem organitzat els models de SI segons els models de SF que accepten, i hem agrupat els SF en funció dels paradigmes de gènere i nombre, en el cas dels noms i adjectius, i en funció de la conjugació, en el cas dels verbs.

Ha estat necessari subagrupar els models dels SF segons les exigències dels SI. Per a la creació dels models de SF s'han de tenir en compte els canvis/alternances que sovint presenten els SI i, a la vegada, els paradigmes de gènere i nombre o de la conjugació.

Vegem un exemple, que pot aclarir el procés, de com es pot resoldre l'anàlisi de les formes verbals de la

primera conjugació que presenten canvis als SI segons els caràcters dels SF que els segueixen:

| | |
|-------|--------------------|
| menj- | -o, -a, |
| meny- | -es, -em, -eu, -en |

S'han de definir dos models de SI per als verbs de la primera conjugació amb aquest tipus de canvi :

| | |
|----------|-----------|
| <VARG> | <VARGU> |
| TV=ARG | TV=ARGU |
| menj- | meng- |
| caç- | cac- |
| naufrag- | naufragu- |
| etc. | etc. |

Fig.3

Com que el SI 'menj-' només accepta uns determinats SF i el SI 'meng-' uns altres, cal separar-los en models diferents de manera que es pugui donar al model l'atribut per validar la flexió. L'atribut que servirà per validar les regles de la flexió és TV i té diferents valors segons el model de SI que es tracta.

En funció de les alternances en els SI cal definir els models dels SF, en aquest cas del present de l'indicatiu, de la manera següent:

| <IP1> | <IP2> |
|-------------------------|--------------------------|
| TEMP=PRES MODE=IND | TEMP=PRES MODE=IND |
| -o PERS=1 NBRE=SG | -es PERS=2 NBRE=SG |
| -a PERS=2 NBRE=PL | -es PERS=1 NBRE=PL |
| | -eu PERS=2 NBRE=PL |
| | -en PERS=3 NBRE=PL |

Fig.4

Són necessaris dos models de present perquè el SI 'meng-' només es combina amb els SF de <IP2> i el SI 'menj-', amb els SF de <IP1>. Els SI com 'salt-', que accepten les formes de <IP1> i de <IP2>, estaran validats per a les dues regles de SF. Si <VAR> és el model de 'salt-' i TV=AR l'atribut que expressa les validacions de la conjugació, l'exemple que hem presentat es resol amb les regles:

- R1 START <VAR> RAR1
- R2 START <VARG> RAR1
- R3 START <VARGU> RAR1
- R4 RAR1 <IP1> RV
condició: TV=AR, TV=ARG
- R5 RAR1 <IP2> RV
condició: TV=AR, TV=ARGU

Així, els verbs com 'salt-ar' es combinaran amb tots dos models de SF i els verbs com 'menjar' tindran un SI ('menj-') que només es combinarà amb <IP1> i un altre SI ('meng-') que només es combinarà amb <IP2>.

Aquesta estratègia d'agrupament de SI i de SF que hem utilitzat per a l'anàlisi dels verbs, també ens ha servit per a l'anàlisi de noms i adjectius. Vegem-ne un cas:

| | |
|---------|-------------|
| <NFAES> | <FAES> |
| CAT=NOM | GEN=FEM |
| TG=FAES | |
| | -a NBRE=SG |
| prems- | |
| faix- | -es NBRE=PL |

que intervenen a les regles:

| | | | |
|----|-------|-------------------|-----|
| R1 | START | <NFAES> | RNA |
| R2 | RNA | <FAES> | RGN |
| | | condició: TG=FAES | |

<NFAES> és un model de SI nominals que es combina amb el model de SF <FAES> segons les regles R1 i R2. Hi ha SI que també es combinen amb <FAES>, però que presenten alternances al SI, com ara 'comarc-' / 'comarqu-', la qual cosa exigeix la creació de dos nous models de SI i de SF :

Models de SI**Models de SF**

<NFA>
 CAT=NDM
 TG=FA

<NFES>
 CAT=NGM
 TG=FES

<FA> <FES>
 GEN=FEM GEN=FEM
 NBRE=SG NBRE=PL

comarc-

comarqu-

-a

-es

Així, es podria prescindir del model de segments finals <FAES>, ja que l'anàlisi de les formes de <NFAES> es resoldria validant amb TG=FAES les regles dels models de SF <FA> i <FES>.

IV.2.3.1.- Els models "BASES".

Tots els SI que no es combinen directament amb un SF es troben agrupats en uns models que anomenem BASES1 i BASES2. La informació referent a la categoria morfològica i als atributs per validar les regles dels SM correspon a cada SI en particular.

IV.2.4.- La informació morfològica

A mesura que avança l'anàlisi d'una forma, l'analitzador va recollint tota la informació que troba. Aquesta informació s'expressa mitjançant atributs que poden tenir diferents valors fins a un màxim de tres-cents.

Quan, al llarg d'aquest procés, un mateix atribut apareix diverses vegades, sempre pren el valor de l'última ocurrència que ha trobat.

Els atributs morfològics tenen la forma ATRIBUT=VALOR, p. e.: CAT=NOM, CAT=VERB, CAT=N/V, NBRE=SG, NBRE=PL, PERS=1, PERS=3, etc.

La informació associada pot ser de diferents tipus:

a)- Estrictament morfològica (categoria, gènere, nombre, persona, etc.).

b)- Necessària per al funcionament de l'autòmat: atributs que, assignats als models o als segments, serveixen per validar les regles de l'autòmat:

| | |
|---------------|----------|
| <VAR> | <SNADOR> |
| CAT=VERB | CAT=N/A |
| | TG=MFBA |
| salt- SN=ADOR | -ador |

| | | | |
|----|-------|----------|-----------------------|
| R1 | START | <VAR> | RAR1 |
| R2 | RAR1 | <SNADOR> | RNA condició: SN=ADOR |

A partir de l'estat RAR1, per passar per la regla R2 cal complir la condició SN=ADOR.

c)- Altres tipus d'informació.

IV.2.4.1.- Nivells d'assignació d'informació

Els atributs morfològics portadors de la informació sobre la categoria, el gènere, el nombre, la persona, etc. poden assignar-se a dos nivells diferents: als models o als segments.

Quan s'assigna un atribut a un model, aquest atribut afecta tots els segments que pertanyen al model. Quan l'atribut s'assigna a un segment només afecta el segment.

Com que els segments inicials (SI) estan agrupats segons un determinat criteri morfològic, que normalment és el seu comportament flexiu (paradigma de conjugació per als verbs i paradigma de gènere i nombre per als noms i adjectius), l'atribut que servirà per validar les regles de flexió s'associa al model, perquè és generalitzable a tots els segments:

<VAR>

(Model de verbs regulars de la 1a. conj.)

CAT= VERB

TV = AR

(Atribut de validació de flexió verbal)

cant-

salt-, etc.

<NMBS>

(Model de noms masculins: 'fruit-0,-s')

CAT=NOM

TG=MBS

(Atribut per validar gènere i nombre)

soldat-

combat-

Contràriament a la flexió, els derivats responen a una selecció prèvia que es dona a partir del SI; així, els atributs que serveixen per validar les regles dels SM, formadors de mots derivats, s'han de donar particularment a cada SI, perquè tot i ser mots de la mateixa classe morfològica poden tenir un comportament derivatiu completament diferent.

Vegeu el que acabo d'exposar amb un exemple:

<VAR>

CAT= VERB
TV = AR

| | | | |
|-------|----------|--------|----------|
| acab- | SN=AMENT | abund- | SN=ANCIA |
| | SN=ALLES | | SA=ANT |
| | SN=ANÇA | | SN=OR |
| | SA=ABLE | | SA=C5 |

Els atributs SN i SA corresponen a 'segment intermedi nominalitzador' i 'segment intermedi adjectivador', respectivament. Els valors corresponen al SM que es vol validar.

Des de la perspectiva de l'anàlisi, aquesta solució representa un control estricte dels derivats, de manera que només sigui possible l'anàlisi de les formes considerades com a correctes. Des de la perspectiva de la generació, s'evita la hipergeneració de formes a partir d'un SI.

IV.2.4.2.- Assignació de la categoria morfològica

Hem preferit utilitzar el terme 'categoria morfològica' en lloc del més habitual de 'categoria sintàctica' perquè l'analitzador és estrictament independent del context, de manera que l'assignació de la categoria a un mot està determinada exclusivament per la seva forma. Quan es produeixen casos d'ambigüitat l'analitzador en dóna totes les interpretacions possibles.

La categoria morfològica dels mots es pot assignar als SI, als SM o bé als segments finals. La majoria de SI no tenen la categoria definida fins que no s'han combinat amb un SM o un SF i, consegüentment, s'hauria d'assignar l'atribut de la categoria al model d'aquests segments. Això no obstant, considerem que hi ha dues bones raons per assignar categoria morfològica als SI:

a)- perquè és un fet que cada SI no accepta combinar-se amb qualsevol SF;

b)- segonament, perquè l'organització en models que presentem descansa sobre la base que hi ha SI nominals, adjectius i verbals.

Els models de SI estan organitzats en funció de la categoria morfològica i del model de flexió corresponent:

| | | | |
|----------|---------|----------|---------|
| <VAR> | <NMBS> | <ADVLL> | <AMFBS> |
| CAT=VERB | CAT=NOM | CAT=ADV | CAT=ADJ |
| pint- | fruit- | TAD=LLOC | legal- |
| | | aquí- | |

Des de la perspectiva del procés d'anàlisi, la categoria d'una forma serà la que té assignada el seu SI, si no és que en un dels passos de l'autòmat es combina amb un SM o amb un SF que tingui l'atribut CAT amb un valor diferent:

| | | |
|----------|----------|--------------------|
| <VAR> | <SMANT> | |
| CAT=VERB | CAT=NOM | |
| cant- | -ant | = cantant, CAT=NOM |
| tomb- | + - | = tombant, CAT=NOM |

Els SI poden canviar de categoria sense necessitat de combinar-se amb un SM: es tracta dels SI que accepten SF corresponents a més d'un model de flexió. Per exemple, a partir del SI 'tomb-' s'obtenen les formes nominals 'tomb, tombs' i les formes verbals de 'tombar'. En aquests casos, els models de SF han de dur l'atribut de la categoria, altrament la informació morfològica de la forma analitzada seria incorrecta:

| | |
|---------|-----------------------|
| <NMBS> | <SF 1a. CONJ> |
| CAT=NOM | CAT=VERB |
| tomb- | -ar, -ava, -ara, etc. |

Malgrat que el SI té l'atribut CAT amb el valor NOM, en combinar-se amb els SF de la primera conjugació preval l'últim valor de l'atribut, en aquest cas VERB.

Hem establert les següents categories morfològiques:

CAT=NOM : noms, comuns i propis

CAT=VERB : verbs

CAT=ADJ : adjectius

CAT=N/A : nom o adjectiu

CAT=ADV : adverbis

CAT=PRC : pronoms

CAT=DET : determinants

CAT=D/PR : determinant o pronom

Les categories CAT=N/A (nom/adjectiu) i CAT=D/PR (determinant/pronom) corresponen a aquelles formes que, per tractar-se d'un analitzador morfològic independent del context, són categorialment ambigües:

'Corredor' CAT=N/A, GEN=MASC, NBRE=SG

'corredora' CAT=N/A, GEN=FEM, NBRE=SG

'aquest' CAT=D/PR GEN=MASC, NBRE=SG

Es pot establir una subclassificació de les categories per tal de donar més informació en l'anàlisi de cada una les formes. A continuació presentem una hipotètica subclassificació per tal de mostrar una

manera possible de jerarquitzar la informació lèxica:

CAT=NOM

TNOM= PERS, per als noms propis de persona
(Salvador Espriu, Jordi Pujol, Felipe
González, etc...)

TNOM=LLOC, per als noms propis de lloc
(Montserrat, Barcelona, ...)

TNOM=ENT, per indicar entitats, partits polítics,
organismes, etc. com ara ONU, CEE, PSOE, ...

TNOM=PAIS, per als noms propis de països (França,
Anglaterra, Estats Units, etc.)

CAT=ADV

TADV=LLOC, adverbis de lloc (aquí, allà, ...)

TADV=MANE, adverbis de manera (tots els adverbis
en (-ment))

TADV=TEMP, adverbis de temps (ara, demà, adés i
ara, etc.)

TADV=NEG, adverbis de negació (no, pas, etc.)

CAT=D/PR, (l'atribut T indica el tipus)

T=POSS, possessius (meu, meva, teu, teva, seu, ...)

T=DEM, demostratius (aquest, aquell, ...)

T=QI, quantificadors indefinits (tot, massa, poc,
més, etc.)...

T=QD, quantificadors definites (dos, tres, ...)

La categoria CAT=DET queda en exclusiva per a l'article, amb la subcategoria T=ART.

CAT=PRO

TPRO= PERS pronoms personals (jo, tu, ell, nosaltres, el, la, se, en, ho, hi, etc.)

TPRO= REL pronoms relatius (que, qui, què, qual, etc.)

IV.2.5.- Disseny de l'analitzador

Com ja hem dit, la finalitat de l'analitzador és resoldre l'anàlisi de les formes flexives i derivades de la llengua mitjançant la combinatòria de segments. Aquesta via de solució no pot significar una despesa excessiva d'atributs, regles i models. Hem procurat seguir un criteri que considerem fonamental: rendibilitzar el màxim cada una de les regles; per això hem tractat amb l'autòmat els aspectes més regulars de les formes i hem donat com a noves entrades de diccionari aquelles formes l'anàlisi de les quals exigiria regles específiques (1).

Respecte als derivats, podíem haver optat per entrar-los com a SI o bé resoldre'ls mitjançant regles a partir del SI primitiu. Amb la primera solució l'autòmat s'hauria simplificat moltíssim, però hauria

(1) Els sistemes d'anàlisi morfològica computacional segueixen sempre aquest criteri.

augmentat de manera important el nombre d'entrades al diccionari; en canvi, si tractavem els derivats mitjançant noves segmentacions i noves regles de l'autòmat, el diccionari es reduïa sensiblement, però el disseny de l'autòmat es complicava.

Sempre que no ha representat complicar exageradament l'autòmat hem resolt el màxim de formes mitjançant regles i hem optat per donar entrades al diccionari quan el tractament de les formes augmentava molt les regles i a més resolïa pocs casos. Un cop construït l'analitzador es pot observar que els aspectes més regulars de la llengua es resolien amb un mínim de regles i models que permeten l'anàlisi de la majoria de les formes, mentre que els aspectes més irregulars exigeixen un aparell tècnic molt més gran i en canvi resolien molts menys casos.

IV.2.5.1.- Abast de l'analitzador

L'abast de l'analitzador està determinat per la seva potència, és a dir, per la quantitat de formes que pot analitzar. La solució òptima és la que amb un mínim d'entrades (atributs, models, regles, SI, SM, SF i SQF) permet analitzar el màxim de formes correctes de la llengua.

Si es dissenya l'analitzador amb l'objectiu que analitzi només les formes acceptades per la normativa, els atributs per validar els SM s'han de donar a cada SI en particular i mai al model, ja que en aquest darrer cas es generalitza el SM per a tots els SI. Els costos de construcció de l'analitzador seran més elevats, però l'abast del corpus estarà totalment controlat i coincidirà exactament amb el de les formes considerades com a correctes per la normativa.

La solució oposada consistiria a donar els atributs per a la validació dels derivats als models i així generalitzar-los per a tots els SI. Amb aquesta generalització seria possible l'anàlisi de moltes formes incorrectes, encara que se simplificaria l'entrada de segments al diccionari.

Una darrera solució pot consistir en l'admissió d'algunes generalitzacions en la combinatòria de SI amb SM. Els avantatges d'aquesta darrera solució són diversos:

a)- S'aconsegueix un certa simplificació en l'entrada de dades de l'analitzador: la generalització de determinats SM mitjançant l'assignació de l'atribut al model evita haver de donar l'atribut a cada SI.

b)- Es preveuen formes possibles, amb la qual cosa disminueixen els costos en temps de manteniment del diccionari.

c)- Només es generalitzen aquells SM que poden donar formes possibles. No es permet l'anàlisi de formes incorrectes. Així, es rebutgen formes com : *'compt-ament-ació', *'solid-ific-or', etc., que s'acceptarien en el cas de generalitzar indiscriminadament els SM; però en canvi s'accepten formes possibles, encara que potser no gaire probables com són: 'empassable' i 'trepitjable'.

IV.2.5.2.- Estratègies generals d'anàlisi

A més de les qüestions de tipus general que acabem d'exposar per al disseny de l'autòmat, en concret ens hem guiat pels següents objectius:

1)- Reduir al màxim el diccionari de SI sempre que ho en resultés una estructura de l'autòmat excessivament complicada.

Hem resolt com a entrades de diccionari:

a)- aquelles formes que representen una variant gràfica d'un segment:

'viu-/viv-a', 'llop/llob-a', '-ació/-ició',
'-ador/-edor/--idor', ...

b)- els SI derivats que normalment resulten de la unió d'un SI amb un SIN, o bé d'un prefix amb un SI:

'llibre-e' / 'llibret-en' ...

'tràfec' / 'atrafec-ar' ...

c)- les formes úniques, com ara "vull" del verb 'voler', 'duu' del verb 'dur' les formes 'sóc' i moltes d'altres del verb 'ser' etc.

Mentre que les formes de (a) i (c) no es poden solucionar mitjançant regles de l'autòmat i s'han de tractar com a noves entrades al diccionari, les formes de (b) podrien resoldre's mitjançant regles. En aquest cas, però, la relació entre una regla i el nombre de casos que resoldria seria molt baixa, fet que ha determinat el rebuig d'aquesta solució.

2)- Fer un disseny de l'autòmat el més senzill possible per tal de facilitar la comprensió de la seva gestió, l'ampliació i/o la restricció del vocabulari.

Aquest darrer punt s'ha concretat en les accions següents:

a)- Hem donat un valor anemotècnic als estats, als atributs i als models per tal de facilitar el seu reconeixement.

b)- Hem evitat sistemàticament la creació de bucles a les regles de l'autòmat perquè, encara que permeten reduir el nombre de regles, creen problemes de control en el reconeixement de les formes. Tot model que intervé en un bucle ha de dur negats tots els atributs que l'autòmat hi hagi pogut reconèixer anteriorment. Vegem-ne un cas:

Tenim el SI nominal 'fruit' amb els atributs TV=AR per a l'anàlisi de les formes del verb 'fruit-ar' i amb els atributs SN=ER ('fruiter', 'fruitera'), SN=ERIA ('fruiteria', 'fruiteries'), etc. El model de 'fruit-' té l'atribut TG=MBS per validar els SF de gènere i nombre. Definim la següent estructura a l'autòmat:

| | | Segments Reconeguts | |
|----|--|---------------------|-----------------------|
| R1 | START <S: NOM.> | RNA | 'fruit-' |
| R2 | RNA <SF GEN/NOMBR> condició: TG=MBS | RGN | 'fruit', 'fruits' |
| R3 | RNA <SF CONJUG> condició: TV=AR | RV | 'fruita', 'fruitan' |
| R4 | RNA <SM NOMIN.'ER'> condició: SN=ER | RNA | 'fruiter', 'fruiters' |
| R5 | RNA <SM NOM.'ERIA'> condició: SN=ERIA | RNA | 'fruiteria', ... |

Aquesta solució permet utilitzar els SF de gènere i nombre tant per als SI ('fruit-) com per als SM ('fruit-er-', 'fruit-eri-') ja que totes aquestes regles tenen RNA com a estat final i d'aquest estat surten e's models de SF de gènere i nombre.

Els segments reconeguts a l'estat RNA són 'fruit-', 'fruit-er-' i 'fruit-eri-'. El SI 'fruit-' té la validació TV=AR, de manera que tots els segments reconeguts fins a RNA accepten combinar-se amb els SF de la flexió verbal, amb la qual cosa s'analitzen com a correctes formes com ara: *'fruit-er-o' 1a pers. sing. pres. ind.; *'fruit-er-es' 2.pers. sg.pres ind.; etc. i també *'fruit-eri-o' 1a. pers. sing. pres. ind; etc.

Per tal d'evitar aquestes anàlisis incorrectes hi ha dues solucions possibles:

1)- Invalidar al SM tots els atributs que l'autòmat hagi pogut trobar anteriorment.

Aquesta, encara que sembla una solució senzilla, no ho és tant si tenim en compte que s'han de preveure tots els possibles atributs que poden tenir els SI que accepten un determinat SM. Actuar per negació sempre implica el risc d'oblidar un atribut, i això, quan es treballa amb un nombre elevat de regles, models i atributs dificulta el govern de l'analitzador.

2)- La segona solució consisteix a crear un estat específic com a estat final de les regles dels SM, per exemple RNA1, d'on surten les regles de SF de gènere i nombre:

| | | Segments Reconeguts |
|----|---|-------------------------------|
| R1 | START <SI NOM.> | RNA 'fruit-' |
| R2 | RNA <SF GEN/NOMBR> condició: TG=MBS | RGN 'fruit', 'fruits' |
| R3 | RNA <SF CONJUG> condició: TV=AR | RV 'fruita', 'fruiten' |
| R4 | RNA <SM NOM.'ER'> condició: SN=ER | RNA1 'fruiter', 'fruiters' |
| R5 | RNA <SM NOM.'ERIA'> condició: SN=ERIA | RNA1 'fruiteria', ... |
| R6 | RNA1 <SF GEN/ NOMBR> condició: TG=FAES | RGN |

Aquesta solució incrementa el nombre de regles, però es guanya en claredat de disseny i control.

c)- Hem creat subconjunts de regles a les quals s'arrriba des de diferents punts de l'autòmat. La segona conjugació verbal pot servir per il·lustrar aquest aspecte.

RAR2 és l'estat final de les regles de models de SI de la segona conjugació. Aquesta conjugació presenta els mateixos SF per al futur i per al condicional en gairebé tots els paradigmes de verbs, que tan sols es diferencien pels diferents SQF que hi ha entre els SI

i els SF : '-dr-' (model <SQFDR>) per a verbs com 'resoldre' (SI='resol-'), '-er-' (model <SQFER>) per a verbs com 'córrer' (SI=corr-), '-r-' (model <SQFR>) per a verbs com 'perdre' (SI='perd-'). Les regles d'anàlisi serien:

| | | | |
|----|-------|------------------|------|
| R1 | START | <SI 2A. CONJ> | RAR2 |
| R2 | RAR2 | <SQFR> | RFC |
| | | condició: SQF=R | |
| R3 | RAR2 | <SQFDR> | RFC |
| | | condició: SQF=DR | |
| R4 | RAR2 | <SQFER> | RFC |
| | | condició: SQF=ER | |
| R5 | RFC | <FUTUR> | F |
| R6 | RFC | <CONDIC> | F |

Així, les regles 5 i 6 són generalitzables per a tots els verbs de la segona conjugació sigui quin sigui el SQF del futur i condicional.

IV.2.5.3.- Esquema de l'analitzador

A continuació presentarem de manera molt esquemàtica el disseny general de l'autòmat que és a la base de l'analitzador per tal d'oferir un marc de lectura per als capítols següents.

Hem agrupat els SI de la conjugació verbal en tres grans grups de models corresponents a les tres conjugacions. Tots aquests models tenen START com a estat inicial i RAR1, RAR2 i RAR3 (Reconeix ARrels de

la primera, segon i tercera conjugació)
respectivament com a estats finals de les regles:

| | | | |
|----|-------|----------------|------|
| R1 | START | <SI 1A. CONJ.> | RAR1 |
| R2 | START | <SI 2A. CONJ.> | RAR2 |
| R3 | START | <SI 3A. CONJ.> | RAR3 |

Els models dels SF verbals tenen RAR1, RAR2, RAR3 com a estats inicials i RV (Reconeix el Verb) com a estat final:

| | | | |
|----|------|---------------------|----|
| R4 | RAR1 | <SF FLEXIO 1A.CONJ> | RV |
| R5 | RAR2 | <SF FLEXIO 2A.CONJ> | RV |
| R6 | RAR3 | <SF FLEXIO 3A.CONJ> | RV |

Els models de SI de noms i adjectius tenen START com a estat inicial i RNA (Reconeix Noms i Adjectius) com a estat final. Els models de SF de gènere i nombre de noms i adjectius tenen RNA com a estat inicial i RGN (Reconeix Gènere i Nombre) com a estat final:

| | | | |
|-----|-------|----------------|-----|
| R7 | START | <SI NOMS> | RNA |
| R8 | START | <SI ADJECTIUS> | RNA |
| R9 | RNA | <SF NOMS> | RGN |
| R10 | RNA | <SF ADJECTIUS> | RGN |
| R11 | RNA | <SF N/A> | F |

Quan al SF es reconeix el caràcter del 'blanc', p.e. el SF dels noms masculins com 'fruit', el model de SF té F (estat definit com a Final) com a estat final.

Els SM verbalitzadors de noms i adjectius tenen RNA com a estat inicial i RAR1 com a estat final, ja que només tractem SM verbalitzadors que donen lloc a verbs de la primera conjugació:

R11 RNA <SM VERBAL> RAR1

Els SM nominalitzadors de noms i adjectius tenen RNA com a estat inicial i RNA1 com a estat final. La funció de l'estat RNA1, com ja s'ha dit, és la d'evitar la formació d'un bucle.

L'estat RNA1 és l'estat que reconeix SI nominals i adjectius 'derivats' mitjançant un SM. Així, els SM nominalitzadors de verbs tenen RAR1, RAR2, i RAR3 com a estats inicials i RNA1 com a estat final:

R12 RAR1 <SM NOMIN/ADJ> RNA1

R13 RAR2 <SM NOMIN/ADJ> RNA1

R14 RAR3 <SM NOMIN/ADJ> RNA1

R15 RNA <SM NOMIN/ADJ> RNA1

De RNA1 surten els models de SF corresponents als segments intermedis:

R16 RNA1 <SF NOM/ADJ> RGN

Aquestes setze regles representen de manera molt general l'estructura de l'autòmat. En els capítols següents tractarem amb detall cada un d'aquests grups.

IV.2.5.4.- Famílies de paraules

Encara que el tema dels derivats el tractem més extensament al capítol VI. En aquest apartat dedicat a la metodologia presentem els recursos de què disposa l'analitzador per reconèixer els derivats i com els hem utilitzats.

El concepte de derivat porta, conseqüentment, al de família de paraules, conjunt de totes aquelles formes que deriven d'un SI comú i que mantenen entre si una relació de significat més o menys propera.

Encara que sigui de vegades arriscat, creiem que és necessari utilitzar el criteri del significat en la definició de les famílies de paraules perquè en cas contrari s'acceptarien com a correctes les derivacions:

| | |
|-----------|--------|
| inter-ior | sol-ar |
| i | edat |
| | ament |
| | ista |

que permeten l'anàlisi de formes totalment allunyades pel significat a partir d'un mateix SI. Si es procedís en aquest sentit, la gestió del diccionari, i per tant de l'analitzador, seria impossible.

El recurs al significat ha estat una exigència del propi sistema per tal de delimitar l'abast de les

famílies de paraules i, en aquest tema, no és possible seguir uns criteris d'absoluta objectivitat. Vegem-ho amb un exemple: de 'mar' és clar que podem derivar-ne 'marinada' ("vent de la banda de mar que bufa a hores determinades"), 'marinatge' ("art de navegar"), 'marinejarse' ("rebre la marinada"), 'marinenc' ('de mar, marí"), etc., però el cas de la forma 'marinar' ("adobar [el peix] per conservar-lo") planteja problemes ja que el seu sentit es troba allunyat del terme inicial 'mar-'.

En aquests casos cal adoptar solucions d'acord amb els resultats que s'esperen obtenir de l'anàlisi morfològica.

Sempre que ha estat possible, hem resolt les formes derivades a partir d'un SI mitjançant regles de l'analitzador. Amb tot, de vegades això no ha estat possible, i hem hagut de definir SI derivats:

a)- quan els SI presenten alteracions formals. En aquests casos, per mantenir la cohesió de la família de paraules hem creat un SI germà, és a dir, un SI vinculat internament a un altre que considerem bàsic:

| SI bàsics | SI germans |
|-----------|-----------------------|
| llob-(0) | llob-(-a,...) |
| viu-(-re) | viv-(-ia, -or,...) |
| pas-(0) | pass-(-ar, -ador,...) |

| | |
|-------------|------------------------|
| nàix-(-er) | nàix-(-er, -ement,...) |
| roj-(0) | roj-(-a), rog-(-es) |
| admet-(-re) | admès-(-0) |
| etc. | |

Hem considerat SI bàsic aquell que serveix de base per formar el singular dels noms, el masculí singular dels adjectius i el gerundi dels verbs.

Es pot donar el cas que els derivats es formin a partir del SI qualificat com a germà:

| | | |
|-----------|-----------------|----------|
| "rog-enc" | SI bàsic= "roj" | CAT=ADJ |
| | SI germà= "rog" | GEN=MASC |
| | | NBRE=SG |

b)- quan el derivat resulta de la unió d'un SI i d'un prefix. El conjunt [prefix-SI] figura al diccionari com un SI 'germà' :

| SI bàsic | SI germans |
|--------------|-----------------|
| genoll-(0) | agenoll-(-ar) |
| inform-(-ar) | desinform-(-ar) |
| aten-(-dr-e) | desaten(-dr-e) |

Si bé l'analitzador permetria el tractament dels prefixos, els problemes d'anàlisi que plantejaria aquesta solució són suficientment importants com per defugir aquesta via.

Això no obstant, presentarem a continuació les dues estratègies, i no les solucions, que permetrien el

tractament dels prefixos amb aquest analitzador per tal de justificar la nostra decisió.

1)- Es podria crear un model que contingues tots els prefixos i intervingues en una regla que tingues START com a estat inicial i final. Els prefixos que contindria el model es combinarien amb qualsevol SI. Els resultats d'emprar aquesta estratègia justifiquen que l'hàgim desestimada. Així, formes com 'abadia' tindrien dues anàlisis possibles:

INTERPRETACIO 1:

| | | |
|----------|------------|---------|
| 'abadia' | SI= abadi- | CAT=NOM |
| | | GEN=FEM |
| | | NBRE=SG |

INTERPRETACIO 2:

| | | |
|----------|------------|---------|
| 'abadia' | SI= badi- | CAT=NOM |
| | Prefix= a- | GEN=FEM |
| | | NBRE=SG |

2)- Una segona possibilitat fóra crear un model amb els prefixos que tingues START com a estat inicial i com a estat final de regla un estat reconeixedor de prefixos, p.e. ERP. D'aquest estat haurien de sortir els models dels SI que accepten un prefix. Els prefixos haurien de dur un atribut específic i els SI que admeten un prefix s'haurien d'organitzar en models en funció del prefix que accepten, de manera que la seva regla tingues com a condició l'atribut del prefix:

| <MODEL PEF> | <MOD. SI A-> | <MOD. SI DES-> |
|--------------|---------------------|-----------------------|
| a- PEF=A | badan- | consol- |
| des- PEF=DES | genoll- grumoil- | f- (fer) composar- |

I les regles serien:

```

R1  START  <MODEL PEF>  ERP
R2  ERP    <MOD. SI A->  E4
      condició: PEF=A
R3  ERF    <MOD. SI DES-> E4
      condició: PEF=DES

```

Com que als models de SI hi hauria barrejats SI nominals, adjectius i verbals, de l'estat E4 haurien de sortir-ne els models de SM i de SF que ja haurien estat definits per a les formes sense prefix.

Aquesta estratègia és igualment inconvenient, perquè s'haurien d'entrar els SI més d'una vegada (per a la seva flexió amb prefix i sense prefix), amb la qual cosa ja no s'estalviarien entrades al diccionari i a més augmentaria considerablement el nombre de regles, així com el nombre de passos per a l'anàlisi de cada forma.

Els mots hereditaris i els mots manllevats han estat relacionats, com a SI 'base' i SI 'germà', sempre que en la llengua actual mantinguin un lligam semàntic:

Mots hereditaris Mots manllevats

| | |
|-----------------|------------------------------|
| 'moble' | 'mobiliari' |
| 'noble' | 'nobiliari' |
| 'vesc' | 'viscari' |
| 'boca' | 'bucal' |
| 'home' | 'human-' (-itat, -itari,...) |
| 'pit' | 'pectoral' |
| 'fruit' | 'fructuós' |
| 'llengua' | 'lingual' |

etc.

Mots com ara: 'estret' / 'estricte', 'eixam' / 'examen', 'regla' / 'rella', que en la llengua actual no tenen altre lligam que l'etimològic, no els hem relacionat com formant part d'una mateixa família de paraules.

IV.2.5.4.1.- Els derivats

Considerem formes derivades les resultants de la combinació d'un SI amb un SM:

rapid(SI)-esa(SM)
ADJ NOM

estupid(SI)-itz(SM)-ació(SM)
ADJ V NOM

o bé un SI amb un SF, sempre que impliqui un canvi de categoria:

rod(SI)-ar(SF) salt(SI)-ar(SF)
NOM V NOM V

Constitueixen el recurs més habitual per formar noves paraules. Un dels grans avantatges de l'ús d'un analitzador com el que presentem és precisament la possibilitat d'obtenir tots els derivats d'un mot a

partir del seu SI mitjançant la seva combinatòria amb SM i SF.

IV.2.6.- Les interpretacions

Un analitzador com el nostre, independent del context, davant d'una forma que pot tenir més d'una interpretació dona totes les possibilitats d'anàlisi. Això es produeix en els casos següents:

a)- En les homonímies donem una interpretació per a cada cas:

tombant Interpretació 1 : VERB, GERUNDI

Interpretació 2 : NOM, MASCULI

roda Interpretació 1 : VERB, PRESENT

Interpretació 2 : VERB, IMPERATIU

Interpretació 3 : NOM, FEMENI

b)- En les formes de la mateixa categoria que tenen diferents valors per als altres atributs, hem definit valors que representen la diversitat d'interpretacions:

'pintaria' CAT=VERB, PERS= 1/3, NBRE=SG

'rebel' CAT=ADJ, GEN=M/F, NBRE=SG

c)- En els casos de polisèmia hem donat tan sols una sortida d'analitzador:

gra CAT=NOM, GEN=MASC, NBRE=SG

banc CAT=NOM, GEN=MASC, NBRE=SG

atracar CAT=VERB, FNP=INF

i tots els derivats, de qualsevol de les accepcions que tingui el SI, es derivaran del mateix SI.

IV.2.7.- El corpus

Amb el present analitzador pretenem mostrar la problemàtica que planteja l'anàlisi morfològica computacional. Així, hem dissenyat un analitzador general mitjançant el tractament exhaustiu de la combinatòria de SI, SM, SQF i SF. Els diccionaris inclouen la totalitat de SM, SQF i SF; en canvi, només hem entrat una mostra representativa de SI triada prèviament amb la finalitat de presentar com a mínim un cas per a cada paradigma flexiu o derivatiu, a més dels SI corresponents a la mostra que analitzem.

Per tal de posar a prova la capacitat de l'analitzador hem entrat a la base de dades textual de l'analitzador textos de diferents autors, representatius de diferents nivells d'ús: un poema de Salvador Espriu de El Caminant i el Mur, "Sentit a la manera de Salvador Espriu", un fragment del Quadern Gris de Josep Pla, una notícia de la secció d'economia del diari "Avui" del dia 14 de juny de 1987 i un fragment d'un article

signat per Vicenç Villatoro del mateix dia.

IV.3.- L'accent i la dièresi

Aquest analitzador no reconeix ni el caràcter de l'accent ni la dièresi.

Per tal que quedi constància que en una forma hi ha un accent, hem utilitzat la convenció de posar al final de la forma el signe '&' si es tracta d'un accent agut i '&1' si es tracta d'un accent greu:

'constancia&1', 'promes&1', 'matis&', etc.

L'accent pot correspondre al SI, al SM o al SF d'una forma. Quan correspon al SF entrem en el diccionari el SF seguit del caràcter de l'accent. Quan l'accent correspon a un SI o a un SM:

a)- forçosament el caràcter de l'accent ha d'anar a continuació del SF,

b)- el model al qual pertany el segment inicial porta el caràcter '&' o bé '&1' així com l'atribut per a la validació i/o el valor:

<NM&1S>

CAT=NOM
T&=M&1S
trafec-
caracter-

<M&1S>

GEN=MASC
-&1 NBRE=SG
-s&1 NBRE=PL

Les regles que permetran la seva anàlisi són:

```
R1  START    <NM&IS>    RNA
R2  RNA      <M&IS>    RGN
      condició: T&=M&IS
```

La idea que ens ha guiat ha estat la creació d'un mitjà fàcil de recuperació dels SI i SM amb accent per si alguna vegada es pot implementar l'analitzador en algun ordinador que el reconegui.

Pel que fa a la dièresi, hem definit l'atribut DIER amb els valors I1, I2, I3, U1, U2, U3 per indicar que la dièresi és a la primera, segona o tercera '-i-' o '-u-' a partir de la dreta, i els valors IA i UA per indicar que la dièresi és a la '-i-' o a la '-u-' del SI.

Com a exemple, tenim les primeres persones del singular del pretèrit perfet i del present de subjuntiu de canviar:

```
cavii&  CAT=VERB      canvii  CAT=VERB
        TEMP=PERF     TEMP=PRES
        MODE=IND      MODE=SUBJ
        PERS=1        PERS=1/3
                   DIER=I1
```

Des del punt de vista dels textos que volem analitzar significa que en entrar el text a la base de dades textual s'ha d'observar la convenció de posar, al final de cada forma amb accent, els caràcters '&' o '&1' segons que es tracti d'un accent agut o greu.

V.- ANALISI DE LA FLEXIÓ

Les formes flexives són aquelles que resulten de combinar-se un SI i un SF, amb o sense un SQF.

La problemàtica que planteja la flexió verbal i la flexió nominal i adjectiva és diversa i per tant hem definit dos subapartats per a la seva descripció: a V.2. la conjugació verbal i a V.3. l'anàlisi de noms i adjectius.

V.1.- Paradigma i model

En el marc d'aquesta investigació el terme 'paradigma' respon estrictament a necessitats expositives per a la descripció de la flexió nominal, adjectiva i verbal.

Com ja hem exposat en el capítol anterior, hem definit models diferents per a aquelles formes que en la seva flexió tenen més d'un SI:

| <NM&> | <NMNS> |
|--------------------|---------|
| CAT=NOM(1) | CAT=NOM |
| TG=&(2) | TG=MNS |
| garni-.....garrin- | |
| pilo-.....pilon- | |
| etc. | |

(1) L'atribut CAT indica la categoria.

(2) L'atribut TG defineix les condicions de les regles dels SF de gènere i nombre.

<VARB>

<VARBU>

CAT=VERB
TV=6

CAT=VERB
TV=6U (1)

arrenc-.....arrenqu-
ajaç-.....ajac-

En aquests casos hem definit un SI com a 'bàsic' i l'altre com a SI 'germà'.

La diferenciació entre SI 'bàsics' i 'germans' i la seva distribució en models diferents, com acabem de veure, fa necessària la diferenciació de dos nivells en el tractament expositiu de la flexió: el de 'paradigma' i el de 'model'.

Entenem per paradigma flexiu el(s) model(s) de SI que, combinat(s) amb models de SF, permet(en) donar compte de totes les formes flexives d'un mot. Per tant, cada paradigma està definit per un o més models de SI. Els models de SF d'un paradigma són implícits en l'atribut del model o del SI que permet validar la flexió.

Es tracta d'un concepte útil només per al nivell expositiu, ja que l'analitzador opera amb models i desconeix per complet la noció de 'paradigma'.

Segons les variants flexives dels SI, hi ha paradigmes d'un sol model, de dos o de tres models.

(1) L'atribut TV defineix les condicions de les regles dels SF de la conjugació.

Com a exemple de paradigmes que consisteixen en un sol model tenim:

Paradigma <VAR>: (verbs regulars de la primera conjugació)

Model <VAR>

Atributs del model: CAT=VERB, TV=AR

SI del model: salt-, agaf-, mir-, etc.

Paradigma <NMBS>: (noms masculins com: 'fruit, fruito')

Model <NMBS>

Atributs del model: CAT=NOM, TG=MBS

SI del model: fruit-, sol-, adob-, tomb-, etc.

Com a paradigmes de dos models, a les formes verbals tenim:

Paradigma <VARG> <VARGU>: (verbs amb alternança al SI: 'caçar/caces')

Model <VARG>

Atributs del model: CAT=VERB, TV=G

SI del model: acluc-, agenç-, agreuj-, etc.

Model <VARGU>

Atributs del model: CAT=VERB, TV=GU

SI (germans) del model: acluqu-, agenc-, agreug-, etc.

i com a exemple de paradigma nominal amb més d'un model tenim:

Paradigma <NM&> <NMNS> (noms amb més d'un SI com:
'matí/matins')

Model <NM&>

Atributs del model: CAT=NOM, TG=&

SI del model: mati-, garri-, porro-, etc.

Model <NMNS>

Atributs del model: CAT=NOM, TG=MNS

SI (germans) del model: matin-, garrin-, porron-,
etc.

Si un SI presenta més de tres variants flexives, ens trobem davant de formes molt irregulars i generalment úniques. En aquests casos hem agrupat les diferents variants de SI en un sol model i hem donat els atributs per validar les formes flexives a cada SI en particular. És el cas dels verbs irregulars com 'haver', 'poder', etc.:

<VHAVER>

CAT=VERB

he- TV=HE
hau- TV=HAU
hav- TV=HAV
heu- TV=HEU
ha- TV=HA

<VPODER>

CAT=VERB

pot- TV=POT
pod- TV=POD
pu- TV=PU
po- TV=PO

En aquests casos i en els de paradigmes d'un sol model com <VAR> i <NMBS>, paradigma i model coincideixen.

V.1.1.- Presentació de les formes flexives

La presentació de la flexió verbal, nominal i adjectiva segueix l'esquema següent:

a)- El paradigma: el(s) model(s) de SI.

b)- Descripció del(s) model(s) de SI: els atributs del model, els SI del model i els atributs dels SI, si en tenen.

c)- Relació de les regles que fan possible l'anàlisi. En primer lloc, les regles del(s) SI, a continuació les regles dels SF.

La numeració de les regles és independent a cada paradigma, per tal que el lector pugui veure fàcilment el nombre de regles necessàries per resoldre cada cas.

A l'apèndix núm. 1 hi figura una relació completa de les regles de la primera, segona i tercera conjugació per tal de facilitar-ne la consulta.

V.2.- Les formes verbals

Considerem formes verbals aquelles que resulten de combinar-se els SI categoritzats com a verbs amb SF portadors dels atributs de nombre (NBRE), persona (PERS), temps (TEMP) i mode (MODE). En el cas de formes no personals del verb (FNP), els participis porten els atributs de gènere (GEN) i nombre (NBRE).

Des de la perspectiva d'aquest analitzador, el tractament de les formes verbals es diferencia de l'anàlisi de la resta de formes per la definició de SQF (Segments Quasi Finals) que corresponen a les modificacions dels SI consistents en la presència de nous caràcters.

V.2.1.- Disseny de l'analitzador de les formes verbals

Considerem com a paradigma de flexió verbal el(s) model(s) de SI que es combinen amb un determinat conjunt de models de SF. El resultat que s'obté de l'anàlisi és el conjunt de totes les formes flexives dels SI verbals inclosos als models.

Hem partit de la classificació tradicional en tres conjugacions:

- primera conjugació: verbs l'infinitiu dels quals acaba en '-ar';
- segona conjugació: verbs que formen l'infinitiu

amb '-er' o '-re' (i també 'dur', 'dir' i 'fer');

- tercera conjugació: verbs caracteritzats per formar l'infinitiu amb el SF '-ir'.

Aquesta classificació ha suposat la definició de tres estats finals diferents per als models de SI verbals:

RAR1 és l'estat final dels models de SI de la primera conjugació;

RAR2 és l'estat final dels models de SI de la segona conjugació i;

RAR3 és l'estat final dels models de SI de la tercera conjugació.

Dels estats RAR1, RAR2 i RAR3 surten les regles dels models de SF de cada una de les tres conjugacions i RV n'és l'estat final. D'aquest estat surt el model que permet el reconeixement del caràcter del blanc i F (1) es l'estat final.

Si START és l'estat inicial de tots els models de SI, l'esquema general de les regles de SI de la conjugació és:

(1) F és un estat definit com a final: quan en el procés d'anàlisi s'arriba a aquest estat, la forma ha estat analitzada.

| | | | |
|----|-------|---------------|------|
| R1 | START | <SI 1a. CONJ> | RAR1 |
| R2 | START | <SI 2a. CONJ> | RAR2 |
| R3 | START | <SI 3a. CONJ> | RAR3 |

Són possibles d'altres esquemes. A continuació avaluarem altres possibles solucions respecte de la que hem definit.

Els esquemes alternatius responen a dues estratègies fonamentals:

a)- definició d'un estat final de regla per a cada model de SI, independentment de la conjugació a la qual pertany ;

b)- definició d'un sol estat final per a les regles dels SI de tota la flexió verbal.

La primera solució es fonamenta en el fet que cada una de les tres conjugacions no representa un tipus únic de comportament flexiu: a cada conjugació s'han hagut de definir diferents paradigmes flexius, tant per canvis en els SI com per variacions en els SF o en la combinatòria d'uns i altres. Només a la primera conjugació, que és la més regular, ha estat necessari definir els paradigmes:

P1- <VAR>: verbs regulars com "saltar" , "ajudar", etc.;

P2- <VARB> <VAR3U>: verbs amb alternances als SI:
"minj-/meng-";

P3- <VCANV>: verbs amb dièresi als SF: "canvi-i" ;

P4- <VESGL>: verbs amb dièresi als SI: "esglai" ;

P5- <VAN>: es tracta d'un verb únic, "anar",
caracteritzat per presentar l'alternança de SI 'an-
/va-' , i

P6- <VESTAR>: també un verb únic que presenta
l'alternança de SI 'est-' / 'esti-'.

La tercera conjugació, a més de presentar variants com
les que acabem de veure, està constituïda per dos
grans grups de verbs que es diferencien per la
presència/ absència del SQF '-eix'. En aquest cas ens
trobem amb una diversitat més gran de models de SI i
de SF que a la primera conjugació: en total són setze
models de SI.

La segona conjugació comprèn paradigmes molt diversos,
alguns dels quals tenen pocs elements en comú.

Malgrat aquesta diversitat hem optat per definir un
estat reconeixedor de SI per a cada conjugació i no un
estat per a cada paradigma. Aquesta decisió ha estat
presa per les raons següents:

a)- perquè encara que cada conjugació pot presentar variants flexives, es pot considerar que existeix un conjunt de SF característic de cada una i per tant es justificable la definició d' un mateix estat inicial per al conjunt de models d'aquests SF;

b)- perquè permet un estalvi important de regles: sovint els models de SF són vàlids per a més d'un paradigma en una mateixa conjugació. Si tots els models de SF d'una determinada conjugació intervenen en regles que tenen el mateix estat inicial, només cal validar les regles dels SF per al model de SI que els correspon, sense d'haver de definir la regla més d'una vegada, fet inevitable si es definís un estat final per a cada paradigma flexiu.

c)- perquè s'augmentaria considerablement el nombre de regles de SM nominalitzadors i adjectivadors de SI verbals ja que de cada un dels estats reconeixadors de SI verbals haurien de sortir totes, o gairebé totes, les regles dels models de SM .

La segona solució, és a dir, la definició d'un sol estat final per a les regles de SI de la conjugació, presentaria l'avantatge de no haver de definir més que una sola vegada les regles de determinats models de SF comuns a més d'una conjugació.

Encara que amb la definició d'un sol estat final per a les regles dels SI es podrien rendibilitzar al màxim els models de flexió, hem definit tres estats finals per als models de SI verbals per les ranns següents:

a)- la reducció d'opcions a partir d'un estat: si definim un sol estat final per als models dels SI verbals, totes les regles dels SF i dels SQF haurien de sortir d'aquest estat, de manera que per a l'anàlisi de cada forma verbal l'analitzador hauria de recórrer totes les regles dels SF de la conjugació.

Amb la definició de tres estats finals, corresponents a les tres conjugacions, per als diferents models de SI, només surten de cada un d'aquests estats les regles de SF d'una determinada conjugació, amb la qual cosa es redueix sensiblement el nombre d'opcions de recerca;

b)- el control dels SM que nominalitzen o adjectiven els SI de cada conjugació;

c)- la necessitat de no excedir el límit de validacions per regla, inevitable en el cas que tots els models de SF de la flexió verbal tinguessin un mateix estat inicial de regla.

El disseny que proposem per a l'anàlisi de la flexió verbal és una solució intermèdia entre les dues que

acaben de descriure: distingim tres grans grups de verbs mitjançant la definició de tres estats finals. Aquesta solució evita tant la màxima unificació, representada per un sol estat final per a les regles dels SI, així com la total diversitat, que correspondria a un estat per a cada paradigma de flexió.

V.2.2.- Segmentació de les formes verbals

En el capítol anterior hem presentat els criteris generals de la segmentació de formes. La flexió verbal, però, presenta una problemàtica particular que exposem a continuació.

Per tal de facilitar l'exposició, seguirem l'exemple de la fig.1.

FUT.

estim-ar-é resol-dr-é pat-ir-é sort-ir-é

P.SUB.

estim-i resol-gu-i pat-er-i surt-i

Fig.1

Els segments 'estim-', 'resol-', 'pat-' 'sort-' i 'surt-' són SI que es mantenen invariables al llarg de la flexió.

Els segments '-é', '-i' són SF que, en aquest marc d'investigació, indiquen, de manera sistemàtica, un temps verbal determinat, una persona gramatical, el mode i el nombre. En l'exemple es pot observar que el SF '-é' representa la primera persona del futur a tots els verbs, que el SF '-i' serveix per formar la primera i tercera persona del present de subjuntiu als diferents paradigmes, etc.

En les formes de l'exemple s'observa que l'element diversificador és la presència/absència de SQF o bé les seves formes diverses.

La conjugació verbal s'hagués pogut resoldre sense necessitat de definir SQF i, en aquest cas:

a)- considerar els SQF com a part integrant dels SI. Això augmentaria el nombre de SI per cada verb i també comportaria l'augment de SI al diccionari.

| SI | SI | SI | SI |
|-----------|------------|------------|-----------|
| 'patir-' | 'resol-' | 'estim-' | 'sort-' |
| 'pat-' | 'resoldr-' | 'estimar-' | 'surt-' |
| 'pateix-' | 'resolgu-' | | 'sortir-' |

b)- Considerar els SQF com a part dels 3F:

| SF-FUT | SF-FUT | SF-FUT |
|---------------|--------|--------|
| '-aré' | '-dré' | '-iré' |
| '-aràs', etc. | | |

SF-SUB

'-i'
'-is', etc.

SF-SUB

'-gui'

SF-SUB

'-eixi'

Aquesta solució diversificaria els models de SF, augmentaria considerablement el seu nombre i, per tant, el de les regles de l'anàlitzador.

La definició de SQF permet reduir el nombre de SI o bé de SF. Les formes de l'exemple es resoldrien definint un sol model de SF per al futur i per al present de subjuntiu:

| <FUT> | <SUBJ> |
|-------|--------|
| -é | -i |
| -às | -is |
| -à | -i |
| -em | -em |
| -eu | -eu |
| -an | -in |

Aquests models serveixen per a tots els paradigmes de flexió, prèvia validació dels SI segons els diferents SQF que accepten. Així, les regles dels SQF són específiques per a cada paradigma, mentre que les regles dels SF són comunes a tots els paradigmes.

| SI | SQF | SF <FUT> |
|--------|------|----------|
| estim- | -ar- | -é |
| | / | -às |
| pat | -ir- | -à |
| | / | -em |
| resol- | -ar- | -eu |
| | | -an |

| SI | SQF | SF <SUBJ> |
|--------|-------|-----------|
| estiu- | ----- | -i |
| | / | -is |
| pat | -eix- | -i |
| | / | -em |
| resol- | -gu- | -eu |
| | | -in |

Les regles de l'anàlisi serien:

| | | | |
|----|-------|---------------|------|
| R1 | START | <SI 1a CONJ> | RAR1 |
| R2 | START | <SI 2a CONJ> | RAR2 |
| R3 | START | <SI 3a CONJ> | RAR3 |
| R4 | RAR1 | <SQF 1a CONJ> | RSQF |
| R5 | RAR2 | <SQF 2a CONJ> | RSQF |
| R6 | RAR3 | <SQF 3a CONJ> | RSQF |
| R7 | RSQF | <SF FUT> | F |
| R8 | RSQF | <SF SUBJ> | F |

Aquesta solució crea una estructura, les regles R7 i R8, que es pot validar per a qualsevol model de SI i que a la vegada permet el tractament específic dels aspectes divergents, les regles R4, R5 i R6.

V.2.3.- Definició dels models

Hem partit del criteri de la flexió verbal per a definir els models de SI. Cada model s'ha configurat segons els SF que accepta.

Hem procurat agrupar el màxim de SI en un mateix model

encara que els SI presentin divergències en algunes formes flexives. Hem resolt aquestes formes mitjançant l'assignació d'atributs a cada SI en particular:

a)- quan les diferències en el comportament flexiu no eren de gran importància, bé per tractar-se de qüestions de normativa gràfica (presència /absència d'accent) , bé perquè es tractava de verbs de comportament molt irregular només en determinades formes, per exemple el participi o el present de l'indicatiu a la segona conjugació.

b)- quan el nombre de verbs afectats era reduït i sense cap mena de possibilitat d'incorporar nous SI per creativitat lingüística, encara que les diferències en el comportament flexiu fossin més grans.

Ha estat en la resolució de la segona conjugació on més sovint hem agrupat en un mateix model segments inicials amb comportaments flexius variats perquè aquesta segona conjugació es caracteritza per la manca total de rendibilitat en la formació de nous verbs. Hem resolt les particularitats mitjançant l'assignació d'atributs als SI.

Normalment la incorporació de nous verbs segueix el model de la flexió regular de la primera conjugació i també, però menys, de la tercera. És per això que en la resolució d'aquests verbs hem procurat resoldre

l'anàlisi de les formes assignant els atributs als models i no als SI, de manera que la tasca de manteniment del diccionari fos el més senzilla possible: donar d'alta el nou SI i assignar-li el model de flexió corresponent.

En els casos de modificacions als SI d'un verb per alternances vocàliques, per canvis ortogràfics o perquè es tractava d'un verb irregular:

a)- si el nombre de SI que presentaven un mateix tipus d'alteració era elevat ('menj-/meng-', 'enaigu-/enaigü-', 'preg-/pregu-', etc.) hem donat un model per a cada variant de SI, de manera que l'atribut per validar la flexió correspon al model:

| VARG> | <VARGU> |
|----------|----------|
| CAT=VERB | CAT=VERB |
| TV=G | TV=GU |
| menj- | meng- |
| preg- | pregu- |
| enaigu- | enaigü- |
| etc. | |

Les formes de <VARG> es combinen amb uns models de SF per a l'atribut TV=G, i les formes de <VARGU> amb la resta de models de SF que configuren el total de formes de flexió, mitjançant l'atribut TV=GU.

b)- quan es tractava de verbs irregulars, que acostumen a tenir un paradigma de flexió propi, hem inclòs en un mateix model totes les variants de SI i

hem donat els atributs per validar la flexió a cada SI:

| | |
|-----------|-------------|
| <VAN> | <VPODER> |
| CAT=VERB | CAT=VERB |
| an- TV=AN | pod- TV=POD |
| va- TV=VA | pot- TV=POT |
| | pu- TV=PU |
| | po- TV=PO |

Normalment els models de SI tenen assignats els atributs de la categoria morfològica (CAT=VERB) i l'atribut per complir les condicions de les regles de la flexió, TV.

A l'exemple de la fig.1, els SI es combinen amb diferents SQF i SF i, per tant, pertanyen a diferents models de SI:

| | | |
|----------|----------|----------|
| <VAR> | <VDRE> | <VPAT> |
| CAT=VERB | CAT=VERB | CAT=VERB |
| TV=AR | TV=DRE | TV=PAT |
| estim- | resol- | pat- |

| | |
|----------|----------|
| <VSORT> | <VSURT> |
| CAT=VERB | CAT=VERB |
| TV=SORT | TV=SURT |
| sort- | surt- |

Hem tractat els models de SF des de la perspectiva de's temps verbals. El model té assignats els atributs de temps i mode, i els SF els atributs de persona i nombre:

| | |
|-----------------------|------------------------|
| <FUT> | <SSPRES> |
| TEMP=FUT | TEMP=PRES |
| MODE=IND | MODE=IND |
| -e PERS=1 NBRE=SG | -i PERS=1/3 NBRE=SG |
| -as PERS=2 NBRE=SG | -is PERS=2 NBRE=SG |
| -a PERS=3 NBRE=SG | -im PERS=1 NBRE=PL |
| etc. | |

Els models de SQF no porten cap mena d'atribut:

| | | |
|---------|--------|--------|
| <ITAR> | <ITDR> | <ITIR> |
| -ar- | -dr- | -ir- |
| <ITEIX> | <ITGU> | |
| -eix- | -gu- | |

Després de presentar els criteris per a la segmentació de formes, la definició de models i l'assignació d'atributs, clourem aquesta introducció descrivint l'esquema de l'analitzador corresponent a l'anàlisi de les formes de la Fig. 1 (pàg. 194).

Les regles per a l'anàlisi d'aquestes formes, fig.2, posen de manifest alguns dels aspectes que considerem més característics del disseny de l'anàlitzador que hem construït:

a)- L'aprofitament d'una mateixa regla de SF o de SQF per a diversos models de SI, regles R15 i R16.

b)- La definició de subsistemes generalitzables per a diversos paradigmes flexius: RFUT2 és l'estat final dels diversos models de SQF de la segona conjugació que permeten la formació del futur i del condicional (<ITER>, 'conn-en-é', <ITR>, 'permet-r-é', etc.) de manera que les regles de SF del futur i del condicional que tenen RFUT2 com a estat inicial són vàlides per a tots els verbs que formen aquests temps amb SQF;

c)- L'ús d'un mateix model a diverses regles és també un sistema de rendibilitzar els models: el model <FUT> és a les regles 9, 10 i 16, i <SBFRES> a les regles 6, 12, 14 i 17.

| | | | |
|----|-------|--------|------|
| R1 | START | <VAR> | RAR1 |
| R2 | START | <VDRE> | RAR2 |
| R3 | START | <VPAT> | REIX |

Fig. 2

| | | | |
|-----|-------|----------|------------------------------------|
| R4 | START | <VSURT> | RAR3 |
| R5 | START | <VSBORT> | RAR3 |
| R6 | RAR1 | <SBPRES> | RV condició:TV=AR |
| R7 | RAR1 | <ITAR> | RFUT1 condició:TV=AR |
| R8 | RFUT1 | <FUT> | RV condició:TV=AR |
| R9 | RAR2 | <ITDR> | RFUT2 condició:TV=DRE |
| R10 | RFUT2 | <FUT> | RV condició:TV=DRE |
| R11 | RAR2 | <ITGU> | RSUBJ condició:TV=DPE |
| R12 | RSUBJ | <SBPRES> | RV condició:TV=DRE |
| R13 | RAR3 | <ITEIX> | REIX condició:TV=PAT |
| R14 | REIX | <SBPRES> | RV condició:TV=PAT |
| R15 | RAR3 | <ITIR> | RFUT3 condició:TV=PAT, TV=SBORT |
| R16 | RFUT3 | <FUT> | RV condició:TV=PAT, TV=SBORT |
| R17 | RAR3 | <SBPRES> | RV condició:TV=SBURT |

Fig.2

La rendibilitat d'un model es determina, com s'acaba de veure, per dos factors :

a)- perquè el model s'utilitza a més d'una regla. Aquest és el cas de <FUT> i <SBPRES>;

b)- perquè la regla en què apareix un determinat model té diverses condicions és a dir, està validada per a diversos models de SI. Aquest és el cas de <ITIR>.

V.2.4.- Presentació de la conjugació

Presentem separatament els diferents paradigmes de flexió de cada una de les tres conjugacions. Cada paradigma consta de:

- a)- el(s) model(s) del(s) SI,
- b)- els atributs del model de SI
- c)- una petita mostra dels SI del model i dels atributs dels SI,
- d)- les regles de l'anàlisi.

La numeració de les regles és independent a cada paradigma.

L'apèndix-1 conté la relació completa de les regles que hem definit per a l'anàlisi de la conjugació.

L'apèndix-2 conté els models de SF i de SQF de la conjugació.

Encara que el tema dels derivats el tractem al capítol VI, presentem, conjuntament amb els atributs de la flexió, alguns dels atributs més freqüents per formar derivats a partir de SI verbals. Hem incorporat aquesta informació en la presentació de la conjugació perquè d'aquesta manera es té una visió més global de l'anàlitzador.

V.2.5.- La primera conjugació

La primera conjugació del català aplega els verbs acabats en '-ar' derivats de la primera conjugació llatina en '-ARE'. "és la conjugació més rica no sols pel cabal originari de verbs llatins i germànics, sinó perquè hi pertanyen gairebé tots els sufixos derivatius per a la formació de nous verbs" (1).

Hi pertanyen la major part dels verbs derivats de:

1. Substantius

amb prefix:

genoll > agenollar, grumoll > agrumollar,

per conversió:

fruit > fruitar, soroll > sorollar, etc.

amb sufix:

finestra > finestrejar, glòria > glorificar

2. Adjectius

amb prefix:

content > acontentar, clar > aclarir, etc.

per conversió:

alegre > alegrar, igual > igualar, etc.

(1) A. M. Badia i Margarit Manual de gramàtica històrica, 1981

amb sufix:

blanc > blanquejar, negre > negrejar, etc

parasintètics:

blanc > emblanquir, negre > ennegrir, etc.

3. Adverbis

lluny > allunyar, davall > davallar, enlaire >
enlairar, arran > arranar, etc.

L'esquema d'anàlisi d'aquesta conjugació és el més senzill i alhora el que més formes resol. L'anàlisi d'aquests verbs requereix tan sols la definició de quatre estats, a més de l'estat axiomàtic START:

- RAR1, que és l'estat final de les regles de SI i l'estat inicial de les regles de models de SF i de SQF;

- RV, que és l'estat final de les regles de models de SF;

- IF1, estat final de les regles dels models de SQF i estat inicial de les regles de models de SF;

- F, estat final que clou l'anàlisi.

L'esquema de l'analitzador dels verbs irregulars 'estar' i 'anar' difereix de la resta de verbs d'aquesta conjugació (1).

(1) Vegeu els paradigmes <VESTAR> i <VANAR>.

V.2.5.1.- Paradigmes de la primera conjugació

Paradigma P1: <VAR>

és el model més productiu de tota la conjugació: la majoria dels verbs catalans es conjuguen segons el paradigma de flexió corresponent a aquest model.

Model: <VAR>

Atributs: CAT=VERB, TV=AR, SA=ABLE, SN=AD

SI: salt-, vol-, xiul-, rod-, mir-, etc.

Regles d'anàlisi:

| | | | | |
|-----|-------|---------|------|--|
| R1 | START | <VAR> | RAR1 | per al SI |
| R2 | RAR1 | <IP1A> | RV | condició: TV=AR |
| R3 | RAR1 | <IP1B> | RV | condició: TV=AR per al present |
| R4 | RAR1 | <IIM1> | RV | condició: TV=AR per a l'imperfet |
| R5 | RAR1 | <IPP1A> | RV | condició: TV=AR |
| R6 | RAR1 | <IPP1> | RV | condició: TV=AR per al preterit perfet |
| R7 | RAR1 | <ITAR> | IF1 | condició: TV=AR |
| R8 | IF1 | <IF> | RV | condició: TV=AR per al futur |
| R9 | IF1 | <IC> | RV | condició: TV=AR per al condicional |
| R10 | RAR1 | <SBPA> | RV | condició: TV=AR |
| R11 | RAR1 | <SBPB> | RV | condició: TV=AR per al pres. de subj. |

| | | | | |
|-----|------|-----------------|----|--------------------------|
| R12 | RAR1 | <SBIM> | RV | |
| | | condició: TV=AR | | per a l'imperf. de subj, |
| R13 | RAR1 | <IMPA1> | RV | |
| | | condició: TV=AR | | |
| R14 | RAR1 | <IMP2B> | RV | |
| | | condició: TV=AR | | |
| R15 | RAR1 | <IMP1B> | RV | |
| | | condició: TV=AR | | per a l'imperatiu |
| R16 | RAR1 | <INF1> | RV | |
| | | condició: TV=AR | | per a l'infinitiu |
| R17 | RAR1 | <GER1> | RV | |
| | | condició: TV=AR | | per al gerundi |
| R18 | RAR1 | <PARTAT> | RV | |
| | | condició: TV=AR | | per al participi |

Paradigma P2: <VARG> <VARGU>

El paradigma P2 està constituït pels V. 2.2 que presenten canvis ortogràfics al seu SI provocats per contacte amb un SF. Els canvis que es produeixen poden ser sintetitzats en els següents parells: ç/c, g/gu, c/qu, j/g, qu/qu, gu/gü.

Aquest paradigma segueix dos models, <VARG> i <VARGU>, segons les característiques gràfiques del SI verbal.

a)- Model: <VARG>

Atributs: CAT=VER3, TV=G, SA=ABLÈ, SN=ADA

SI:

abalaç-

torç-

abasseg- SN=AMNT

abdic- SN=CIO&

abreuj- SN=AMN /
adequ- SN=CIOü
enaigu- SN=NO, etc.

b)- Model: <VARGU>

Atributs: CAT=VERB, TV=GU

SI: abalanc-, torc-, abassegu-, abdiqu-,
abreug-, adequ-, enaigü-, etc.

Regles d'anàlisi.

| | | | | |
|-----|-------|-----------------|------|------------------------|
| R1 | START | <VARG> | RAR1 | |
| R2 | START | <VARGU> | RAR1 | |
| R3 | RAR1 | <IP1A> | RV | |
| | | condició: TV=G | | |
| R4 | RAR1 | <IP1B> | RV | |
| | | condició: TV=GU | | per al present |
| R5 | RAR1 | <IIM1> | RV | |
| | | condició: TV=G | | per a l'imperfet |
| R6 | RAR1 | <IPP1> | RV | |
| | | condició: TV=G | | |
| R7 | RAR1 | <IPP1A> | RV | |
| | | condició: TV=GU | | per al pret.perf. |
| R8 | RAR1 | <ITAR> | IF1 | |
| | | condició: TV=G | | |
| R9 | IF1 | <IF> | RV | |
| | | condició: TV=G | | per al futur |
| R10 | IF1 | <IC> | RV | |
| | | condició: TV=G | | per al condicional |
| R11 | RAR1 | <SBPA> | RV | |
| | | condició: TV=GU | | |
| R12 | RAR1 | <SBPB> | RV | |
| | | condició: TV=GU | | per al pres. de subj. |
| R13 | RAR1 | <SBIM> | RV | |
| | | condició: TV=GU | | per a l'impf. de subj. |
| R14 | RAR1 | <IMPA1> | RV | |
| | | condició: TV=G | | |
| R15 | RAR1 | <IMP2B> | RV | |
| | | condició: TV=GU | | |
| R16 | RAR1 | <IMP1B> | RV | |
| | | condició: TV=GU | | per a l'imperatiu |
| R17 | RAR1 | <INF1> | RV | |
| | | condició: TV=G | | per a l'infinitiu |
| R18 | RAR1 | <GER1> | RV | |
| | | condició: TV=G | | per al gerundi |

R19 RAR1 <PARTAT> RV
condició: TV=G per al participi

Paradigma P3: <VDONAR>

El model <VDONAR> inclou el SI 'don-' i el SI 'd-',
corresponents als verbs 'donar' i 'dar',
respectivament.

L'única diferència de 'donar' respecte del paradigma
<VAR> és la presència de l'accent diacrític al
present i a l'imperatiu.

Hem resolt l'anàlisi d'aquests verbs definint el model
<VDONAR> que conté els SI d'ambdós verbs. Els atributs
per validar la flexió són als SI.

Model: <VDONAR>

Atributs: CAT=VERB

SI:

don- TV=DON, T&=DON&

d- TV=DAR

Regles d'anàlisi:

R1 START <VDONAR> RAR1 per als SI

R2 RAR1 <IPDON> RV
condició: TV=DON

R3 RAR1 <IPDON&> RV
condició: TV=DON

R4 RAR1 <IPB> RV
condició: TV=DAR per al present

R5 RAR1 <IIM1> RV
condició: TV=DON, TV=DAR per a l'imperf.

| | | | | |
|-----|------|----------|-----|---|
| R6 | RAR1 | <IPP1> | RV | |
| | | | | condició: TV=DON |
| R7 | RAR1 | <IPP1A> | RV | |
| | | | | condició: TV=DON |
| R8 | RAR1 | <IPP1DA> | RV | |
| | | | | condició: TV=DAR per al pret.perf. |
| R9 | RAR1 | <ITAR> | IF1 | |
| | | | | condició: TV=DON, TV=DAR |
| R10 | IF1 | <IF> | RV | |
| | | | | condició: TV=DON, TV=DAR per al fut. |
| R11 | IF1 | <IC> | RV | |
| | | | | condició: TV=DON, TV=DAR per al cond. |
| R12 | RAR1 | <SBPA> | RV | |
| | | | | condició: TV=DON |
| R13 | RAR1 | <SBPB> | RV | |
| | | | | condició: TV=DON, TV=DAR per al pres.de subj. |
| R14 | RAR1 | <SBIM> | RV | |
| | | | | condició: TV=DON |
| R15 | RAR1 | <SBIMDA> | RV | |
| | | | | condició: TV=DAR per a l'imperf. de subj. |
| R16 | RAR1 | <IMPA1&> | RV | |
| | | | | condició: T&=DON& |
| R17 | RAR1 | <IMP2B> | RV | |
| | | | | condició: T&=DON& |
| R18 | RAR1 | <IMP1B> | RV | |
| | | | | condició: TV=DON, TV=DAR per a l'imper. |
| R19 | RAR1 | <INF1> | RV | |
| | | | | condició: TV=DON, TV=DAR per a l'infin. |
| R20 | RAR1 | <GER1> | RV | |
| | | | | condició: TV=DON, TV=DAR per al ger. |
| R21 | RAR1 | <PARTAT> | RV | |
| | | | | condició: TV=DON, TV=DAR per al part. |

Paradigma P4: <VCANV>

Els verbs del paradigma <VCANV> es caracteritzen per la presència de dièresi al SF de les tres persones del singular i de la tercera persona del plural del present de subjuntiu i a les terceres persones de l'imperatiu. Com que l'analitzador no reconeix el caràcter de la dièresi, l'hem representada per un atribut, DIER, que pren diferents valors segons que la dièresi sigui al SI o al SF, a la 'i' o la 'u'. En

el cas d'aquest paradigma el valor de l'atribut és I1,
(dièresi a la primera 'i' a partir de la dreta).
L'atribut que indica la presència de dièresi és als
models de SF (1).

Model: <VCANV>

Atributs: CAT=VERB, TV=CANV, SA=ABLE

SI: canvi-, accentu-, evacu-, llo-, menyspre-,
aline-, conre-, arre-, arri-, escoli-,
espoli-, po-, empro-, inco-, su-,
entresu-, actu-, fluctu-, etc.

Regles d'anàlisi:

R1 START <VCANV> RAR1
R2 RAR1 <IP1A> RV
condició: TV=CANV
R3 RAR1 <IP1B> RV
condició: TV=CANV per al present
R4 RAR1 <IIM1> RV
condició: TV=CANV per a l'imperfet
R5 RAR1 <IPP1A> RV
condició: TV=CANV
R6 RAR1 <IPP1> RV
condició: TV=CANV per al pret.perf.
R7 RAP1 <SBPAD> RV
condició: TV=CANV
R8 PAR1 <SBPB> RV
condició: TV=CANV per al pres.subj.
R9 RAR1 <EBIM> RV
condició: TV=CANV per a l'imperf.subj.
R10 RAR1 <IMPA1> RV
condició: TV=CANV
R11 RAR1 <IMP2BD> RV
condició: TV=CANV
R12 RAR1 <IMP1B> RV
condició: TV=CANV per a l'imperatiu
R13 RAR1 <ITAR> IF1
condició: TV=CANV
R14 IF1 <IF> RV
condició: TV=CANV
R15 IF1 <IC> RV
condició: TV=CANV per al fut. i cond.

(1) Vegeu a l'apèndix II la descripció dels models
<SBPAD> i <IMP2BD>.

R16 RAR1 <INF3> RV
 condició: TV=CANV
 R17 RAR1 <INF1> RV
 condició: TV=CANV per a l'infinitiu
 R17 RAR1 <GER1> RV
 condició: TV=CANV per al gerundi
 R18 RAR1 <PARTAT> RV
 condició: TV=CANV per al participi

Paradigma P5: <VESGL>

Els verbs del paradigma <VESGL> es caracteritzen perquè tenen un accent gràfic al SI de la primera persona del pretèrit perfet i perquè presenten el caràcter de la dièresi al SI de les tres persones del singular i a la tercera del plural del present de subjuntiu, i a les terceres persones de l'imperatiu.

Model: <VESGL>

Atributs del model: CAT=VERB, TV=ESGL, SA=ABLE,
 T&=IPP&

SI: aboi-, monjoi-, volei-, estormai-, etc.

Regles d'anàlisi:

R1 START <VESGL> RAR1 per als SI
 R2 RAR1 <IP1A> RV
 condició: T/=ESGL
 R3 RAR1 <IP1B> RV
 condició: TV=ESGL per al present
 R4 RAR1 <IIM1> RV
 condició: TV=ESGL per a l'imperfet
 R5 RAR1 <IPP1&> RV
 condició: T&=IPP&
 R6 RAR1 <IPP1> RV
 condició: TV=ESGL per al pret. perf.
 R7 RAR1 <SBPD> RV
 condició: TV=ESGL
 R8 RAR1 <SBPB> RV
 condició: TV=ESGL per al subj. pres.

- R9 RAR1 <SBIM> RV
condició: TV=ESGL per al subj. imperf.
- R10 RAR1 <IMPA1> RV
condició: TV=ESGL
- R11 RAR1 <IMPD> RV
condició: TV=ESGL
- R12 RAR1 <IMP1B> RV
condició: TV=ESGL per a l'imperatiu
- R13 RAR1 <ITAR> IF1
condició: TV=ESGL
- R14 IF1 <IF> RV
condició: TV=ESGL per al futur
- R15 IF1 <IC> RV
condició: TV=ESGL per al condicional
- R16 RAR1 <INF1> RV
condició: TV=ESGL per a l'infinitiu
- R17 RAR1 <GER1> RV
condició: TV=ESGL per al gerundi
- R18 RAR1 <PARTAT> RV
condició: TV=ESGL per al participi

Paradigma P6: <VESTAR>

Aquest paradigma inclou un verb únic en la seva modalitat. Aquest verb es caracteritza per la presència del SQF '-gu-' al pretèrit perfet, al present i a l'imperfet de subjuntiu i a l'imperatiu. Per a la resolució d'aquestes formes hem creat un estat, RGU1, que funciona com a estat final de la regla del SQF '-gu-', model <ITGU>. De l'estat RGU1 surten les regles dels models de SF que es combinen amb aquest SQF (regles 11-17).

La primera forma del present 'estic' es resol combinant el SI 'esti-' amb el SF '-c', model <IPIC>, que també utilitzem per resoldre d'altres formes verbals com 'resolc', 'vinc', etc.

Els atributs per validar els SM nominalitzadors i adjectivadors no es poden donar, en aquest cas, al model, perquè aquest aplega SI amb comportaments derivatius i flexius diferents. Aquests atributs s'han de donar a cada SI en particular.

Model: <VESTAR>

Atributs: CAT=VERB

SI:

est- TV=EST, SA=ABLE, SN=ADA

esti- TV=ESTI

Regles d'anàlisi:

- R1 START <VESTAR> RAR1
- R2 RAR1 <IPIC> RV
condició: TV=ESTI
- R3 RAR1 <IPEST> RV
condició: TV=EST per al present
- R4 RAR1 <IIM1> RV
condició: TV=EST per a l'imperfet
- R5 RAR1 <ITAR> IF1
condició: TV=EST
- R6 IF1 <IF> RV
condició: TV=EST per al futur
- R7 IF1 <IC> RV
condició: TV=EST per al condicional
- R8 RAR1 <INF1> RV
condició: TV=EST per a l'infinitiu
- R9 RAR1 <GER1> RV
condició: TV=EST per al gerundi
- R10 RAR1 <PART1> RV
condició: TV=EST per al participi
- R11 RAR1 <ITGU> RGU1
condició: TV=ESTI
- R12 RGU1 <IPP> RV
condició: TV=ESTI per al pret. perf.
- R13 RGU1 <SBPA> RV
condició: TV=ESTI
- R14 RGU1 <SBPB> RV
condició: TV=ESTI per al pres. subj.
- R15 RGU1 <SBIM> RV
condició: TV=ESTI per a l'imperf. de subj.

R16 RGU1 <IMPDI> RV
condició: TV=ESTI
R17 RGU1 <IMP1> RV
condició: TV=ESTI per a l'imperatiu

La regla 11 (R11) reconeix el SQF '-gu-' que serveix per formar el pretèrit perfet, el subjuntiu i l'imperatiu. Hauríem pogut resoldre aquestes formes augmentant el nombre de SI: 'estic-', 'estigu-', però atès que el model <ITGU> permet resoldre formes semblants a d'altres paradigmes de la segona i tercera conjugació, hem desestimat aquesta solució.

Paradigma P7: <VAN>

El model <VAN> aplega tan sols els SI del verb 'anar'. Aquest verb es caracteritza perquè:

a)- Presenta una alternança de dos SI 'an-' i 'va-'. Atès que el gerundi es forma sobre el SI 'an-', hem considerat aquest SI com a forma de base i 'va-' com a segment "germà";

b)- La seva primera forma del present de l'indicatiu presenta un SF específic '-ig-' que també apareix a alguns verbs de la segona conjugació ('ve-ig') al qual he assignat el model <IPIG>.

c)- Presenta el SQF '-ir-' al futur i al condicional en lloc del SQF '-ar-', habitual a la resta de verbs de la primera conjugació. L'estat IF1, que reconeix <ITAR>, reconeix també <ITIR> (R.8) i

així les regles dels models <IF> (R.9) i <IC> (R.10) que surten de IF1 són vàlides per a tots els verbs de la primera conjugació independentment del SQF que tinguin.

d)- Forma les tres persones del singular i la tercera del plural del present de subjuntiu i les dues terceres persones de l'imperatiu amb el SQF '-g-'. Per resoldre aquestes formes hem utilitzat l'estructura que ja hem creat per al verb 'estar': de RGU1, que és l'estat final de la regla del SQF '-g-', en surten les regles dels models dels SF.

Model: <VAN>

Atributs: CAT=VERB

SI:

an- TV=AN, SN=ADA

va- TV=VA

Regles d'anàlisi:

| | | | | |
|----|-------|-----------------|------|--------------------|
| R1 | START | <VAN> | RAR1 | |
| | | | | per als SI |
| R2 | RAR1 | <IPIG> | RV | |
| | | condició: TV=VA | | |
| R3 | RAR1 | <IPHA> | RV | |
| | | condició: TV=VA | | |
| R4 | RAR1 | <IPB> | RV | |
| | | condició: TV=AN | | per al present |
| R5 | RAR1 | <IIM1> | RV | |
| | | condició: TV=AN | | per a l'imperfet |
| R6 | RAR1 | <IPP1A> | RV | |
| | | condició: TV=AN | | |
| R7 | RAR1 | <IPP1> | RV | |
| | | condició: TV=AN | | per al pret. perf. |
| R8 | RAR1 | <ITIR> | IF1 | |
| | | condició: TV=AN | | |

| | | | | |
|-----|------|-----------------|------|-------------------------|
| R9 | IF1 | <IF> | RV | |
| | | condició: TV=AN | | per al futur |
| R10 | IF1 | <IC> | RV | |
| | | condició: TV=AN | | per al condicional |
| R11 | RAR1 | <ITG> | RGU1 | |
| | | condició: TV=VA | | |
| R12 | RGU1 | <SBPA> | RV | |
| | | condició: TV=VA | | |
| R13 | RAR1 | <SBPB> | RV | |
| | | condició: TV=AN | | per al present de subj. |
| R14 | RAR1 | <SBIM> | RV | |
| | | condició: TV=AN | | per a l'imp. de subj. |
| R15 | RAR1 | <IMP;E> | RV | |
| | | condició: TV=AN | | |
| R16 | RGU1 | <IMP2B> | RV | |
| | | condició: TV=VA | | per a l'imperatiu |
| R17 | RAR1 | <INF1> | RV | |
| | | condició: TV=AN | | per a l'infinitiu |
| R18 | RAR1 | <GER1> | RV | |
| | | condició: TV=AN | | per al gerundi |
| R19 | RAR1 | <PARTAT> | RV | |
| | | condició: TV=AN | | per al participi |

V.2.6.- La segona conjugació

La segona conjugació aplega verbs procedents de la segona i tercera conjugacions llatines, que en català han donat verbs a b l'infinitiu en '-er' o '-re'.

Aquesta conjugació, sobre el model de la qual ja no es formen nous verbs, presenta, però, un nombre molt elevat de models flexius. Aquestes dues característiques, la manca de productivitat i la varietat de models, han condicionat el disseny de l'analitzador que, per tal de reduir el nombre de models de SI i, a la vegada, posar de manifest els aspectes comuns a diversos subparadigmes, es caracteritza per:

a)- la inclusió en un mateix model de segments inicials que divergeixen en la combinatòria amb SF, sovint amb els SF del participi, del present i de l'imperatiu. Per exemple, 'tém-er' i 'constrèny-er' pertanyen a un mateix model, <V2>, encara que el primer forma el participi amb el mateix SI, 'tem-ut' i el segon el forma amb un SI germà: 'empès'; el primer té accent agut a l'infinitiu, i el segon el té greu.

b)- l'assignació d'atributs als SI per a les validacions de les seves formes particulars:

<VDRE>

CAT=VERB
TV=DRE

ven- BLP=S

dol- TV1=DOL
BLP=S

dissol- TV1=T, BLP=S
etc.

<V2>

CAT=VERB
TV=V2

tem- T?=&
PART=UT

empeny-T&=&1

L'atribut del model que serveix per validar la flexió és sempre TV. Les formes flexives particulars dels segments inicials estan validades amb l'atribut TV1 i els valors corresponents. Quan es tracta de formes amb accent utilitzen l'atribut T& amb els valors & 1 &1. L'atribut PART serveix per validar les formes del participi que es poden formar a partir del SI.

Des de la perspectiva d'aquest analitzador, la segona conjugació es caracteritza:

a)- per la presència de SQF (tots els models de SQF comencen pels caràcters <IT...>):

1)- al futur i al condicional

<ITR>

-r- 'mou-r-é', 'mou-r-ia'

<ITER>

-er- 'corr-er-é', 'corr-er-ia'

<ITDR>

-dr- 'resol-dr-é', 'resol-dr-ia'

Els models de SQF de futur i condicional tenen RAR2 com a estat inicial de regla i IF2 com a estat final. De l'estat IF2 surten els models de SF de futur i condicional;

2)- al pretèrit perfet, present de subjuntiu, imperfect de subjuntiu i l'imperatiu:

<ITEGU>
-egu- 'corr-egu-i', 'corr-egu-em', 'corr-egu-és'

<ITGU>
-gu- 'vol-gu-i', 'vol-gu-em', 'vol-gu-és'

<ITSQU>
-squ- 'na-squ-i', 'na-squ-em', 'na-squ-és'

Aquests models de SQF tenen RAR2 com a estat inicial de regla i RGU2 com a estat final.

3)- al participi:

<ITEG>
-eg- 'corr-eg-ut'

<ITG>
-g- 'vol-g-ut'

<ITSC>
-sc- 'na-sc-ut'

<ITT>
TG=GNT
-t- 'resol-t'

<IT4>
Tf,=GN4
-t- 'mòl-t'

Els models de SQF dels participis tenen RAR2 com a estat inicial de regla i INTP com a estat final.

b)- perquè molts verbs formen el participi sobre la base d'un SI germà: 'empènyer'/ empès', 'constrènyer/ constret', 'pondre/ post', 'romandre/ romàs', 'contendre/ contès', 'defendre/ defès', 'admetre/admès', 'raure/ras', 'concloure/conclòs', etc.

Els SI germans pertanyen a diferents models segons els SF de gènere i nombre amb què es combinen:

| | | | |
|-----------|---------|-------|---------|
| <VET> | <VES&1> | <VES> | <VES&> |
| TG=GNT | TG=GN&1 | TG=GN | TG=GN& |
| constret- | empes- | mes- | confos- |
| restret- | ates- | post- | infos- |
| etc. | | | |

Des de la perspectiva d'aquest analitzador, cal distingir dos grups de verbs a la segona conjugació:

a)- Verbs que presenten un SQF a totes les seves formes: 'nàixer', 'créixer', 'aparèixer', etc.

b)- Verbs que presenten SQF només a algunes de les seves formes.

Els verbs del primer grup es caracteritzen per la presència del SQF '-ix-', que els és específic, mentre que la resta de SQF ('-gu-', '-g-', '-dr-', etc.) els comparteixen amb els verbs de l'altre grup.

Per tal de poder utilitzar el mateix sistema de regles de SF per a tota la segona conjugació, hem definit dos estats reconeixadors de SI:

1)- l'estat RVIX, estat final dels models de SI del grup a),

2)- l'estat RAR2, estat final dels models de SI dels verbs del grup b).

De l'estat RVIX surt el model del SQF '-ix-', que té com estat d'arribada RAR2, l'estat d'arribada dels models de SI del grup b). Així, totes les regles que tenen RAR2 com a estat inicial de regla són igualment vàlides per als verbs del grup a) i b).

Com ja s'ha vist en tractar els models de SQF, els estats finals de les regles en què intervenen els models de SQF són RGU2, IF2 I INTP. Aquests models, segons l'esquema que acabem de proposar, poden tenir com a estat inicial de regla RAR2 o bé RVIX.

RGU2 és l'estat inicial de regles de SF del pretèrit perfet, del present de subjuntiu, de l'imperfet de subjuntiu i de l'imperatiu.

IF2 és l'estat inicial de les regles dels SF del futur, del condicional i dels infinitius en '-re'.

INTP és l'estat inicial de les regles de SF dels participis formats amb SQF.

Vegem amb un exemple prou representatiu l'esquema general de la segona conjugació:

| | | |
|---------------------|--------------------|--------------------|
| <V:IXGU> | <VTEM> | <VDRE> |
| CAT=VERB TV=IXGU | CAT=VERB TV=TEM | CAT=VERB TV=DRE |
| apare- | tem- | resol- |

El verb del model <VIXGU> correix (en el grup a) i els altres dos models, al grup b). A continuació donem algunes de les regles necessàries, només les que poden ser significatives, per il·lustrar la definició dels estats, per a l'anàlisi d'aquestes formes:

RAR2 és l'estat final dels models de SI del grup b) i del SQF '-ix-', model <ITIX>, del grup a). De RAR2 surten els models de flexió i els models de SQF del futur i condicional, que són comuns als dos grups, ja que els verbs del grup a) formen el futur i el condicional amb tots dos SQF:

'apare-ix(SQF)-er(SQF)-é

Els verbs del grup a) no presenten el SQF '-ix-' a totes les seves formes. Alguns verbs d'aquest grup formen el subjuntiu, el pretèrit perfet i l'imperatiu amb d'altres SQF, com ara 'gu-', model <ITGU>. Hi ha

verbs del grup b) que també es combinen amb aquests SQF. Els models que resolen l'anàlisi dels SQF comuns als dos grups intervenen en regles que tenen com a estat inicial RAR2, per al grup b), i RVIX, per al grup a).

SEGMENT RECONEGUT

| | | | | |
|-----|-------|-----------|-------------------------------------|---------------------------------------|
| R1 | START | <VIXGU> | RVIX | 'apare-' |
| R2 | START | <VTEM> | RAR2 | 'tem-' |
| R3 | START | <VDRE> | RAR2 | 'resol-' |
| R4 | RVIX | <ITIX> | RAR2 | 'apareix-' |
| | | | condició: TV=IXGU | |
| R5 | RVIX | <ITGU> | RGU2 | 'aparegu-' |
| | | | condició: TV=IXGU | |
| R6 | RAR2 | <ITGU> | RGU2 | 'resolgu-' |
| | | | condició: TV=DRE | |
| R7 | RAR2 | <ITER> | IF2 | 'apareixer-', 'teme' |
| | | | condició: TV=TEM, TV=IXGU | |
| R8 | RAR2 | <ITDR> | IF2 | 'resoldr-' |
| | | | condició: TV=DRE | |
| R9 | IF2 | <IF> | RV | 'resoldré', 'apareixeré', 'temeré' |
| | | | condició: TV=DRE, TV=IXGU, TV=TEMER | |
| R10 | RGU2 | <SF SUBJ> | RV | 'aparegui', 'resolgui' |
| | | | condició: TV=IXGU, TV=DRE | |
| R11 | RAR2 | <SF SUBJ> | RV | 'temi' |
| | | | condició: TV=TEM | |

La raó d'haver definit estats finals de regla diferents per a cada SQF ha estat evitar la creació de bucles. A més, aquesta estructura permet controlar per a cada SQF quins SF accepta.

En la presentació dels paradigmes de la segona conjugació donem tots, o gairebé tots, els SI de cada model perquè en la resolució d'aquesta conjugació bona part de la informació s'ha d'associar als SI i perquè el seu nombre és reduït.

V.2.6.1.- Paradigmes de la segona conjugació

Paradigma P1: <V2> <VES&1> <VET>

El model <V2> aplega els verbs de la segona conjugació que:

a)- en la forma de l'infinitiu el SI té accent gràfic :

- agut a les formes: 'témer', 'fúmer', etc.
- greu a les formes: 'complànyer', 'plànyer' etc.

Els atributs T&=INF& i T&=INF&1 als SI permeten l'anàlisi d'aquestes formes.

b)- Per a la formació del futur, els SI es combinen amb el SQF '-er-', model <ITER>.

Alguns verbs d'aquest paradigma formen el participi amb els SF '-ut, -uda, -uts, -udes', d'altres el formen amb un SI germà. L'atribut PART=UT permet complir les condicions de les regles de SF dels

participis del primer grup. Els models <VET> i <VES&1> apleguen els SI germans del participi corresponents als verbs del segon grup. Cada un d'aquests models de SI de participi correspon a un model de flexió de gènere i nombre.

El verb 'pertànyer' presenta la forma de participi 'pertangut', que per ser una forma única i no poder-se resoldre a partir dels SI definits, s'ha entrat al diccionari al model <ED> com a SI germà de 'pertany-'.
</p></div>

a)- Model: <V2>

Atributs: CAT=VERB, TV=V2

SI:

| | |
|------------|------------------|
| tem- | PART=UT, T&=INF& |
| fum- | PART=UT, T&=INF& |
| esprem- | PART=UT, T&=INF& |
| prem- | PART=UT, T&=INF& |
| complany- | PART=UT, T&=IN&1 |
| plany- | PART=UT, T&=IN&1 |
| feny- | PART=UT, T&=IN&1 |
| trem- | PART=UT, T&=IN&1 |
| terratrem- | PART=UT, T&=IN&1 |
| pertany- | PART=UT, T&=IN&1 |
| empeny- | T&=IN&1 |
| ateny- | T&=IN&1 |
| constreny- | T&=IN&1 |
| estreny- | T&=IN&1 |

destrany- T&=IN&1

restrany- T&=IN&1

b)- Model: <VES&1>

Atributs: CAT=VERB, FNP=PART, TG=GN&1

SI:

empes-

ates-

c)- Model: <VET>

Atributs: CAT=VERB, FNP=PART, TG=GNT

SI:

constret-

estret-

destret-

restret-

Regles d'anàlisi:

| | | | | |
|-----|-------|-----------------|------|--------------------|
| R1 | START | <V2> | RAR2 | |
| R2 | START | <VES&1> | INTP | |
| R3 | START | <VET> | INTP | per als SI |
| R4 | RAR2 | <PI1> | RV | |
| | | condició: TV=V2 | | |
| R5 | RAR2 | <IPCA> | F | |
| | | condició: TV=V2 | | |
| R6 | RAR2 | <IPB> | RV | |
| | | condició: TV=V2 | | per al present |
| R7 | RAR2 | <IIM> | RV | |
| | | condició: TV=V2 | | per a l'imperfet |
| R8 | RAR2 | <IPP> | RV | |
| | | condició: TV=V2 | | |
| R9 | RAR2 | <IPP1A> | RV | |
| | | condició: TV=V2 | | per al pret. perf. |
| R10 | RAR2 | <ITER> | IF2 | |
| | | condició: TV=V2 | | |
| R11 | IF2 | <IF> | RV | |
| | | condició: TV=V2 | | per al futur |
| R12 | IF2 | <IC> | RV | |
| | | condició: TV=V2 | | per al condicional |

| | | | | |
|-----|------|----------|-------------------|-------------------------|
| R13 | RAR2 | <SBPA> | RV | |
| | | | condició: TV=V2 | |
| R14 | RAR2 | <SBPB> | RV | |
| | | | condició: TV=V2 | per al pres. de subj. |
| R15 | RAR2 | <SBIM> | RV | |
| | | | condició: TV=V2 | per al imperf. de subj. |
| R16 | RAR2 | <IMPA> | F | |
| | | | condició: TV=V2 | |
| R17 | RAR2 | <IMP9> | RV | |
| | | | condició: TV=V2 | per a l'imperatiu |
| R18 | RAR2 | <INF&> | RV | |
| | | | condició: T&=INF& | |
| R19 | RAR2 | <INF&1> | RV | |
| | | | condició: T&=IN&1 | per a l'infinitiu |
| R20 | RAR2 | <GER> | RV | |
| | | | condició: TV=V2 | per al gerundi |
| R21 | RAR2 | <PARTUT> | RV | |
| | | | condició: PART=UT | |
| R22 | INTP | <GN&1> | RV | |
| | | | condició: TG=GN&1 | |
| R23 | INTP | <GNT> | RV | |
| | | | condició: TG=GNT | per al participi |

Paradigma P2: <VCO>

Els verbs del model <VCO> es caracteritzen perquè:

a)- es combinen, al present de l'indicatiu, amb els SF '-es' i '-e', corresponents a la segona i tercera persones del singular i, a la segona persona de l'imperatiu, amb el SF '-e', models <IPE> i <IMPCC>, respectivament;

b)- es combinen amb el SQF '-egu-', model <ITEGU>, per formar el pretèrit perfet, el subjuntiu i part de l'imperatiu, enfront del SQF '-gu-', present a d'altres paradigmes de la conjugació.

Model: <VCO>

Atributs: CAT=VERB, TV=CO, TV1=F, T&=INF&

SI: 'córrer' i els seus derivats: 'deccórrer',
 'descórrer', 'discórrer', 'encórrer',
 'escórrer', etc.

Regles d'anàlisi:

| | | | | |
|-----|-------|-------------------|------|---------------------------|
| R1 | START | <VCO> | RAR2 | |
| R2 | RAR2 | <PI1> | RV | |
| | | condició: TV=CO | | |
| R3 | RAR2 | <IPE> | RV | |
| | | condició: TV=CO | | |
| R4 | RAR2 | <PIB> | RV | |
| | | condició: TV=CO | | per al present |
| R5 | RAR2 | <IIM> | RV | |
| | | condició: TV1=F | | per a l'imperfet |
| R6 | RAR2 | <ITEGU> | RGU2 | |
| | | condició: TV=CO | | |
| R7 | RGU2 | <IPP1A> | RV | |
| | | condició: TV=CO | | |
| R8 | RGU2 | <IPP> | RV | |
| | | condició: TV=CO | | per al pret. perf. |
| R9 | RAR2 | <ITER> | IF2 | |
| | | condició: TV=CO | | |
| R10 | IF2 | <IF> | RV | |
| | | condició: TV1=F | | per al futur |
| R11 | IF2 | <IC> | RV | |
| | | condició: TV1=F | | per al condicional |
| R12 | RAR2 | <SBPA> | RV | |
| | | condició: TV=CO | | |
| R13 | RAR2 | <SBPB> | RV | |
| | | condició: TV=CO | | |
| R14 | RGU2 | <SBPB> | RV | |
| | | condició: TV=CO | | per al pres. de subj. |
| R15 | RGU2 | <SBIM> | RV | |
| | | condició: TV=CO | | per a l'imperfet de subj. |
| R16 | RAR2 | <IMPCO> | RV | |
| | | condició: TV=CO | | |
| R17 | RAR2 | <IMP1> | RV | |
| | | condició: TV=CO | | |
| R18 | RGU2 | <IMP2C> | RV | |
| | | condició: TV=CO | | per a l'imperatiu |
| R19 | RAR2 | <INF&> | RV | |
| | | condició: T&=INF& | | per a l'infinitiu |
| R20 | RAR2 | <GER> | RV | |
| | | condició: TV1=F | | per al gerundi |
| R21 | RAR2 | <ITEG> | RIP2 | |
| | | condició: TV=CO | | per al participi |
| R22 | RIP2 | <PARTUT> | RV | |
| | | condició: TV=CO | | |

Paradigma P3: <VTORC> <VTOR/>

Aquest paradigma inclou els verbs que presenten canvis ortogràfics en els SI per contacte amb determinades formes de SF, o per manca de SF.

Només presenta el SQF '-er-' per a la formació del futur i del condicional.

a)- Model: <VTOPC>

Atributs: CAT=VERB, TV=TORC, TV1=F(1),
T&=IN&1

SI: torc-, retorc-, estorc-, colltorc-,

b)- Model: <VTOR/>

Atributs: CAT=VERB, TV=TOR/, PART=UT

SI: torç-, retorç-, estorç-, colltorç-,

Regles d'anàlisi:

R1 START <VTORC> RAR2

R2 START <VTOR/> RAR2 per als SI

R3 RAR2 <IP/> RV
condició: TV=TOR/

R4 RAR2 <IP1B> RV
condició: TV=TORC per al present

R5 RAR2 <IIM> RV
condició: TV1=F per a l'imperfet

(1) Cada regla té, com a màxim, setze validacions. Com que les regles del futur, condicional i gerundi ultrapassen aquest límit, he donat l'atribut TV1=F a verbs que formen aquests temps amb els mateixos SF i així es redueix el nombre de validacions per a aquestes regles.

- R6 RAR2 <IPP> RV
condició: TV=TORC
- R7 RAR2 <IPP1A> RV
condició: TV=TORC per al pret.perf.
- R8 RAR2 <ITER> IF2
condició: TV=TORC
- R9 IF2 <IF> RV
condició: TV1=F per al futur
- R10 IF2 <IC> RV
condició: TV1=F per al condicional
- R11 RAR2 <SBPA> RV
condició: TV=TORC
- R12 RAR2 <SBPB> RV
condició: TV=TORC per al pres. de subj.
- R13 RAR2 <SBIM> RV
condició: TV=TORC per a l'imperf. de subj.
- R14 RAR2 <IMPEP> F
condició: TV=TOR/
- R15 RAR2 <IMP1> RV
condició: TV=TORC per a l'imperatiu
- R16 RAR2 <INF&1> RV
condició: T&=IN&1 per a l'infinitiu
- R17 RAR2 <GER> RV
condició: TV1=F per al gerundi
- R18 RAR2 <PARTUT> RV
condició: PART=UT per al participi

Paradigma P4: <VSAB> <VSAP>

Els models <VSAP> i <VSAB> resolen tan sols l'anàlisi dels verbs 'caber' i 'saber'. Aquests dos verbs presenten entre si algunes diferències a la primera persona del present d'indicatiu, a l'imperatiu i a l'infinitiu, que hem resolt amb atributs als SI (TV1=SAP, TV1=CAP, CAB=S). Ambdós verbs es caracteritzen per:

- a)- formar el futur i el condicional amb el SQF '-r-', model <ITR>;

b)- combinar-se amb el SQF '-igu-' al present de subjuntiu i a l'imperatiu, i amb el SQF '-ig-' per formar el participi.

En els temps assenyalats a l'apartat b), divergeixen de la resta de verbs de la segona conjugació, de manera que alguns models de SF els són exclusius. Les formes d'aquests temps s'haurien pogut resoldre també com a entrades directes al diccionari.

a)- Model: <VSAB>

Atributs: CAT=VERB, TV=SAB, TV1=F, PART=UT

SI:

sab-

cab- CAB=S

b)- Model: <VSAP>

Atributs: CAT=VERB, TV=SAP

SI:

sap- TV1=SAP

cap- TV1=CAP

Regles d'anàlisi:

R1 START <VSAB> RAR2

R2 START <VSAP> RAR2 per als SI

R3 RAR2 <PI1> RV
condició: CAB=S

R4 RAR2 <IPEP> RV
condició: TV=SAP

R5 RAR2 <IPEB> RV
condició: TV=SAB per al present
R6 RAR2 <IIM> RV
condició: TV1=F per a l'imperfet
R7 RAR2 <IPP> RV
condició: TV=SAB
R8 RAR2 <IPP1A> RV
condició: TV=SAB per al pret.perf.
R9 RAR2 <ITR> IF2
condició: TV=SAB
R10 IF2 <IF> RV
condició: TV1=F per al futur
R11 IF2 <IC> RV
condició: TV1=F per al condicional
R12 RAR2 <ITIG> RIG
condició: TV=SAP
R13 RAR2 <ITIGU> RIG
condició: TV=SAP
R14 RIG <SBIG&1> RV
condició: TV=SAP
R15 RIG <SBIG&A> RV
condició: TV=SAP
R16 RIG <SBIG> RV
condició: TV=SAP per al present de subj.
R17 RAR2 <SBIM> RV
condició: TV=SAB per al imperf. de subj.
R18 RGU2 <IMPSA&> RV
condició: TV1=SAP
R19 RGU2 <IMP2P> RV
condició: TV1=SAP
R20 RIG <IMP2C> RV
condició: TV=SAP
R21 RAR2 <IMPEP> F
condició: TV1=CAP
R22 RAR2 <IMP2P> RV
condició: CAB=S per a l'imperatiu
R23 RAR2 <INF> RV
condició: SAB=S
R24 IF2 <INFE> RV
condició: CAB=SI per a l'infinitiu
R25 RAR2 <GER> RV
condició: TV1=F per al gerundi
R26 RAR2 <PARTUT> RV
condició: PART=UT per al participi

Paradigma P5: <VREB> <VREP>

Els models <VREP> i <VREB> inclouen aquells verbs que canvien la '-b-' final del SI per una '-p-' quan es troba en posició final de mot o va seguida de consonant: cal donar dos SI per a l'anàlisi de les seves formes.

Només presenten el SQF '-r-' per a la formació del futur, condicional i infinitiu.

a)- Model: <VREB>

Atributs: CAT=VERB, TV=REB, TV1=F, PART=UT

SI: perceb-, conceb-, decsb-, aperceb-, reb-, etc.

b)- Model: <VREP>

Atributs: CAT=VERB, TV=REP

SI: percep-, concep-, decep-, apercep-, rep-, etc.

Regles d'anàlisi:

R1 START <VREB> RAR2
R2 START <VREP> RAR2 per als SI
R3 RAR2 <PI1> RV
condició: TV=REB
R4 RAR2 <IPEB> RV
condició: TV=REB
R5 RAR2 <IPEP> RV
condició: TV=REP per al present
R6 RAR2 <IIM> RV
condició: TV1=F per a l'imperfet
R7 RAR2 <IPP> RV
condició: TV=REB
R8 RAR2 <IPP1A> RV
condició: TV=REB per al pret.perf.

| | | | | |
|-----|------|----------|-----|---|
| R9 | RAR2 | <ITR> | IF2 | |
| | | | | condició: TV=REB |
| R10 | IF2 | <IF> | RV | |
| | | | | condició: TV1=F per al futur |
| R11 | IF2 | <IC> | RV | |
| | | | | condició: TV1=F per al condicional |
| R12 | RAR2 | <SBPA> | RV | |
| | | | | condició: TV=REB |
| R13 | RAR2 | <SBPB> | RV | |
| | | | | condició: TV=REB per al pres. de subj. |
| R14 | RAR2 | <SBIM> | RV | |
| | | | | condició: TV=REB per a l'imperf. de subj. |
| R15 | RAR2 | <IMPEP> | F | |
| | | | | condició: TV=REP |
| R16 | RAR2 | <IMP1> | RV | |
| | | | | condició: TV=REB per a l'imperatiu |
| R17 | RAR2 | <GER> | RV | |
| | | | | condició: TV1=F per al gerundi |
| R18 | RAR2 | <PARTUT> | RV | |
| | | | | condició: PART=UT per al participi |
| R19 | IF2 | <INFE> | RV | |
| | | | | condició: TV=REB per a l'infinitiu |

Paradigma P6: <VOLER>

Aquest paradigma correspon tan sols al verb 'voler'. Les diferents formes de SI s'han inclòs en un mateix model, <VOL>. Cada forma de SI té associats els atributs corresponents a les condicions de les regles de SF.

La forma 'vull', corresponent a la primera persona del present d'indicatiu, com que no té paral·lelisme amb cap altre verb, l'hem resolta com a entrada directa al model <ED>, amb els atributs corresponents:

<ED>

vull CAT=VERB, TEMP=PRES,
MODE=IND, PERS=1, NBRE=SG

Aquest verb es caracteritza per :

- a)- presentar l'alternança 'vol-/vul-' en els SI;
- b)- combinar-se amb el SQF '-gu-' per formar el pretèrit perfet, tot el subjuntiu i tot l'imperatiu;
- c)- combinar-se amb el SQF '-dr-' per al futur i el condicional;
- d)- combinar-se amb el SQF '-g-' per formar el participi.

Model: <VOL>

Atributs: CAT=VERB

SI:

vol- TV=VOL, TV1=T

vul- TV=VUL

Regles d'anàlisi:

- R1 START <VOL> RAR2
per als SI
- R2 RAR2 <IPCA> F
condició: TV=VOL
- R3 RAR2 <IPB> F
condició: TV=VOL per al present
- R4 RAR2 <IIM> RV
condició: TV1=F
- R5 RAR2 <ITGU> RGU2
condició: TV=VOL, TV=VUL
- R6 RGU2 <IPP> RV
condició: TV=VOL
- R7 RGU2 <IPP1A> RV
condició: TV=VOL per al pret. perf.
- R8 RAR2 <ITDR> IF2
condició: TV=VOL
- R9 IF2 <IF> RV
condició: TV1=F per al futur

| | | | | |
|-----|------|----------|------------------|-------------------------|
| R10 | IF2 | <IC> | RV | |
| | | | condició: TV1=F | per al condicional |
| R11 | RGU2 | <SBPA> | RV | |
| | | | condició: TV=VUL | |
| R12 | RGU2 | <SBPB> | RV | |
| | | | condició: TV=VUL | per al pres. de subj. |
| R13 | RGU2 | <SBIM> | RV | |
| | | | condició: TV=VUL | per al imperf. de subj. |
| R14 | RGU2 | <IMPDI> | RV | |
| | | | condició: TV=VUL | |
| R15 | RGU2 | <IMP1> | RV | |
| | | | condició: TV=VUL | per a l'imperatiu |
| R16 | RAR2 | <INF> | RV | |
| | | | condició: TV=VOL | per a l'infinitiu |
| R17 | RAR2 | <GER> | RV | |
| | | | condició: TV1=F | per al gerundi |
| R18 | RAR2 | <ITG> | RIP2 | |
| | | | condició: TV=VOL | |
| R19 | RIP2 | <PARTUT> | RV | |
| | | | condició: TV=VOL | per al participi |

Paradigma P7: <VPOD>

El model <VPOD> inclou un únic verb, 'poder'. Per tal de poder utilitzar models de SF ja definits, l'anàlisi del verb requereix la definició dels SI: 'pod-', 'pu-', 'po-' i 'pot-'.

El verb 'poder' es caracteritza:

a)- per combinar-se amb el SQF '-gu-' per formar el pretèrit perfet, el present i l'imperfet de subjuntiu i algunes formes de l'imperatiu:

'po-gu-i', 'pu-gu-i', 'po-gu-és', etc.

b)- per combinar-se amb el SQF '-r-' per a la formació del futur i del condicional: '-pod-r-é', 'pod-r-ia';

c)- per formar el participi segons el model regular de la segona conjugació, <PARTUT>, amb el SQF '-g-':

'po-g-ut', 'po-g-uda', 'po-g-uts', 'po-g-udes'.

Model: <VPOD>

Atributs: CAT=VERB

SI:

pcd- TV=POD, TV1=F

pot- TV=POT

pu- TV=PU

po- TV=PO

Regles d'anàlisi:

| | | | | |
|-----|-------|------------------------|------|--------------------|
| R1 | START | <VPOD> | RAR2 | |
| | | | | per als SI |
| R2 | RAR2 | <IPIC> | RV | |
| | | condició: TV=PU | | |
| R3 | RAR2 | <IPEP> | RV | |
| | | condició: TV=POT | | |
| R4 | RAR2 | <IPEB> | RV | |
| | | condició: TV=POD | | per al present |
| R5 | RAR2 | <IIM> | RV | |
| | | condició: TV1=F | | per a l'imperfet |
| R6 | RAR2 | <ITGU> | RGU2 | |
| | | condició: TV=PO, TV=PU | | |
| R7 | RGU2 | <IPP> | RV | |
| | | condició: TV=PO | | |
| R8 | RGU2 | <IPP1A> | RV | |
| | | condició: TV=PO | | per al pret. perf. |
| R9 | RAR2 | <ITR> | IF2 | |
| | | condició: TV=POD | | |
| R10 | IF2 | <IF> | RV | |
| | | condició: TV1=F | | per al futur |
| R11 | IF2 | <IC> | RV | |
| | | condició: TV1=F | | per al condicional |
| R12 | RGU2 | <SBPA> | RV | |
| | | condició: TV=PU | | |

| | | | | |
|-----|------|------------------|------|-------------------------|
| R13 | RGU2 | <SBPB> | RV | |
| | | condició: TV=PU | | per al pres. de subj. |
| R14 | RGU2 | <SBIM> | RV | |
| | | condició: TV=PO | | per al imperf. de subj. |
| R15 | RGU2 | <IMPDI> | RV | |
| | | condició: TV=PU | | |
| R16 | RGU2 | <IMP1> | RV | |
| | | condició: TV=PU | | per a l'imperatiu |
| R17 | RAR2 | <INF> | RV | |
| | | condició: TV=POD | | per a l'infinitiu |
| R18 | RAR2 | <GER> | RV | |
| | | condició: TV1=F | | per al gerundi |
| R19 | RAR2 | <ITG> | RIP2 | |
| | | condició: TV=PO | | |
| R20 | RIP2 | <PARTUT> | RV | |
| | | condició: TV=PO | | per al participi |

Paradigma PB: <VDRE> <VES> <VES&> <VES&1>

Els verbs del paradigma <VDRE> <VES> <VES&> <VES&1> es caracteritzen perquè:

a)- formen l'infinitiu, el futur i el condicional amb el SQF '-dr-': apren-dr-e, dol-dr-e, mol-dr-e, etc;

b)- formen el pretèrit perfet, el present de subjuntiu i algunes formes de l'imperatiu amb els SQF '-gu-': 'apren-gu-i', 'mol-gu-i', 'dol-gu-i', etc;

El model <VDRE> inclou verbs que presenten certes divergències que no hem considerat suficientment importants com per definir diferents paradigmes, ja que:

- el nombre de verbs que s'aplegarien seria molt reduït;

- no hi ha perspectives de formació de nous verbs segons aquests models de flexió;

- els verbs de <VDRE> tenen prou característiques comunes com per tractar-los conjuntament.

Les diferències entre els verbs de <VDRE> són les següents:

a)- La presència/absència d'accent, greu o agut, a la tercera persona del present d'indicatiu i a la segona persona de l'imperatiu:

'ven', 'pren', 'fon', 'descompon', 'defèn',
expèn', 'ofèn', 'distèn', 'encén', 'átèn', etc.

Les formes sense accent porten al SI l'atribut BLP=S per validar la regla del SF, <IPCA>:

recompon- BLP=S,

Les formes que tenen accent agut validen la regla del SF, model <IPC&>, amb l'atribut T&=RE& al SI:

conten- T&=RE&

Finalment, les formes que tenen accent greu porten al SI l'atribut T&=RE&1 per validar la regla del SF, model <IPC&1>:

defen- T&=RE&1

b)- Diferents maneres de formar el participi:

1- Mitjançant el SQF '-g-', model <ITG>, seguit dels SF '-ut, -uda', model <PARTUT>:

| | |
|-------|------------------------|
| <ITG> | <PARTUT> |
| -g- | CAT=VERB FNP=PART |
| | -ut GEN=MASC, NBRE=SG |
| | -uda GEN=FEM, NBRE=SG |
| | -uts GEN=MASC, NBRE=PL |
| | -udes GEN=FEM, NBRE=PL |

p.e.: 'dol-g-ut', 'val-g-ut', etc.

L'atribut TV1=DOL al SI permet validar la regla del SQF '-g-'.

2- Mitjançant el SF '-ut, -uda, -uts, -udes', model <PARTUT>:

p.e.: 'ven-ut', 'carven-ut', etc.

L'atribut PART=UT als SI permet la validació de la regla del model de SF <PARTUT>, que en aquest cas té RAR2 com a estat inicial.

3- Mitjançant el SQF, '-t-', model <ITT>, seguit dels SF de gènere i nombre.

L'atribut per validar la regla del SQF és TV1=T;

resol- TV1=T

El model <ITT> té RAR2 com a estat inicial i INTP com a estat final. L'atribut per validar la regla del model dels SF de gènere i nombre és a' model <ITT>:

| | |
|----------|----------------------|
| <ITT> | <GN> |
| CAT=VERB | -0 GEN=MASC, NBRE=SG |
| FNP=PART | -a GEN=FEM, NBRE=SG |
| -t- | -s GEN=MASC, NBRE=PL |
| | -es GEN=FEM, NBRE=PL |

p.e.: 'resol-t-0', 'absol-t-0', etc.

4- Mitjançant un SI germà. Hem definit tres models de SI germans segons els SF de gènere i nombre que accepten. Aquests models de SI germans porten l'atribut per validar les regles de gènere i nombre que els corresponen:

| | | |
|----------|----------|----------|
| <VES> | <VES&> | <VES&1> |
| TG=GN | TG=GN& | TG=GN&1 |
| CAT=VERB | CAT=VERB | CAT=VERB |
| FNP=PART | FNP=PART | FNP=PART |
| post- | confos- | romas- |
| pres- | infos- | contes- |
| etc. | | |

El model <VES> correspon als SI germans sense accent, <VES&> als SI amb accent agut, i

<VES&i> als SI amb accent greu.

a)- Model: <VDRE>

Atributs: CAT=VERB, TV=DRE

SI:

ven- BLP=S, PART=UT

dol- BLP=S, TV1=DOL

condol- BLP=S, TV1=DOL

sol- BLP=S, TV1=DOL

tol- BLP=S, TV1=T

absol- BLP=S, TV1=T

col- BLP=S, TV1=T

dissol- BLP=S, TV1=T

reabsol- BLP=S, TV1=T

resol- BLP=S, TV1=T

mol- BLP=S, TG=GN4

pon- BLP=S

correspon- BLP=S

val- BLP=S

descompon- BLP=S

compon- BLP=S

recompon- BLP=S

respon- BLP=S

fon- BLP=S

ton- BLP=S

confon- BLP=S

infon- BLP=S

marfon- BLP=S

reton- BLP=S

refon- BLP=S

transfon- BLP=S

difon- BLP=S

roman- BLP=S

conten- T&=RE&

desaten- T&=RE&

preten- T&=RE&

desenten- T&=RE&

disten- T&=RE&

encen- T&=RE&

enten- T&=RE&

esten- T&=RE&

malenten- T&=RE&

sobreenten-T&=RE&

| | |
|------------|------------------|
| aten- | T&=RE& |
| retroven- | T&=RE&1, PART=UT |
| reven- | T&=RE&1, PART=UT |
| malven- | T&=RE&1, PART=UT |
| carven- | T&=RE&1, PART=UT |
| defen- | T&=RE&1 |
| depen- | T&=RE&1 |
| aprehen- | T&=RE&1 |
| despen- | T&=RE&1 |
| expen- | T&=RE&1 |
| ofen- | T&=RE&1 |
| suspen- | T&=RE&1 |
| apren- | T&=RE&1 |
| compren- | T&=RE&1 |
| corpren- | T&=RE&1 |
| desapren- | T&=RE&1 |
| empren- | T&=RE&1 |
| malpren- | T&=RE&1 |
| mampren- | T&=RE&1 |
| repren- | T&=RE&1 |
| salpren- | T&=RE&1 |
| sobrepren- | T&=RE&1 |
| sorpren- | T&=RE&1 |
| ullpren- | T&=RE&1 |
| despren- | T&=RE&1 |

Regles d'anàlisi:

| | | | | |
|-----|-------|-----------|---------|------------------|
| R1 | START | <VDRE> | RAR2 | |
| R2 | START | <VES> | INTP | |
| R3 | START | <VES&> | INTP | |
| R4 | START | <VES&1> | INTP | per als SI |
| R5 | RAR2 | <IPIC> | RV | |
| | | condició: | TV=DRE | |
| R6 | RAR2 | <IPCA> | F | |
| | | condició: | BLP=S | |
| R7 | RAR2 | <IPC&> | F | |
| | | condició: | T&=RE& | |
| R8 | RAR2 | <IPC&1> | F | |
| | | condició: | T&=RE&1 | |
| R9 | RAR2 | <IPP> | RV | |
| | | condició: | TV=DRE | per al present |
| R10 | RAR2 | <IIM> | RV | |
| | | condició: | TV=DRE | per a l'imperfet |
| R11 | RAR2 | <ITGU> | RGU2 | |
| | | condició: | TV=DRE | |
| R12 | RGU2 | <IPP> | RV | |
| | | condició: | TV=DRE | |

R13 RGU2 <IPP1A> RV
 condició: TV=DRE per al pret. perf.
 R14 RAR2 <ITDR> IF2
 condició: TV=DRE
 R15 IF2 <IF> RV
 condició: TV=DRE per al futur
 R16 F2 <IC> RV
 condició: TV=DRE per al condicional
 R17 RGU2 <SBPA> RV
 condició: TV=DRE
 R18 RGU2 <SBPB> RV
 condició: TV=DRE per al pres. de subj.
 R19 RGU2 <SBPM> RV
 condició: TV=DRE per al imperf. de subj.
 R20 RAR2 <IIF> F
 condició: TV=DRE
 R21 RAR2 <IIP> RV
 condició: TV=DRE
 R22 RAR2 <IMPA&1> RV
 condició: TV=DRE
 R23 RGU2 <IMPB> RV
 condició: TV=DRE per a l'imperatiu
 R24 IF2 <INFE> RV
 condició: TV=DRE
 R25 RAR2 <INF> RV
 condició: TV1=DOL per a l'infinitiu
 R26 RAR2 <GER> RV
 condició: TV=DRE per al gerundi
 R27 RAR2 <ITG> RIP2
 condició: TV1=DOL
 R28 RAR2 <PARTUT> RV
 condició: PART=UT
 R29 INTP <GN> RV
 condició: TG=GN
 R30 INTP <GN&1> RV
 condició: TG=GN&1
 R31 INTP <GN&1> RV
 condició: TG=GN&1
 R32 INTP <GN4> RV
 condició: TG=GN4
 R33 RAR2 <ITT> INTP
 condició: TV1=T per al participi

Paradigma P9: <VTRE> <VES> <VES&1>

El model <VTRE> es combina amb els mateixos models de SF que <VDRE>. <VTRE> i <VDRE> es diferencien perquè:

a)- <VDRE> forma el futur i el condicional amb el SQF '-dr-' i <VTRE> amb el SQF '-r-', model <ITR>.

b)- <VTRE> no es combina amb cap altre SQF.

Alguns verbs de <VTRE> formen el participi amb un SI germà. Els models que apleguen els SI del participi són:

- <VES> per a les formes sense accent: 'mes', 'mesa', etc.

- <VES&i> per a les formes amb accent greu: 'admès', 'promès', etc.

La resta de verbs forma el participi amb els SF '-ut', '-uda', '-uts', '-udes', inclosos al model <PARTUT>. L'atribut PART=UT als SI serveix per complir la condició de la regla d'aquest model.

a)- Model: <VTRE>

Atributs: CAT=VERB, TV=TRE, T&=IM&

SI:

admet-
arremet-
compromet-
emet-
malmet-
promet-

met-

bat- PART=UT
debat- PART=UT

rebat- PART=UT

bot- PART=UT

rebot- PART=UT

pard- PART=UT

ret- PART=UT

romp- PART=UT

irromp- PART=UT

b)- Model: <VES&1>

Atributs: CAT=VERB, FNP=PART, TG=GN&1

SI:

admes-

emes-

malmes-

permes-

remes-

etc.

c)- Model: <VES>

Atributs: CAT=VERB, FNP=PART, TG=GN

SI:

mes-

Regles d'anàlisi:

R1 START <VTRE> RAR2

R2 START <VES> INTP

R3 START <VES&1> INTP per als SI

R4 RAR2 <PI1> RV

condició: TV=TRE

R5 RAR2 <IPCA> F

condició: TV=TRE

R6 RAR2 <IPB> RV

condició: TV=TRE per al present

R7 RAR2 <IIM> RV

condició: TV=TRE per a l'imperfet

R8 RAR2 <IPP> RV

condició: TV=TRE

| | | | | |
|-----|------|----------|-----|---|
| R9 | RAR2 | <IPP1A> | RV | |
| | | | | condició: TV=TRE per al pret. perf. |
| R10 | RAR2 | <ITR> | IF2 | |
| | | | | condició: TV=TRE |
| R11 | IF2 | <IF> | RV | |
| | | | | condició: TV=TRE per al futur |
| R12 | IF2 | <IC> | RV | |
| | | | | condició: TV=TRE per al condicional |
| R13 | RAR2 | <SBPA> | RV | |
| | | | | condició: TV=TRE |
| R14 | RAR2 | <SBPB> | RV | |
| | | | | condició: TV=TRE per al pres. de subj. |
| R15 | RAR2 | <SBIM> | RV | |
| | | | | condició: TV=TRE per a l'imperf. de subj. |
| R16 | RAR2 | <IMPA> | RV | |
| | | | | condició: TV=TRE |
| R17 | RAR2 | <IMPB> | RV | |
| | | | | condició: TV=TRE per a l'imperatiu |
| R18 | IF2 | <INFE> | RV | |
| | | | | condició: TV=TRE per a l'infinitiu |
| R19 | RAR2 | <GER> | RV | |
| | | | | condició: TV=TRE per al gerundi |
| R20 | RAR2 | <PAKTUT> | RV | |
| | | | | condició: PART=UT |
| R21 | INTP | <GN> | RV | |
| | | | | condició: TG=GN |
| R22 | INTP | <GN&1> | RV | |
| | | | | condició: TG=GN&1 per al participi |

Els verbs que descrivim a continuació es caracteritzen perquè presenten tres formes de SI. Així, els paradigmes de flexió d'aquests verbs inclouen tres models. Dos d'aquests models són comuns a tots els paradigmes, que es diferencien per un tercer model.

<VCRE> i <VCREU> són els models comuns. <VCREI>, <VQUE>, <VBEV> i <VESCRV> són els models de SI que determinen les diferències i defineixen el quatre paradigmes de flexió.

Vegem una mostra representativa dels SI d'aquests verbs:

- 1- 'cre-' 'escri-' 'be-' 'cai-', model <VCRE>;
- 2- 'creu-, 'escriu-' 'beu-, 'cau-', model <VCREU>;
- 3- 'crei-' , model <VCREI>,
'escriv-', model <ESCRIV>,
'bev-', model <VBEV>
'que-', model <VQUE>.

Els paradigmes són:

- Paradig.1 <VCRE> <VCREU> <VCREI>
Paradig.2 <VCRE> <VCREU> <VSCRIV>
Paradig.3 <VCRE> <VCREU> <VBEV>
Paradig.4 <VCRE> <VCREU> <VQUE>

Aquests verbs presenten les característiques comunes següents:

a)- Les formes del model <VCRE> es combinen amb el SDF '-gu-' per formar el pretèrit perfet, el present de subjuntiu i algunes formes de l'imperatiu.

b)- El model <VCREU> permet l'anàlisi del futur, el condicional i d'algunes formes del present d'indicatiu i de l'imperatiu.

El futur i condicional es formen amb el SQF '-r-',
model <IYF>, seguit dels models de SF.

Els SI de <VCRE> presenten algunes diferències en la
seva flexió que hem resolt mitjançant atributs als SI
per validar les regles corresponents.

Paradigma 1 del grup <VCRE> <VCREU>:

Paradigma P10: <VCRE> <VCREU> <VCREZ>

Els verbs d'aquest paradigma es caracteritzen perquè:

a)- tenen un accent al SI de la primera i segona
persones del plural de l'imperfet. Les formes de
l'imperfet sense accent, model de SF <IIMFE>, estan
validades amb l'atribut TV1=CRE, les formes
accentuades tenen l'atribut T&=IM&1;

b)- es combinen amb el SQF '-g-' per formar el
participi. L'atribut TV1=CRE permet la validació
d'aquestes formes.

a)- Model: <VCRE>

Atributs: CAT=VERB, TV=CRE

SI:

se- TV1=CRE, T&=IM&1
sobrese- TV1=CRE, T&=IM&1
cre- TV1=CRE, T&=IM&1
descre- TV1=CRE, T&=IM&1
ile- TV1=CRE, T&=IM&1

b)- Model: <VCREU>

Atributs: CAT=VERB, TV=CREU

SI:

seu-
sobreseu-
creu-
descreu-
lleu-

c)- Model: <VCREI>

Atributs: CAT=VERB, TV=CREI

SI:

sai-
sobresai-
crei-
descrei-
lleai-

Regles d'anàlisi:

- R1 START <VCRE> RAR2
R2 START <VCREU> RAR2
R3 START <VCREI> RAR2 per als SI
R4 RAR2 <IPIC> RV
condició: TV=CRE
R5 RAR2 <IPCA> RV
condició: TV=CREU
R6 RAR2 <IPB> RV
condició: TV=CREI per al present
R7 RAR2 <IIMFE> RV
condició: TV=CRE
R8 RAR2 <IIM&1> RV
condició: T=IM&1 per a l'imperfet
R9 RAR2 <ITGU> RBU2
condició: TV=CRE
R10 RBU2 <IPP> RV
condició: TV=CRE
R11 RBU2 <IPP1A> RV

| | | | |
|-----|------|-------------------|--------------------------|
| | | condició: TV=CRE | per al pret. perf. |
| R12 | RAR2 | <ITR> | IF2 |
| | | condició: TV=CREU | |
| R13 | IF2 | <IF> | RV |
| | | condició: TV=CREU | per al futur |
| R14 | IF2 | <IC> | RV |
| | | condició: TV=CREU | per al condicional |
| R15 | RGU2 | <SBPA> | RV |
| | | condició: TV=CRE | |
| R16 | RGU2 | <SBPB> | RV |
| | | condició: TV=CRE | per al pres. de subj. |
| R17 | RGU2 | <SBIM> | RV |
| | | condició: TV=CRE | per a l'imperf. de subj. |
| R18 | RGU2 | <IMPB> | RV |
| | | condició: TV=CRE | |
| R19 | RAR2 | <IMPEP> | F |
| | | condició: TV=CREU | |
| R20 | RAR2 | <IMP2P> | RV |
| | | condició: TV=CREI | per a l'imperatiu |
| R21 | IF2 | <INFE> | RV |
| | | condició: TV=CREU | per a l'infinitiu |
| R22 | RAR2 | <GER> | RV |
| | | condició: TV=CREI | per al gerundi |
| R23 | RAR2 | <ITG> | RIP2 |
| | | condició: TV=CRE | |
| R24 | RIP2 | <PARTUT> | RV per al participi |

Paradigma 2 del grup <VCRE> <VCREU>:

Paradigma P11: <VCRE> <VCREU> <VESCRV>

Aquest paradigma resol l'anàlisi de verbs com 'viure' i 'escriure' que, encara que presenten un nombre important de característiques comunes, es diferencien:

a)- en la formació del pretèrit perfet: 'escriv-i' i 'vi-squ-i'. El verb 'escriure' i els seus derivats formen el pretèrit perfet a partir del SI del model <VESCRV> que es combina amb un model de SF exclusiu, <IPP3>. L'anàlisi d'aquestes formes es resol amb l'atribut CRIV=SI als SI.

El verb 'viure' i els seus derivats resolen l'anàlisi del pretèrit perfet a partir del SI contingut al model <VCRE>. Els segments inicials d'aquests verbs tenen l'atribut TV=VI, que valida la regla del SQF '-squ-' amb què forma el pretèrit perfet.

b)- en la formació del participi: 'escrit' i 'vi-sc-ut'. El participi es resol a partir dels SI de <VCRE>. L'atribut TV=VI als SI del verb 'viure' i els seus derivats valida el SQF '-sc-' que permet l'anàlisi d'aquesta forma. L'atribut TVI=T als SI del verb 'escriure' i els seus derivats valida la regla del SQF '-t-'. Aquest atribut s'ha utilitzat a d'altres paradigmes amb la mateixa finalitat.

c)- en la formació del subjuntiu: 'escri-gu-i', 'escriu-ís' i 'vi-squ-i', 'vi-squ-és'.

Les formes del present de subjuntiu del verb 'escriure' es formen a partir del SI de <VCRE>. L'atribut del model TV=CRE valida les regles del model de SF. Les formes de l'imperfet de subjuntiu del verb 'escriure' es formen amb els SI del model <VSCRIV> degudament validats amb l'atribut CRIV=S.

El verb 'viure' i els seus derivats formen el subjuntiu amb el SI de <VCRE>. L'atribut del SI TV=VI valida la regla del SQF '-squ-' amb què es forma el present i l'imperfet de subjuntiu.