

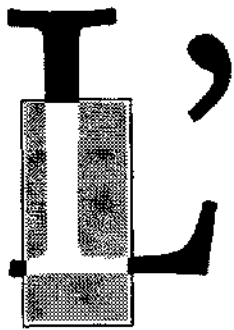
---

**TONI BADIA**

---

**LEXICOGRAFIA  
I MODELS LINGÜÍSTICS:  
LES TEORIES LINGÜÍSTIQUES  
I EL LÈXIC**

---



objectiu fonamental d'aquest treball és posar de manifest com la teoria lingüística, qualsevol teoria lingüística, comporta una visió particular sobre el lèxic i, per tant, proposa unes tasques lexicogràfiques determinades. Per a això, ens aturarem en alguns dels models teòrics de la lingüística d'aquest segle i n'observarem la visió del lèxic. Els models escollits no són, evidentment, els únics que ens haurien pogut servir: la proliferació de models lingüístics teòrics és una de les característiques de l'evolució d'aquesta ciència durant aquest segle. Ara bé, la seva successió representa prou adequadament els grans moments pels quals ha passat la lingüística en la seva història moderna. Així doncs, després d'uns mots breus sobre el que hem convingut d'anomenar gramàtica tradicional, començarem el nostre recorregut en l'estructuralisme, per passar després a la gran família de teories generativistes; primer observarem tres moments de l'evolució del pensament de Chomsky (la teoria prestàndard, la teoria estàndard i la teoria de la recció i el lligam), i després ens aturarem a dues altres teories generatives sorgides dels ambients més directament influïts per la lingüística computacional (la gramàtica lèxico-funcional i la gramàtica d'estructura sintagmàtica orientada al nucli).

## 1. INTRODUCCIÓ

Si acceptem que els plantejaments científics tenen lloc en la lingüística i que el seu objectiu fonamental és aconseguir una explicació del comportament lingüístic humà, haurem de convenir que allò que pertoca és una caracterització global d'aquest comportament. De fet, les teories lingüístiques tenen com a objectiu l'oferir aquesta descripció global, en què cada component troba el seu lloc en relació als altres. D'aquesta manera, en cada una d'elles, el lèxic ocupa el lloc que li correspon, al costat dels altres elements que la conformen. Per tant, no és estrany que en cada una el lèxic hi tingui un paper diferent: no es tracta d'una posició absoluta, sinó relativa al conjunt de la teoria, o del sistema que aquesta proposa. Així, un dels criteris usats en la classificació de les teories lingüístiques és el pes relatiu que atorguen al lèxic en el sistema global que proposen (Droste i Joseph, 1991:19).

No hi ha dubte que qualsevol proposta de sistema lingüístic ha de concedir un lloc al lèxic, al diccionari. Ara bé, aquest pot ser simplement el mínim necessari (com en les primeres versions de la Gramàtica generativa transformacional –vegeu, més avall, §6), o, en canvi, pot constituir una peça essencial en el sistema proposat (com en la Gramàtica lèxico-funcional –vegeu §8). Aquestes dues teories són probablement dos extrems en la dicotomia entre lèxic i gramàtica. Ara bé, serveixen per a mostrar d'una manera força gràfica com hi ha una relació inversa entre la quantitat de feina atorgada al lèxic i la que correspon a la resta de la gramàtica. Així, en els primers estadis de la gramàtica generativa les regles gramaticals ocupen molts dels espais que en la gramàtica lèxico-funcional ocupa el lèxic. A la vegada, també mostren (en dues versions distintes) com les propietats del lèxic són una conseqüència de les característiques de la teoria lingüística en general. D'aquesta manera, una teoria en la qual tots els elements sistemàtics són formulats en regles exteriors al lèxic no permet que aquest tingui cap altre paper que el d'un veritable calaix de sastre (on es recullen precisament els aspectes que no han pogut ser sotmesos a la sistematicitat). En canvi, una teoria en la qual es preveuen dos nivells de sistematicitat diferents, l'un exterior i l'altre interior al lèxic, està imposant unes condicions al lèxic quant a la seva estructura i al seu enllaç amb els altres elements de la gramàtica.

En el paràgraf anterior hem partit d'una visió teòrica del llenguatge, en la qual ocupa un lloc principal la distribució de la feina entre els diversos components de la teoria. És una característica general de les aproximacions teòriques al llenguatge de proposar un model explicatiu en el qual hi ha diverses estructures independents però interrelacionades (anomenades mòduls). En aquesta visió, el lèxic és un d'aquests mòduls, enfront dels altres components de la gramàtica. Ara bé, és només entrat aquest segle que comencen a sorgir visions teòriques del llenguatge en què es posa a la pràctica una estructuració així. Però, fins i tot ara, a les acaballes del segle xx, el que solem anomenar *gramàtica tradicional*, que té les seves arrels en una mala popularització

de la gramàtica clàssica i medieval, encara perdura en les nostres aules i publicacions. Així doncs, encara que probablement no és massa correcte, sembla procedent de fer una petita menció a la situació del lèxic en la gramàtica tradicional.

## 2. LA GRAMÀTICA TRADICIONAL

Entenem per gramàtica tradicional aquelles visions del llenguatge sense sistematitzar, en les quals es juxtaposen explicacions d'ordre diferent de manera gens rigorosa, barrejant, per exemple, descripcions de tipus semàntic amb característiques morfològiques. Com que, a més, hi manca una perspectiva global del fenomen lingüístic, aquests diferents elements no troben el seu lloc en un marc teòric general i són usats, simplement, en funció de la conveniència del redactor de la gramàtica. Per altra banda, la preocupació fonamental d'aquestes aproximacions gramaticals és determinar els usos lingüístics correctes (enfront dels barbarismes, localismes, vulgarismes...), de manera que es converteixen a la pràctica en manuals prescriptius del bon ús de la llengua. Com que gairebé només interessa establir una normativa, a la quals els parlants (de fet, els escriptors) d'una llengua s'haurien de sotmetre, en aquests manuals, els estudis lingüístics hi tenen una importància relativa.

Aquestes dues característiques que acabem de mencionar (la manca de perspectiva global i la finalitat únicament normativista) fan que molt sovint la gramàtica tradicional caigui en plantejaments absolutament casuístics: es tracta de caracteritzar els usos correctes dels mots i expressions i de distingir-los dels incorrectes, i de fer-ho de manera que pugui ser utilitzada per persones no entrenades especialment per a l'anàlisi lingüística. El resultat n'és un model gramatical no sistemàtic, centrat en els aspectes més fàcilment delimitables de la llengua.

No és estrany que en un model d'aquestes característiques el tema de la relació entre el lèxic i la gramàtica quedi molt difús. En primer lloc, perquè la mateixa noció de gramàtica no hi és gaire clara i, en segon lloc, perquè es fonamenta sempre en els aspectes més superficials de la llengua. Així, el centre d'atenció fonamental és la paraula. Des d'aquesta perspectiva, els aspectes que poden ser tractats més fàcilment són l'ortografia, la morfologia i la semàntica lèxica, que correspondrien, les primeres, a la forma de les paraules i, la segona, al seu significat. I aquesta és precisament la distribució de la feina entre les gramàtiques i els diccionaris: les primeres determinen les normes d'ortografia i les formes morfològiques de les paraules, i els segons fixen els seus significats. Com que les relacions entre les paraules en les emissions lingüístiques no són estudiades de manera sistemàtica, els pocs elements sintàctics o estructurals que s'hi tracten ho són en contextos de paraula. La paraula, de fet, queda convertida en la unitat d'anàlisi lingüística, de manera que els estudis al voltant del lèxic són els que ocupen la major part de les pàgines de les gramàtiques tradicionals. El punt d'enllaç



entre la informació proporcionada per les gramàtiques i els diccionaris és ben dèbil: la breu categorització dels mots en el diccionari segueix els criteris classificatoris proposats per les gramàtiques.

La tasca de la lexicografia, doncs, queda limitada a la recopilació de les paraules de la llengua, a la seva fixació i a la determinació dels seus usos (sentits) correctes, i a la seva categorització d'acord amb els pocs principis fixats per la gramàtica.

### 3. L'ESTRUCTURALISME

Des dels primers intents d'elaboració d'una teoria del llenguatge, apareix la necessitat de buscar una estructura que superi el marc estricte de les paraules. L'observació que hi ha aspectes de la llengua (tant sintàctics, com semàntics) que no poden ser explicats des d'una perspectiva lèxica únicament ha dut els lingüistes a plantejar la necessitat d'una estructura per sobre i més enllà de les paraules. La història de la lingüística en aquest segle és, de fet, la història de la sistematització de les estructures lingüístiques, amb la finalitat de poder descriure i predir els fets lingüístics amb prou objectivitat i amb prou caràcter explicatiu. En aquest procés, el primer estadi és el representat per l'estructuralisme. Amb ell, podríem dir que neix la lingüística com a ciència, ja que és quan, per primer cop, es disposa d'un aparell teòric rigorós per a començar a emprendre seriosament l'anàlisi lingüística.

És clau en el plantejament estructuralista la caracterització de les relacions que els diversos elements de la llengua poden contreure entre ells, les quals són agrupades en dos blocs, el de les relacions sintagmàtiques i el de les paradigmàtiques (Saussure, 1915). Com ja és prou conegut, les primeres són les que estableixen entre ells els elements en la cadena fònica, mentre que les segones són les que tenen amb altres elements del sistema (que podrien ocupar el seu lloc en la cadena fònica, que s'hi oposen, que s'hi complementen, etc.).

La caracterització d'aquestes relacions es fa a partir de dues operacions fonamentals: la segmentació i la classificació (Bierwisch, 1966). El procés de segmentació s'aplica a les cadenes fòniques de manera que s'hi descobreixen els elements que les conformen; com és obvi, els resultats d'aquest procés es poden trobar en nivells diferents de descripció lingüística (fonemes i síl·labes, morfemes, paraules, sintagmes, etc.). Aquests elements són després classificats segons les seves propietats formals i funcionals. En aquest procés de classificació hi juguen un paper principal les posicions relatives que ocupen els diversos elements en l'estructura abstracta formada per tots ells, en el sistema que és la llengua. Naturalment els criteris rellevants per a la classificació són determinats en gran part pel nivell de descripció que es pretén aconseguir: així, si hom pretén obtenir una descripció de tipus fonològic, es tindran en compte factors d'ordre fònic, mentre que per a una descripció morfosintàctica, s'hauran de tenir en compte sobretot factors de tipus sintàctic.

Així doncs, el valor relatiu d'un element en el sistema és el resultat del conjunt de relacions sintagmàtiques i paradigmàtiques que contreu. A més, la mateixa porció física de la cadena fònica pot ser subjecte de diferents tipus de relacions, segons el nivell de descripció que es tingui en compte. Sens dubte, des del punt de vista estructuralista, qualsevol fragment de la cadena fònica (simple o complex) pot ser un element subjecte a diversos tipus de relacions. Ara bé, també és cert que la paraula hi ocupa una posició central, ja que és a partir d'aquesta noció que es caracteritzen completament els conceptes més adequats de *morfema*, *lexema*, *sintagma*, etc. En gran part, és el tractament semàntic el que provoca aquest fet. Vegem-ho amb una mica de detall, ja que és a la base de la concepció estructuralista del diccionari.

El signe lingüístic té un vessant formal i un altre de semàntic, està format per elements físics (els fonemes) que representen elements no físics (els conceptes, en la visió estructuralista majoritària). Aquesta visió del signe lingüístic determina fortament el pes que tindran les paraules en la descripció general de la llengua. Així, trobem en el pla del contingut el mateix biaix que hi ha en el pla de l'expressió, ja que en ambdós es dóna preferència a la descripció dels elements i a la seva classificació: així com les relacions entre els fonemes en la cadena fònica no són explicades a fons (Bierwisch, 1966), el tractament semàntic estructuralista es redueix, gairebé sempre, a una semàntica de la paraula. Que aquest paral·lelisme va ser entès d'una manera quasi literal ho prova el fet que es va intentar aplicar de seguida a la semàntica de les paraules el mètode de la descomposició en trets essencials i universals que tant d'èxit havia tingut en fonologia. És cert que molt aviat apareix la divisió entre *lexema* i *morfema*, de manera que resulta possible de distingir, en una paraula, entre els factors pròpiament de contingut i els que determinen l'enllaç de les paraules entre elles. Però aquests elements són ja les unitats mínimes en què es pot descompondre la cadena fònica des del punt de vista del contingut; qualsevol pas posterior en l'anàlisi es fa per un procés de descomposició en trets que no és composicional en el sentit modern del terme.<sup>1</sup> El camí que hauria pogut obrir la porta de l'estudi de les relacions semàntiques horitzontals hauria estat l'estudi integrat de les relacions sintàctiques entre els elements de les cadenes fòniques. De totes maneres, des de la perspectiva estructuralista era pràcticament impossible arribar a la generalitat necessària per a aconseguir-ho, ja que es tractaven els diversos sistemes morfemàtics com a estructures aïllades les unes de les altres.

Fins i tot, la mateixa noció d'arbitrarietat del signe lingüístic sembla confirmar aquest pes important que tenen les paraules en l'anàlisi lingüística: si comparem l'estructura fonològica d'una paraula amb el seu contingut, la seva arbitrarietat salta a la vista (en la majoria dels casos, no hi ha res que relacioni l'una amb l'altre); en canvi, si el que comparem són cadenes fonològiques de més d'una paraula, l'explicació de l'arbitrarietat es fa molt més complexa, ja que cal fer referència a les diverses estructures que hi intervenen i a la manera com s'entrellacen les unes amb les altres. Des d'aquesta perspectiva, doncs, resulta també natural de centrar la descripció lingüística en la paraula.

(1) Entenem que la divisió d'un tot és composicional quan la interpretació del tot és una funció de la interpretació de les parts: és a dir, quan per a interpretar el tot n'hi ha prou amb conèixer la interpretació de les parts i la manera com aquestes es combinen per a formar el tot.

Una altra característica fonamental de l'estructuralisme, que en determina profundament la visió sobre el lèxic, és l'establiment de la relació entre la forma i el contingut del signe a un nivell absolutament superficial: són els diversos elements de la cadena fònica (tan segmentats com calgui) els que són portadors de significat, de manera que tots els elements de contingut de les expressions lingüístiques han de ser relacionats de manera directa amb elements de la cadena fònica (siguin morfemes o lexemes; de fet, l'única variació admesa és la del morfema zero). Des del punt de vista de la descripció gramatical, això comporta problemes pràctics en el tractament d'alguns fenòmens (Tusón, 1980), com els constituents discontinus, la sinonímia sintàctica (tal com apareix, per exemple, en oracions actives transitives i les seves corresponents passives), o l'ambigüitat sintàctica (com la que es mostra en expressions com *llapis de fusta de colors*). Des del punt de vista de l'organització de la gramàtica, però, té conseqüències extraordinàriament importants. En no haver de donar compte de realitats lingüístiques abstractes, la descripció lingüística es redueix a la caracterització dels elements en què es descompon la cadena fònica (elements que resulten sempre del procés de segmentació lineal que hem apuntat abans) i a l'establiment de les possibilitats de combinació dels uns amb els altres. De manera natural, doncs, els morfemes i els lexemes es converteixen en el centre de l'anàlisi estructuralista. Les relacions horitzontals (sintagmàtiques) que contreen entre ells serveixen fonamentalment al lingüista per a descobrir les relacions verticals (paradigmàtiques) amb els altres elements del sistema; d'aquesta manera la classificació (taxonomia) dels elements es converteix en una activitat bàsica. No hi ha dubte que una perspectiva així potencia el paper del lèxic (del diccionari) i de la morfologia, en detriment del de la resta de la gramàtica: la màxima activitat descriptiva lingüística es concentra en la caracterització dels diversos elements lèxics (lexemes o morfemes).

Per una banda, els morfemes són caracteritzats en descriure el (sub)sistema del qual formen part, de manera que, per exemple, els morfemes de plural són descrits en el subsistema del nombre (o en els diversos subsistemes del nombre), i els morfemes passius són descrits en el subsistema de la veu verbal.<sup>2</sup> Aquesta caracterització, és clar, es fonamenta en propietats formals, i és només a través de la seva posició en el subsistema corresponent, que cada morfema aporta un significat determinat: així, els morfemes de plural en català aporten un significat diferent del que tenen en grec clàssic a causa que en català l'oposició és entre dos únics valors (singular i plural), mentre que en grec clàssic ho és entre tres valors (singular, dual i plural), de manera que el valor de nombre de l'expressió catalana *els rius que limiten Barcelona* hauria de ser expressat en grec clàssic en dual, i no en plural.

Per l'altra banda, els lexemes són caracteritzats per les relacions de tipus semàntic que contreen entre ells, entre les quals podem destacar dos grans grups de relacions, que Lyons (1977) anomena *camp semàntic* i *relacions de sentit*. De fet, l'estudi d'aquestes relacions constituirà la tasca central de la semàntica estructuralista. Els

(2) Notem de passada que això ens mostra una limitació de l'aparell estructuralista d'anàlisi. La passivitat d'una oració no és, simplement, una característica del seu verb, sinó una propietat general de tota la construcció, en la qual juguen un paper, a més de la morfologia verbal, els règims preposicionals i/o de cas, les relacions semàntiques entre el verb i els seus complements, etc.

campes semàntics conformen els subsistemes en què s'agrupen els lexemes, de manera que diverses de les relacions semàntiques estudiades, en ser-los aplicades, permeten d'establir-ne l'estructura. Indubtablement, les relacions paradigmàtiques que es posen de manifest d'aquesta manera constitueixen el cos de l'anomenada semàntica lèxica, i gran part d'aquesta informació coincideix amb la continguda en els diccionaris que coneixem. Així, de la mateixa manera que els aspectes pròpiament sintàctics són tinguts menys en compte en la tradició estructuralista que els morfosintàctics, els aspectes de semàntica general (o oracional) gairebé no hi són presents.

Totes aquestes opcions tendeixen a afavorir el tractament de la paraula (del lexema) com a nucli de la descripció lingüística, en detriment dels aspectes més generals (tant sintàctics i semàntics, com morfològics). En certa manera, aquesta visió no posa en qüestió la pràctica lexicogràfica tradicional. Això és així perquè, en el fons, manté la posició central de la paraula i no elabora d'una manera prou sistemàtica els aspectes generals de la llengua. Podríem dir, doncs, que l'estructuralisme no proposa encara una visió del llenguatge en la qual tots els ingredients juguin el seu paper conjuntadament, de manera que no ofereix un model (en el sentit modern del terme) del llenguatge humà.

Per a nosaltres, una derivació interessant de la visió estructuralista la constitueixen les primeres versions de la gramàtica de dependència (Tesnière, 1959), ja que se n'ha derivat una de les pràctiques lexicogràfiques més vinculades a una visió teòrica del llenguatge, la dels diccionaris de valències. Tesnière titula el seu llibre *Éléments de syntaxe structurale* i el comença afirmant que l'objecte de la sintaxi estructural és la *phrase*, el sintagma, així com que la noció de sintagma és lògicament anterior a la de mot (Tesnière, 1959:11,25). A la vegada, descriu les relacions entre els diversos elements de les oracions a partir de la noció de dependència, que no ha ser identificada amb la relació de precedència en la cadena fònica; així, com assenyala Baez San José (1988), hi ha una certa similitud entre aquest plantejament i el de la primera gramàtica generativa, en el sentit que les dues pressuposen una realitat abstracta per sobre de la cadena fònica, que en mostra la veritable estructura.

Tot i aquest reconeixement d'una estructura abstracta, la sintaxi de Tesnière està estrictament centrada en la paraula; una simple fullejada del seu llibre permet d'adonar-se que els fets sintàctics són explicats al voltant de les paraules, de les categories majors, sense postular processos sintàctics generals i independents dels lexemes als quals afecten. Així ens trobem davant d'una sistematització de les relacions que els elements superficials de la cadena fònica contreen entre ells, que revela una taxonomia més adequada que les motivades per criteris estrictament superficials, però que, en definitiva, encara està lluny de poder oferir explicacions generalitzadores de les diferents realitzacions lexemàtiques. En els diccionaris de valències inspirats en la teoria de dependències de Tesnière, hi podem trobar nombrosos exemples d'aquest fet.

Considerem les diverses entrades lèxiques corresponents a l'adjectiu alemany *alt*

('vell'), consignades en el diccionari de Sommerfeldt i Schreiber (1974). Hi trobem nou entrades diferents, la majoria de les quals tenen la mateixa realització sintàctica i, per tant, es diferencien només per la caracterització semàntica dels seus arguments. A més, hi apareix com un fenomen casual un fet important de l'alemany, com és que (molts d)els seus adjectius atributius poden tenir una realització predicativa si no són usats en sentit valoratiu. A la vegada, la relació existent entre l'adjectiu *alt* usat absolutament, com a (1), i amb especificadors, com a (2), queda totalment amagada, quan es tracta d'un fenomen habitual en les llengües germàniques pel qual s'admeten especificadors quantitius per als adjectius que poden expressar una mesura en una escala graduada (altres adjectius que admeten aquest tipus d'especificador són, per exemple, *hoch* ('alt'), *lang* ('llarg') o *schwer* ('pesat')).

- (1) a. Die Frau ist alt ('La dona és vella')  
 b. die alte Frau ('la dona vella')
- (2) a. Die Frau ist fünfzig Jahre alt (lit.: 'La dona és vella cinquanta anys')  
 b. die fünfzig Jahre alte Frau (lit.: 'la dona vella de cinquanta anys')

L'existència d'entrades diferents i paral·leles per a una mateixa cadena fònica no suposa que tinguin cap relació determinada. Des d'un punt de vista teòric, es podria tractar de simples cadenes fòniques homònimes, de manera que el seu tractament en el diccionari hauria estat el mateix: aparèixer com a entrades distintes sota la mateixa etiqueta de cadena fònica.

Aquest exemple ens mostra el problema principal de la visió estructuralista quant al disseny d'un model gramatical: els fets descoberts tenen un límit en la possibilitat de ser sistematitzats, límit que és imposat per la impossibilitat d'expressar relacions abstractes entre diverses realitzacions oracionals o sintagmàtiques. Expressions complexes que els parlants d'una llengua tenim com a pròximes o relacionades (sigui com a sinònimes, sigui mitjançant altres tipus de relació –implicació, conseqüència, de contrari, de contradicció, etc.) no poden ser relacionades a causa de la manca d'abstracció de la cadena fònica. És només en les propostes generatives que aquesta limitació podrà ser superada. Passem, doncs, a observar el paper que s'ha atorgat al lèxic en les diverses teories generatives que s'han anat produint des de les seves primeres manifestacions per part de Chomsky els anys cinquanta.

#### 4. EL GENERATIVISME

Els plantejaments generativistes són els primers que han permès el desenvolupament d'una visió del lèxic integrada dins la teoria lingüística general. Això ha estat així a causa dels plantejaments inicials del generativisme: no es tracta de descriure



els fenòmens lingüístics com fets donats, exteriors als parlants, sinó com a produccions lingüístiques d'aquests. En lloc de partir de les dades lingüístiques tal com es poden trobar recollides en els corpus lingüístics (en una visió fonamentalment estàtica del llenguatge), els generativistes parteixen de les expressions lingüístiques tal com són produïdes pels parlants. D'aquesta manera, els objectius bàsics de la lingüística canvien: en lloc de pretendre posar de manifest les relacions estructurals entre els diversos elements presents en els textos (o manifestacions lingüístiques en general), la lingüística pretén explicar la capacitat humana de fer aquestes manifestacions; es tracta, doncs, de produir un model teòric de com el parlant d'una llengua és capaç de produir i de comprendre textos. Conseqüentment, la posició relativa de la lingüística respecte d'altres ciències també canvia: si Saussure la veia com una branca d'una ciència dels signes, d'abast més general, que ell anomenava *semiologia*, Chomsky la veu com una branca de la ciència cognitiva.

Des d'aquesta nova perspectiva, es poden explicar també les relacions estructurals entre els elements lingüístics, però això es fa en un marc global més ampli. Certament, la pretensió dels generativistes és que l'explicació dels fenòmens lingüístics que en resulta és molt més adequada; segons les nocions establertes per Chomsky el generativisme no pretén aportar ja una mera descripció (elaborada fonamentalment a partir d'una taxonomia), sinó una veritable explicació dels fenòmens lingüístics en oferir una visió general del fet lingüístic humà. Ens trobem, doncs, amb el primer intent seriós d'oferir una visió global dels diversos factors que intervenen en l'activitat lingüística humana; conseqüentment, és en aquest marc on comencen a aparèixer visions diferents sobre com s'ha de relacionar el lèxic amb els altres components de la gramàtica.

Tot i que vist des de l'exterior apareix com un tot monolític, el corrent generativista, des de ben al començament de la seva història, ha estat molt ric en tendències i molt ràpid en la seva evolució. En general, la història de la gramàtica generativa ha estat molt lligada a la figura del seu màxim exponent, N. Chomsky, sigui perquè es tracta de models desenvolupats per ell (o per seguidors immediats seus), sigui perquè es proposen models alternatius als seus. Aquesta visió de la lingüística neix a partir dels primers treballs de Chomsky, en els quals incorpora els principis i la metodologia que havien contribuït amb èxit al desenvolupament dels llenguatges artificials durant els últims anys del segle XIX i la primera part d'aquest segle (treballs de Frege, Russell i Whitehead, Tarski i Carnap, entre molts altres). Com a conseqüència, es concep la capacitat lingüística com una teoria formal, amb els seus diversos components, que és capaç de generar (és a dir, de produir) la totalitat de les oracions gramaticals de la llengua en qüestió. És clar que en una aproximació així a la teoria lingüística, el lloc que hi ocuparà el lèxic resultarà fàcilment determinat, ja que tots els components hi tenen especificat el seu paper. Així doncs, des de la nostra perspectiva repassarem alguns d'aquests models i hi observarem el paper que correspon al lèxic. D'aquesta manera, podrem determinar les tasques que les diverses teories lingüístiques assignen al lèxic i, per tant, el paper que en cada cas hi pot tenir la lexicografia.

## 5. EL PRIMER MODEL GENERATIVISTA: EL PRESTÀNDARD

La primera versió àmpliament coneguda de la gramàtica generativa és la que va presentar Chomsky el 1957 al seu llibre *Syntactic Structures*. En ella, el lèxic no hi té un paper diferent del de la sintaxi, és a dir, no constitueix un component diferenciat de la gramàtica. La gramàtica (la sintaxi), doncs, es compon només de dos mòduls diferenciats: les regles d'estructura de frase i les regles transformacionals.<sup>3</sup> Així, tant les relacions de constituència com la introducció dels elements lèxics es produeixen per regles del mateix tipus, sintagmàtiques o de reescriptura de frase. Ho veiem en el conjunt de regles que li serveix d'exemple de l'anàlisi de constituents (Chomsky, 1957:26):<sup>4</sup>

- |     |       |                              |      |  |
|-----|-------|------------------------------|------|--|
| (3) | (i)   | Oració $\rightarrow$ SN + SV | (iv) | T $\rightarrow$ <i>el, la, etc.</i>            |
|     | (ii)  | SN $\rightarrow$ T + N       | (v)  | N $\rightarrow$ <i>noi, nena, pilota, etc.</i> |
|     | (iii) | SV $\rightarrow$ Verb + SN   | (vi) | Verb $\rightarrow$ <i>agafa, xuta, etc.</i>    |

En aquesta llista de regles no hi ha distinció (ni teòrica, ni pràctica) entre les de (3,i-iii) i les restants; ens trobem sempre amb regles que reescriuen un símbol per un altre (o uns altres): el símbol *SN* és reescrit per *T* i *N*, i *N* és reescrit per un dels símbols de la llista *noi, nena, etc.* És clar que, fins i tot en una teoria fortament sintàctica com aquesta, els elements lèxics hi han de tenir un lloc, ja que altrament ens trobaríem amb estructures vàlides (per a la llengua en qüestió), que podrien ser representades per diagrames vàlids (és a dir, generats per les regles corresponents), però que no podrien servir per a la caracterització de les seves oracions gramaticals; dit simplement, generariem estructures, però no oracions; per tant, un llenguatge no podria ser definit per la llista de les seves oracions (és a dir, de les generades pel conjunt de regles en consideració), sinó per la llista de les seves estructures. Així doncs, la presència de regles com les de (3,iii-vi) és totalment justificada. El problema rau en què molt aviat apareixen problemes que difícilment poden ser resolts en aquest marc formal tan estret. El tractament mateix de les variacions morfològiques de les paraules hi pot ser resultat d'una manera relativament simple només perquè es tracta de donar compte de l'anglès, que és una llengua amb molt poques variacions flexives.<sup>5</sup> A la vegada, de seguida apareixen dificultats justament per la mala caracterització dels elements lèxics en aspectes sintàctics (i, fins i tot, semàntics): els verbs, per exemple, han de ser classificats i tractats de manera diferent segons els tipus sintàctics dels complements que poden tenir (transitius, intransitius, preposicionals, etc.), i també segons els trets lèxics dels seus complements (així, *la pedra beu* és inacceptable perquè el verb *beure* demana com a subjecte sintagmes nominals amb el tret lèxic *+animal*).

(3) En aquesta obra, es preveu la incorporació d'un mòdul addicional que donés compte de l'estructura fonològica de les oracions (això es fa a partir de la interpretació fonològica de l'estructura sintàctica resultant de les transformacions); a la vegada, no es nega la possibilitat d'incloure-hi un mòdul semàntic, interpretatiu també, tot i que s'argumenta a bastament que l'estructura sintàctica de les oracions és independent de la seva interpretació semàntica.

(4) En aquest cas, com en tots aquells en què ha estat possible, hem adaptat les propostes originals per a l'anglès, de manera que resultin adequades per al català.

(5) Per a adonar-se de la complexitat que un tractament d'aquest tipus comportaria per al català només cal intentar adaptar les regles de l'auxiliar verbal (Chomsky, 1957:39).

## 6. EL MODEL ESTÀNDARD

Fonamentalment per les raons que acabem d'aduir, en la segona versió de la gramàtica generativa posada a punt per Chomsky (1965), el lèxic entra a formar part, com un component més, de l'estructura general de la gramàtica. De fet és aquesta versió la que estableix els fonaments de la concepció generativa del llenguatge, ja que, al costat de la sintaxi, amb els seus dos nivells (d'estructura superficial i profunda), inclou també la fonologia i la semàntica, com a mòduls addicionals, interpretatius de les estructures sintàctiques superficials i profundes, respectivament.

La major part dels canvis que s'introdueixen a aquest model respecte al prestàndard es relacionen amb el lèxic i la semàntica. És natural que ben aviat es vagi considerar necessària la introducció de la semàntica en el model, ja que altrament es fa difícil de mantenir la idea que es tracta d'un model de la competència lingüística dels parlants, atès que hi ha fenòmens que formen part de la capacitat lingüística i que no poden ser explicats si no és des d'una perspectiva semàntica (per exemple, la sinonímia entre diverses expressions lingüístiques, moltes de les ambigüitats pròpies de les expressions, moltes de les restriccions de selecció, etc.). La semàntica, doncs, s'introdueix en la teoria lingüística perquè sense ella queden sense explicar aspectes importants del comportament lingüístic dels parlants; de totes maneres, es fa tot intentant que no envaeixi el camp de la gramàtica formalitzada, que es reserva a la sintaxi. Es pren aquesta actitud perquè, en el fons, hi ha el convenciment que el llenguatge natural en la seva totalitat no pot ésser formalitzat (almenys en l'estadi actual de desenvolupament de la lingüística). Com a conseqüència, Chomsky intenta introduir en la sintaxi tants elements com pot de la semàntica, i deixa per a aquesta tot allò que no pot ser rigorosament formalitzat.

En la visió de la teoria estàndard, el lèxic apareix com un component més del model lingüístic general. Ara bé, en el lèxic és on s'introdueixen la majoria de les característiques de tipus semàntic que s'incorporen a la gramàtica; així, la informació que fa possible l'adopció del mecanisme de les restriccions de selecció prové del lèxic. I aquesta informació, juntament amb la resta de la informació lèxica, és considerada idiosincràtica (Chomsky, 1965:87,136,142): el lèxic és el lloc on es consignen els aspectes irregulars (és a dir, que no poden ser sotmesos a les regles) del llenguatge, entre els quals hi ha els trets semàntics que caracteritzen les entrades lèxiques. Vegem el contingut d'unes entrades lèxiques, segons queden especificades per Chomsky (1965:107):

- (4) a. (pilota, [+N, +Det -, +Comptable, -Abstracte...])
- b. (noi, [+N, +Det -, +Comptable, +Animat, +Humà...])
- c. (xuta, [+V, + \_ SN, +[+Animat] Auxiliar \_ Det [-Animat], +Eliminació-objecte...])

En elles, s'hi recullen els aspectes que Chomsky mateix indica (1965:87) que formen part del contingut idiosincràtic de les entrades lèxiques: els aspectes de l'estructura fonètica que no són predictibles per regla general (en les entrades anteriors, aquesta informació està representada per la forma ortogràfica), les propietats rellevants per al funcionament de les regles transformacionals (a l'entrada de *xuta*, el tret *+Eliminació-objecte* indica que pot ser objecte de la transformació d'eliminació de l'objecte), les propietats que són rellevants per a la caracterització semàntica de la paraula (a les entrades anteriors, els trets pròpiament semàntics no hi són indicats: haurien de ser especificats en el lloc dels punts suspensius),<sup>6</sup> i els trets que indiquen en quines posicions pot aparèixer el mot en qüestió (la part sensitiva al context de les entrades lèxiques: a les dels noms de (4) s'indica que han d'aparèixer després d'un determinant (*+Det \_*), mentre que a la del verb *xuta* hi ha, per una banda, informació sobre la seva subcategorització (*+ \_SN*) i, per l'altra, sobre els contextos en què pot aparèixer, segons les seves restriccions de selecció (*[+Animat] Auxiliar \_ Det [-Animat]*)).

En el lèxic, doncs, es consigna la informació que no pot ser tractada de manera general. Així, per exemple, si hi hagués una regularitat en el fet que els verbs transitius, o sia, els que subcategoritzen per un objecte directe, poguessin aparèixer sense aquest complement (com *menjar*, per exemple), no caldria indicar aquesta informació en l'entrada lèxica de cada verb transitiu, ja que una regla general (en aquest cas transformacional) contindria la informació corresponent. Ara bé, Chomsky mateix es preocupa per possibles simplificacions del lèxic. A la secció *L'estructura del lèxic* (1965:164ss.), discuteix algunes possibilitats de simplificar les entrades lèxiques, a base de compactar la seva informació. Proposa dos tipus de mecanismes: unes convencions d'interpretació de les gramàtiques i unes regles de redundància sintàctica: les primeres són universals i les altres, específiques de cada llengua.

Les dues convencions proposades per Chomsky són les següents. En primer lloc, un mecanisme per a no haver d'especificar tots els trets possibles en relació a la subcategorització i a les restriccions de selecció: els primers apareixen explícitament a la regla només si són positius, i els segons, només si són negatius; els altres trets (amb valors negatius o positius, respectivament) són omplerts automàticament si no hi ha conflicte amb els trets formulats explícitament. I en segon lloc, un mecanisme per a evitar la duplicació d'informació en cadenes de trets com *[+Animat, +Humà]* o *[+N, +Comú]*, en què la presència d'un tret a la cadena implica necessàriament la presència, amb els valors indicats, dels trets que el precedeixen, així, la presència, en una entrada lèxica del tret *-Humà* indica que el tret *+Animat* és pertinent; i la indicació que un element lèxic és *+Comú* implica que es tracta d'un nom (és a dir, d'un element lèxic marcat *+N*).

Per altra banda, les regles de redundància sintàctica, específiques de cada llengua, serveixen per a simplificar les entrades lèxiques en casos com el de les implicacions entre regles de subcategorització estricta. Així, en català (com en anglès) hi ha verbs

(6) Probablement molts estariem d'acord amb què trets com *humà*, *abstracte* o fins i tot, *comptable* són ja trets semàntics; Chomsky, però, considera que els trets que són usats en alguna regla sintàctica són realment trets sintàctics i no semàntics (1965: 142).

transitius que poden tenir un complement adverbial de manera (per exemple, *netejar*, *llegir...*, com a (5a)), mentre que n'hi ha d'altres que no l'admeten (com *témer*, *va-ler...*, com es mostra a (5b)).

- (5) a. Va llegir el llibre (amb molta atenció / despreocupadament)  
b. Tem les feres (\*amb molta atenció / \*despreocupadament)

Hi ha, doncs, una implicació en el sentit que si un verb transitiu pot portar un complement adverbial de manera, a més del sintagma nominal objecte, també pot aparèixer només amb l'objecte (o sia, sense el complement adverbial); en canvi, la implicació en sentit contrari no és correcta (per al català o l'anglès), ja que no és cert que si un verb subcategoritza per un complement d'objecte també pot portar un complement adverbial de manera (com *témer*). Per tant, podem establir una regla de redundància sintàctica com la de (6).

- (6) [+ \_ SN ^ Compl\_Manera] → [+ \_ SN]

Així, l'entrada lèxica del verb *netejar* es pot formular com s'indica, parcialment, a (7).

- (7) (*netejar*, [+ \_ SN ^ Compl\_Manera, ...])

Una entrada lèxica així, per aplicació de la regla de redundància (6) serà convertida en la de (8a) (també especificada parcialment, només) o fins i tot la de (8b), si ens atenim a la convenció anterior sobre valors negatius en les regles de subcategorització estricta.

- (8) a. (*netejar*, [+ ^ SN\_Cmpl\_Manera, + \_ SN, ...])  
b. (*netejar*, [+ ^ SN\_Cmpl\_Manera, + \_ SN, - \_ Compl\_Manera, - \_, ...])

Així doncs, Chomsky preveu ja la possibilitat d'introduir alguna sistematicitat en la informació continguda en el lèxic, que ha de ser complementada per les idees desenvolupades en els dos apartats dedicats, respectivament, a la morfologia flexiva i a la derivacional. De totes maneres, a partir dels plantejaments d'aquest capítol, resulta bastant clar que l'objectiu d'aquesta sistematicitat és més de simplificar la manera de codificar la informació de les entrades lèxiques, que d'establir generalitzacions gramaticals de valor sintàctic; conseqüentment no hi ha cap contradicció entre les afirmacions generals sobre el caràcter idiosincràtic de la informació continguda en el lèxic i aquestes propostes sobre la possibilitat de simplificar el lèxic. Així i tot, és interessant de remarcar que des dels primers moments de l'establiment formal de la teoria generativa es pren en consideració el paper que hi pot jugar el lèxic i que hi apareixen en



germen els mecanismes de l'herència que hauran de tenir un paper tant rellevant en els plantejaments lexicològics més recents (veure, més avall, §9).<sup>7</sup> Ens trobem, per tant, amb una teoria lingüística que fa un plantejament teòric general de com explicar el fenomen de la llengua; per a això, proposa una sèrie de components de la gramàtica, entre els quals hi ha el lèxic. És, doncs, una teoria que aporta una caracterització clarament definida del lèxic, de manera que les tasques lexicogràfiques en aquest marc poden quedar força delimitades en relació a la informació morfosintàctica (és clar que pel que fa a la semàntica la teoria estàndard no aporta cap idea interessant).

Quasi immediatament després de la publicació de Chomsky (1965) sorgeixen els primers intents de desenvolupar una teoria seriosa del lèxic. Aquests esforços adopten dues direccions diferents: per una banda, hi ha el desenvolupament d'una teoria de les relacions entre paraules (o lexemes), de manera que s'afronten els aspectes relacionats amb la formació de paraules (derivació i composició); per l'altra, apareixen intents de millorar el tractament de la informació semàntica dels elements lèxics i, consegüentment, d'estructurar més clarament la totalitat de la informació que ha d'aportar el lèxic. Un bon exemple del primer moviment el constitueix Botha (1968), l'objectiu fonamental del qual és precisament justificar la necessitat de tractar de manera sistemàtica la informació lèxica (superant, doncs, la visió que el lèxic conté només informació de tipus idiosincràtic); en aquest estudi, aquesta justificació es fonamenta sobretot en la necessitat que té la teoria lingüística d'explicar les relacions existents entre diverses paraules (a través dels processos de derivació i, sobretot, de composició). Un altre exemple, aquest des de la lexicografia, el constitueix Corbin (1981), on s'analitzen les pràctiques lexicogràfiques d'alguns diccionaris francesos, des del marc teòric aportat per la teoria estàndard.

El segon tipus de reacció pot ser exemplificat amb la posició defensada per Fillmore (1971), on estudia precisament els diversos tipus d'informació que ha de contenir el lèxic, entre els quals inclou el rol que cada argument juga en la situació indicada per les paraules predicatives, les pressuposicions (o condicions) que fan que la paraula sigui usada adequadament, el tipus de relació conceptual o morfològica que la paraula té amb altres elements del lèxic, i el seu significat. Tot i que es tracta fonamentalment d'una proposta programàtica, ens permet de fer-nos una idea de com semblava en aquell moment que els aspectes semàntics i pragmàtics podrien ser incorporats en les entrades lèxiques. La proposta sembla un avenç important en la direcció d'incorporar nous aspectes a la part tractada (és a dir, formalitzada) del llenguatge. Ara bé, en moltes ocasions (i això va ocórrer en els moments històrics posteriors a la publicació de Chomsky (1965)), la introducció de la semàntica (i, és clar, de la pragmàtica) ha comportat una pèrdua de rigorositat en el plantejament explicatiu i formal; serà només fins a la generalització de les tècniques semàntiques desenvolupades a la lògica i a les ciències de la computació que aquest risc podrà ser superat. De totes maneres, des del punt de vista del lexicògraf, un plantejament així apunta en una direcció molt atractiva, ja que

(7) Chomsky se'n revela, un cop més, com un geni, que és capaç d'intuir, molts anys abans de disposar de l'aparell formal adequat, la rellevança que tindran els mecanismes d'organització del lèxic.

sense oblidar els aspectes sintàctics fonamentals (tal com han estat recollits en la versió estàndard de la teoria generativa) s'indica un camí per a la introducció de la informació semàntica, que tot diccionari ha de contenir, lligada d'alguna manera amb la informació pròpiament sintàctica.

La tumultuosa història de la gramàtica generativa els últims anys 60 (Newmeyer, 1986) va centrar-se en la relació entre la sintaxi i la semàntica (específicament, en si les estructures semàntiques eren també generades –és a dir, si hi havia regles que les produïen– o eren simplement el resultat de la interpretació de les estructures sintàctiques),<sup>8</sup> i també en la de la sintaxi i la morfologia. En part les dues qüestions anaven lligades, en tant que la descomposició de les paraules (tasca fonamentalment morfològica) podia anar directament a la interpretació semàntica o podia donar lloc a regles del component de base. Més que la polèmica en si, el que ens interessa aquí és de fer constar el seu resultat, en la que fou l'opció que ha perdurat posteriorment. El corrent central de la gramàtica generativa, és a dir el liderat per Chomsky, va acabar adoptant el principi exposat a Chomsky (1970) que es coneix amb el nom d'*hipòtesi lexicalista* i que ha donat peu a la majoria dels estudis en morfologia dins d'aquest corrent central de la teoria generativa.

La hipòtesi lexicalista, com suggereix el seu nom, concedeix al lèxic una importància més gran, ja que tracta com a lèxiques unes relacions que s'havien considerat gramaticals. Com menciona Newmeyer (1986:107), la funció essencial de la hipòtesi va ser prohibir les transformacions que provoquessin un canvi de categoria; d'aquesta manera, s'evitava que, per exemple, un verb o un adjectiu fos transformat en un nom i es restringia fortament l'abast del camp d'aplicació de les regles transformacionals. Així, la relació entre *Jaume ha arribat* i *l'arribada de Jaume* ha de ser formulada dins el lèxic i no per mitjà de regles transformacionals en la gramàtica. Naturalment, com Chomsky (1970:22) indica, una posició com aquesta només pot ser presa quan el lèxic ha estat separat del component de base i quan les entrades lèxiques contenen informació contextual (a través de les subcategoritzacions i de les restriccions de selecció); hi pot haver entrades lèxiques de categories diferents que mantinguin les mateixes condicions contextuais.

Com és natural, una posició així va obrir les portes a tota una sèrie d'estudis morfològics en el marc de la teoria generativa. Si fins ara, en general els generativistes no distingien entre els fenòmens morfològics i els sintàctics, l'acceptació d'aquesta hipòtesi comporta una distinció molt clara entre les relacions lèxiques i les gramaticals, i entre els seus constructes: els de les primeres no poden ser afectats per les regles de la gramàtica. D'aquesta manera s'engega la primera gran onada de treballs morfològics de la gramàtica generativa: es tracta de precisar la naturalesa, els límits i les condicions d'aplicació de les regles lèxiques, centrats sempre en els mecanismes de la derivació i la composició.

(8) D'aquí els noms de *semàntica generativa* i *semàntica interpretativa*; el primer designa una teoria que va tenir els seus moments de màxim esplendor aquests anys, i el segon és el lema dels generativistes que es van mantenir en l'ortodòxia i van continuar defensant l'autonomia de la sintaxi.

## 7. LA TEORIA DE LA RECCIÓ I EL LLIGAM

El plantejament teòric actual del corrent generativista majoritari és el conegut sota el nom de *Teoria de la recció i el lligam*, normalment abreujat sota les sigles angleses GB (de *Government and Binding*). En aquesta nova versió, han desaparegut les regles de reescriptura com a element central en la construcció de les representacions sintàctiques; el seu lloc és ocupat per una sèrie de principis generals, que són els que determinen la gramaticalitat de les oracions. L'estructura general de la teoria és encara molt similar a la de la teoria estàndard, amb un doble nivell sintàctic (superficial i profund), i dos mòduls interpretatius (la forma fonètica i la forma lògica).

Ara bé, la manera com es relacionen els constructes de cada estructura és força diferent de la que trobàvem a la teoria estàndard; allà podíem especificar d'una manera bastant clara el tipus de regla que corresponia a cada relació: així, per exemple, l'estructura profunda i l'estructura superficial estaven relacionades a partir de les regles transformacionals, que no intervenien en cap altra relació. Aquí, en canvi, són els diversos principis els que determinen les relacions que s'estableixen entre les diferents estructures (teoria X', teoria  $\theta$ , moviment d' $\infty$ , principi de projecció, criteri  $\theta$ , etc.); cosa que es veu complicada, sovint, pel fet que actuen en més d'un nivell.

Convé que vegem quin és el paper del lèxic en aquesta estructura complexa, tenint en compte que presentacions generals de la teoria com la de Sells (1985) no l'hi situen.<sup>9</sup> El lèxic a GB, com a totes les teories generatives antecessores seves, és el lloc on s'emmagatzema la informació pertinent relacionada amb cada entrada lèxica. Pel que fa a la sintaxi (ja que se segueix deixant de banda la semàntica), aquesta informació es concreta en quatre apartats: una matriu fonològica, un tret categorial, un esquema sintàctic, i una graella de papers temàtics (rols  $\theta$ ).

Així, l'entrada lèxica del verb *donar* seria aproximadament com segueix (Zubizarreta, 1986):

(9)	donar: +V.	(AGENT,	destí.	tema)
			SP	SN

En el llistat dels papers temàtics dels predicats s'estableix una distinció entre els papers interns i els externs (Williams, 1981), que queda reflectida a (9), on l'argument extern és representat en majúscules i els interns, en minúscules. Aquesta distinció, doncs, és especificada en el lèxic i una classificació dels verbs segons els seus arguments i el corresponent esquema sintàctic ho ha de tenir en compte. Hi ha verbs que no tenen arguments interns (com *viure*), mentre que n'hi ha d'altres que en tenen un (com *veure*) o més (com *donar*), dels quals alguns tenen argument extern (com *veure* i *donar*) i d'altres no (com *arribar*). La connexió entre els arguments interns i la seva caracterització sintàctica es fa en el lèxic, mentre que la manera com els arguments

(9) Vegeu, però, la presentació de de Geest i Jaspers (1991), que seguim en la nostra descripció.



externs es realitzen a la superfície (si és que s'hi realitzen) depèn de factors de tipus sintàctic que operen a nivell posterior.

Seguint l'establert a la teoria de la X barrada (Jackendoff, 1977), a GB l'enllaç entre les entrades lèxiques i les representacions sintàctiques que se'n deriven, ha de seguir el principi de projecció, segons el qual les representacions de tots els nivells sintàctics es projecten a partir del lèxic d'acord amb les propietats de subcategorització dels elements lèxics (Chomsky, 1981:29). D'aquesta manera, les propietats dels constructes sintàctics són determinades per les propietats dels elements lèxics dels quals són una projecció. A la vegada, es manté la hipòtesi lexicalista, de manera que les transformacions no poden alterar les estructures interiors dels elements lèxics. És important de notar que aquests dos principis atorguen al lèxic un lloc preeminent en la teoria, que abans no havia tingut; el mecanisme de les regles d'inserció lèxica que apareixia a la teoria estàndard, no garantia aquesta immutabilitat de la informació lèxica al llarg dels processos posteriors, sobretot transformatius. Així, amb GB, el lèxic adquireix un paper preponderant en la teoria generativa com mai no havia tingut des del seu naixement. Sens dubte aquests plantejaments suposen un capgirament important en el generativisme, que havia representat una preponderància absoluta de la sintaxi enfront de les altres subdisciplines lingüístiques (com la morfologia o la semàntica, més tradicionalment lligades a la lexicografia).

GB, doncs, atorga una posició determinada al lèxic, amb unes funcions molt concretes en relació a les operacions de tipus sintàctic que es duen a terme en el procés de generació de les oracions. Un plantejament així permet de posar sobre la taula diferents hipòtesis i contrastar les unes amb les altres. No és estrany, doncs, que un dels debats més rics en els darrers anys hagi estat sobre la relació de la sintaxi amb la morfologia; d'una manera especial, l'explicació dels fenòmens de la derivació i de la composició (tradicionalment inclosos en la morfologia) ha centrat bona part d'aquesta polèmica. En general, el centre de la discussió ha estat el veure fins a quin punt hi ha uns processos morfològics independents (i diferents) dels processos sintàctics; la discussió, és clar, pren tot el seu interès quan es comencen a considerar factors universals (o, almenys, multilingüals), ja que hi ha fenòmens que en unes llengües prenen una manifestació sintàctica i en d'altres en prenen una de morfològica. Un dels centres d'interès a GB actualment és la relació entre les estructures predicatives dels elements lèxics de base i les dels seus derivats (per exemple, entre la dels verbs i les dels noms que se'n deriven): es tracta de veure com es projecta l'estructura argumental del predicat de base en els seus derivats i quines modificacions ha de tenir el seu esquema sintàctic, de manera que no es perdin les propietats que convé traspassar de l'un als altres i que es puguin mantenir els principis essencials de la teoria. En gran part, els plantejaments en aquest tema estan lligats al problema general del lligam entre la graella temàtica i l'esquema de subcategorització sintàctica dels predicats (com a Grimshaw, 1990).

Per tant, segons els principis de GB, el lèxic constitueix un element central de la teoria lingüística, de manera que les peces lèxiques usades en la generació d'oracions

resulten fonamentals en la forma final d'aquestes. A la vegada, hi ha un seguit de regles de caire morfològic que contribueixen a enllaçar uns elements lèxics amb altres. Aquestes regles fan possible la reducció del nombre d'entrades lèxiques finals, ja que permeten d'explicar l'aparició d'elements lèxics nous (creativitat en la producció lèxica). A la vegada representen un primer pas en l'explicació de les relacions semàntiques entre entrades lèxiques, ja que en principi s'accepta que la productivitat lèxica va lligada a la composicionalitat en la interpretació; així, les paraules aparentment derivades, que no poden ser interpretades composicionalment (a partir de la interpretació dels seus morfemes i lexemes), són considerades morfològicament irregulars i, conseqüentment, s'ha proposat que siguin llistades en el diccionari (di Sciullo i Williams, 1987).

## 8. LA GRAMÀTICA LÈXICO-FUNCIONAL

En els moments més àlgids de la gramàtica transformacional, en què les transformacions eren utilitzades de manera molt lliure, el seu excessiu poder expressiu va conduir molts lingüistes a buscar-hi alternatives. Es tractava d'intentar construir una gramàtica expressivament adequada, és a dir, que no generés més oracions de les que són considerades naturals en la llengua en qüestió i que no generés la mateixa oració per dos camins diferents. A la vegada, es considerava que el tractament de la informació lingüística en termes d'estructura de constituents, que en principi apareixia com a molt adequat per a l'anglès, no resultava natural per a moltes altres llengües en impedir-hi generalitzacions interessants. Probablement la primera teoria que va aparèixer com una alternativa seriosa fou la Gramàtica relacional de D. Perlmutter. De totes maneres, la primera proposta que ha tingut una gran influència en el desenvolupament posterior de la lingüística teòrica ha estat la Gramàtica lèxico-funcional (LFG, de *Lexical-Functional Grammar*), que ha iniciat el darrer moviment lexicalista en la teoria gramatical. L'obra clàssica de presentació de LFG és Bresnan (1982); de totes maneres, n'existeixen diverses descripcions breus, com Sells (1985) i Wescoat i Zaenen (1991).

Com el seu mateix nom indica, ja des dels seus inicis, LFG ha concedit una gran importància al lèxic.<sup>10</sup> Dit breument, el lèxic és el lloc on s'expressen totes les relacions locals (Wescoat i Zaenen, 1991:119): en el lèxic es descriuen tots els processos que alteren l'estructura argumental dels predicats; així, els fenòmens gramaticals que en altres teories generatives s'explicaven mitjançant transformacions, es descriuen amb operacions de tipus lèxic. Un bon exemple d'això és el tractament de la passiva. Tant a la teoria estàndard, com a GB, la passiva és explicada en termes d'una transformació que implica (entre altres coses) el moviment del sintagma nominal objecte a la posició de subjecte; es tracta, doncs, d'una relació entre estructures sintàctiques: l'estructura generada per les regles de base és relacionada (per mitjà d'una transformació o d'un moviment) amb una altra estructura sintàctica.<sup>11</sup>

(10) L'altra etiqueta, *funcional*, fa referència a la noció de funció sintàctica, que LFG considera un primitiu de la teoria gramatical; així, nocions com la de subjecte o objecte adquireixen un paper descriptiu i un estatus precís: són generalitzacions sobre la manera com els predicats es relacionen amb els seus arguments, independentment de la seva forma sintàctica (cas, règim preposicional...).

(11) Notem, de passada, que aquesta segona estructura sovint és d'unes característiques tals que hauria pogut ser generada directament per les regles de base; així, per exemple, l'estructura configuracional de la passiva *El president fou rebut amb eufòria* és idèntica a la de *El president ha arribat amb retard*, és a dir, la mostrada per l'esquema següent:  $I_C[S_N I_{S_N} I_V I_{S_N} I_V I_{S_P} III]$ .



- (12) a. Els nois semblen tenir gana  
 b. I expected John to arrive on time (lit.: 'Esperava J. arribar a l'hora')

La segona línia de cada entrada serveix per a indicar que l'element no subcategoritzat pel verb d'elevació (el marcat SUBJ per a *semblar* i OBJ per a *expect*) són el subjecte de la clàusula d'infinitiu que té la funció de XCOMP. Notem, de passada, que aquesta representació dels verbs d'elevació comporta, també, un tractament lèxic d'aquest fenomen: és en la mateixa entrada lèxica que s'estableix la relació entre la realització superficial de l'argument i la seva posició en el component de base. A LFG es preveu l'aplicació de les regles morfològiques abans de la construcció de les estructures sintàctiques; per tant, es tracta d'operacions que es duen a terme a l'interior del lèxic, el resultat de les quals serveix d'entrada a les regles sintàctiques. Així, es poden escriure entrades lèxiques per a afixos, com la següent, adaptada de Sells (1985), per a la tercera persona del singular del present dels verbs de la primera conjugació.

- (13) -a AFIX           (↑ SUBJ PERS) = 3  
                           (↑ SUBJ NOM) = sing  
                           (↑ TEMPS) = pres

Naturalment, les formes irregulars no poden ser tractades d'aquesta manera, sinó que han de ser formulades explícitament en el lèxic. Per tal que no hi hagi interferències entre els processos regulars i els irregulars, cal una bona estructuració de tota aquesta informació. Així, en general, cal pensar en els processos regulars com processos per defecte, és a dir, processos que es realitzen sempre que no hi ha una indicació en el sentit contrari, o sia, sempre que no hi ha una formació irregular (vegeu els mecanismes descrits breument a § 9).

Finalment, LFG preveu l'existència de regles lèxiques, que serveixen per a expressar relacions entre elements lèxics. Aquestes regles s'apliquen també a l'interior del lèxic i, per tant, és només el seu resultat que és utilitzat per les regles sintàctiques. Un exemple característic de regla lèxica a LFG és la de la passiva, la part essencial de la qual trobem a (14):

- (14) (SUBJ) →  $\phi$  / (COMPL\_AG)  
 (OBJ) → (SUBJ)

D'acord amb la regla de (14), l'entrada lèxica d'un verb que subcategoritza per a les funcions de subjecte i objecte pot ser convertida en una entrada lèxica, l'esquema de subcategorització de la qual és com s'indica a (15): sense complement agent a (15a) i amb aquest complement a (15b).

- (15) a. <(↑ SUBJ)>  
 b. <(↑ COMPL\_AG) (↑ SUBJ)>

Els altres fenòmens que comporten una alteració en l'estructura argumental dels predicats són formulats també per mitjà de regles lèxiques; així, des d'aquest punt de vista, fenòmens com la causativització, la reflexivitat, la intransitivització, l'eliminació de l'objecte, o les construccions de doble objecte, serien tractades com a fenòmens lèxics mitjançant regles lèxiques.

Així doncs, LFG ens porta a una visió del lèxic molt més rica que el simple llistat d'entrades lèxiques. En primer lloc, la informació que aporta el lèxic està integrada plenament en el conjunt de la informació lingüística tractada per la gramàtica, fins al punt que molta d'aquesta darrera deriva, de manera directa o indirecta, de la primera. En segon lloc, els elements en el lèxic estan organitzats, de manera que s'estableixen relacions entre ells: per una banda, els processos flexius permeten de derivar unes formes lèxiques d'unes altres i, per l'altra, les regles lèxiques relacionen unes entrades lèxiques amb les altres, permetent així de formular semblances i paral·lelismes entre elles. D'aquesta visió del lèxic a una concepció lexicogràfica totalment estructurada hi va només un pas. La propera (i darrera) teoria que tractarem ens mostrarà ja una organització totalment estructurada del lèxic.

## 9. LA GRAMÀTICA D'ESTRUCTURA SINTAGMÀTICA ORIENTADA AL NUCLI

HPSG (de *Head-driven Phrase Structure Grammar*) és la més recent de les gramàtiques d'unificació (Shieber, 1985). Els seus aspectes fonamentals poden ser estudiats amb un cert detall gràcies a la publicació de dos volums de presentació general (Pollard i Sag, 1987; en premsa). Ja des dels plantejaments inicials, aquesta teoria manté deliberadament concepcions i tractaments elaborats originalment en d'altres marcs teòrics. Com Pollard i Sag (1987:1) assenyalen no es tracta tant de copiar d'ací i d'allà, com d'oferir un punt de vista nou a molts plantejaments que han estat desenvolupats en la gran tradició intel·lectual que hi ha darrera l'evolució de les gramàtiques formals els darrers anys.

El primer tret distintiu de HPSG al qual hem de fer referència és que codifica la informació lingüística en estructures de trets. En això, adopta una posició comuna a tots els formalismes i teories basats en la unificació (Shieber, 1985), que ja es troba, encara que parcialment, a LFG. Les estructures de trets són concebudes com *entitats que descriuen o representen un altre objecte tot especificant valors per a diversos atributs de l'objecte descrit* (Pollard i Sag, 1987:28); en general, anomenem *tret* qual-sevol parell d'atribut i valor. Normalment, les estructures de trets contenen només informació parcial sobre l'objecte descrit. Així, podem tenir una estructura de trets amb l'única informació que el nom (atribut) de l'objecte és *Barcelona* (valor), que descriurà (parcialment) un nombre indeterminat d'objectes, entre els quals diversos

pobles i ciutats; a aquesta estructura de trets li podem afegir algun tret nou, com que la població (atribut) de l'objecte és de més d'un milió d'habitants (valor) i que és del país (atribut) Catalunya (valor), de manera que llavors l'estructura de trets ja no descriu, en el món actual, cap altre objecte que la capital de Catalunya; encara que aquesta darrera descripció només tingui un sol objecte en el món que se li adequi, també es tracta només d'una descripció parcial de Barcelona! En general, aquestes estructures de trets es representen amb matrius d'atributs i valors, com la de (16), que recull la informació (parcial) que acabem de mencionar sobre Barcelona:

(16)	<table style="border-collapse: collapse; margin: 0 auto;"> <tr> <td style="padding-right: 10px;">nom:</td> <td>Barcelona</td> </tr> <tr> <td>població:</td> <td>&gt; 1.000.000 hab</td> </tr> <tr> <td>país:</td> <td>Catalunya</td> </tr> </table>	nom:	Barcelona	població:	> 1.000.000 hab	país:	Catalunya
nom:	Barcelona						
població:	> 1.000.000 hab						
país:	Catalunya						

Quan en una teoria lingüística s'usen aquestes estructures, es fa amb la finalitat de representar les propietats formals que tenen els objectes lingüístics. Per als creadors de HPSG els objectes lingüístics són signes (i citen explícitament Saussure), el significat i significat dels quals reinterpreten, respectivament, com un tipus de situació de parla i com una forma determinada de ser les coses en el món. Convé que notem que en el vessant semàntic del signe adopten una postura absolutament realista, declarant-se així hereus de la tradició semàntica que, provinent de la lògica i la filosofia, va entrar en lingüística a través de l'obra de Montague (1974). De fet, les opcions semàntiques dels creadors de HPSG s'encaminen més cap a la versió anomenada de la semàntica de les situacions (Barwise i Perry, 1983), que cap a versions més clàssiques d'aquesta tradició semàntica.

El segon tret distintiu de HPSG que cal que mencionem és que, a causa de la representació de la informació lingüística amb estructures de trets, aquesta teoria no demana diversos nivells de representació. Tots els aspectes descriptius d'un signe lingüístic determinat poden coexistir en una única estructura. Així, en un únic signe hi pot aparèixer, alhora, informació de tipus fonològic (o gràfic), morfosintàctic, de sintaxi relacional i semàntic. Considerem, per exemple, l'entrada lèxica de la forma verbal catalana *menjà*, adaptada (i simplificada) de les que proposen Pollard i Sag:

(17)	<table style="border-collapse: collapse; margin: 0 auto;"> <tr> <td colspan="2" style="padding: 5px;">fonologia menjà</td> </tr> <tr> <td style="padding: 5px;">sintaxi   local</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;"> <table style="border-collapse: collapse; margin: 0 auto;"> <tr> <td style="padding-right: 10px;">nucli</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;"> <table style="border-collapse: collapse; margin: 0 auto;"> <tr> <td style="padding-right: 10px;">major</td> <td>V</td> </tr> <tr> <td>forma-verb</td> <td>finita</td> </tr> <tr> <td>auxiliar</td> <td>-</td> </tr> </table> </td> </tr> <tr> <td style="padding: 5px;">subcat</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">SN[ACC, NORM]<sub>2</sub>, SN[NOM, NORM, SING, 3PERS]<sub>1</sub></td> </tr> <tr> <td style="padding: 5px;">lex</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">+</td> </tr> </table> </td> </tr> <tr> <td style="padding: 5px;">semàntica   contingut</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;"> <table style="border-collapse: collapse; margin: 0 auto;"> <tr> <td style="padding-right: 10px;">relació</td> <td>menjar</td> </tr> <tr> <td>arg1</td> <td>1</td> </tr> <tr> <td>arg2</td> <td>2</td> </tr> </table> </td> </tr> </table>	fonologia menjà		sintaxi   local	<table style="border-collapse: collapse; margin: 0 auto;"> <tr> <td style="padding-right: 10px;">nucli</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;"> <table style="border-collapse: collapse; margin: 0 auto;"> <tr> <td style="padding-right: 10px;">major</td> <td>V</td> </tr> <tr> <td>forma-verb</td> <td>finita</td> </tr> <tr> <td>auxiliar</td> <td>-</td> </tr> </table> </td> </tr> <tr> <td style="padding: 5px;">subcat</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">SN[ACC, NORM]<sub>2</sub>, SN[NOM, NORM, SING, 3PERS]<sub>1</sub></td> </tr> <tr> <td style="padding: 5px;">lex</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">+</td> </tr> </table>	nucli	<table style="border-collapse: collapse; margin: 0 auto;"> <tr> <td style="padding-right: 10px;">major</td> <td>V</td> </tr> <tr> <td>forma-verb</td> <td>finita</td> </tr> <tr> <td>auxiliar</td> <td>-</td> </tr> </table>	major	V	forma-verb	finita	auxiliar	-	subcat	SN[ACC, NORM] <sub>2</sub> , SN[NOM, NORM, SING, 3PERS] <sub>1</sub>	lex	+	semàntica   contingut	<table style="border-collapse: collapse; margin: 0 auto;"> <tr> <td style="padding-right: 10px;">relació</td> <td>menjar</td> </tr> <tr> <td>arg1</td> <td>1</td> </tr> <tr> <td>arg2</td> <td>2</td> </tr> </table>	relació	menjar	arg1	1	arg2	2
fonologia menjà																									
sintaxi   local	<table style="border-collapse: collapse; margin: 0 auto;"> <tr> <td style="padding-right: 10px;">nucli</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;"> <table style="border-collapse: collapse; margin: 0 auto;"> <tr> <td style="padding-right: 10px;">major</td> <td>V</td> </tr> <tr> <td>forma-verb</td> <td>finita</td> </tr> <tr> <td>auxiliar</td> <td>-</td> </tr> </table> </td> </tr> <tr> <td style="padding: 5px;">subcat</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">SN[ACC, NORM]<sub>2</sub>, SN[NOM, NORM, SING, 3PERS]<sub>1</sub></td> </tr> <tr> <td style="padding: 5px;">lex</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">+</td> </tr> </table>	nucli	<table style="border-collapse: collapse; margin: 0 auto;"> <tr> <td style="padding-right: 10px;">major</td> <td>V</td> </tr> <tr> <td>forma-verb</td> <td>finita</td> </tr> <tr> <td>auxiliar</td> <td>-</td> </tr> </table>	major	V	forma-verb	finita	auxiliar	-	subcat	SN[ACC, NORM] <sub>2</sub> , SN[NOM, NORM, SING, 3PERS] <sub>1</sub>	lex	+												
nucli	<table style="border-collapse: collapse; margin: 0 auto;"> <tr> <td style="padding-right: 10px;">major</td> <td>V</td> </tr> <tr> <td>forma-verb</td> <td>finita</td> </tr> <tr> <td>auxiliar</td> <td>-</td> </tr> </table>	major	V	forma-verb	finita	auxiliar	-																		
major	V																								
forma-verb	finita																								
auxiliar	-																								
subcat	SN[ACC, NORM] <sub>2</sub> , SN[NOM, NORM, SING, 3PERS] <sub>1</sub>																								
lex	+																								
semàntica   contingut	<table style="border-collapse: collapse; margin: 0 auto;"> <tr> <td style="padding-right: 10px;">relació</td> <td>menjar</td> </tr> <tr> <td>arg1</td> <td>1</td> </tr> <tr> <td>arg2</td> <td>2</td> </tr> </table>	relació	menjar	arg1	1	arg2	2																		
relació	menjar																								
arg1	1																								
arg2	2																								

En ella, hi trobem els quatre tipus d'informació que mencionàvem: la fonològica (representada, de manera simplificada, pel tret FONOLOGIA = menjà), la morfosintàctica (expressada per una part dels trets de l'atribut *SINTAXI|LOCAL|NUCLI*), la de sintaxi relacional (continguda fonamentalment en el tret *SINTAXI|LOCAL|SUBCAT*), i la semàntica (expressada com a valor de l'atribut *SEMÀNTICA*).

Un tractament així té l'avantatge que es poden establir d'una manera molt fàcil i elegant les relacions que hi ha entre els complements exigits sintàcticament i els arguments de la relació expressada; a la vegada, un element rector pot imposar restriccions als seus complements tant d'ordre sintàctic com semàntic. Tots dos factors apareixen a l'entrada lèxica de *menjà*: els índexs 1 i 2 indiquen, respectivament, que el contingut del sintagma nominal subjecte és el mateix que l'argument primer del contingut de *menjar*, i que el contingut de l'objecte és idèntic a l'argument segon del seu contingut.

Finalment, el tercer tret distintiu de HPSG que ens interessa de remarcar és el lloc central que hi ocupa la unificació. Com en els altres formalismes d'unificació, per a HPSG la unificació és una operació fonamental a l'hora de relacionar estructures de trets. La informació continguda en elles es combina gràcies a aquesta operació, de manera que en resulta una altra estructura de trets. Considerem les estructures de trets de (18):

- (18) a. 

categoria:	sn
concordança:	[nombre: singular]

  
 b. 

categoria:	sn
concordança:	[persona: tercera]

  
 c. 

categoria:	sn
concordança:	[nombre: plural]

La informació de les estructures de trets de (18a) i (18b) és diferent, però compatible; en canvi, la informació de (18a) i (18c) es incompatible. Per tant, no és possible de pensar en una estructura de trets que contingui tota la informació de (18a) i (18c) alhora, mentre que sí que es pot construir una estructura de trets que contingui la informació tant de (18a) com de (18b): aquesta estructura de trets resultant serà la unificació de les dues originals, que queda consignada a (19):

- (19) a. 

categoria:	sn
concordança:	[nombre: singular persona: tercera]

En general, i informalment, podem definir la unificació de dues estructures de trets A i B com l'estructura de trets C que conté tota la informació de A i B i només la informació de A i B.

L'operació d'unificació pot ser usada per a explicar bona part de les condicions que regulen les relacions sintagmàtiques entre elements lingüístics. Així, en els forma-

lismes d'unificació més simples, com les DCGs (Pereira i Warren, 1980) i PATR (Shieber, 1986), les regles d'estructura sintagmàtica usuals són complementades amb informació addicional respecte a les característiques dels elements lingüístics que hi intervenen, i la informació corresponent a la categoria mare és construïda a partir de la unificació de la informació de les filles.

Equipats amb aquestes dades sobre HPSG, podem passar a considerar-ne la visió sobre el lèxic. HPSG, com LFG (vegeu, abans, §8) i com la gramàtica categorial (Pollard i Sag, 1987:191), adopta la postura de lexicalitzar al màxim la informació lingüística, cosa que li permet aconseguir *una reducció dràstica del nombre i la complexitat de les regles d'estructura sintagmàtica*. Naturalment, aquesta opció comporta un augment notable de la complexitat dels signes lingüístics. Com a mostra d'aquest fet, podem observar a (20) l'entrada lèxica anglesa que proposen per a la forma verbal *likes* (tercera persona del singular de *like*, 'agradar').

(20)	fonologia	likes												
	sintaxilocal	nucli	<table border="1"> <tr> <td>major</td> <td>V</td> </tr> <tr> <td>forma_verb</td> <td>finita</td> </tr> <tr> <td>auxiliar</td> <td>-</td> </tr> </table>	major	V	forma_verb	finita	auxiliar	-					
major		V												
forma_verb	finita													
auxiliar	-													
		subcat	SN[ACUSATIU] 2, SN[NOMINATIU]	<table border="1"> <tr> <td>1</td> <td>persona</td> <td>tercera</td> </tr> <tr> <td></td> <td>nombre</td> <td>singular</td> </tr> </table>	1	persona	tercera		nombre	singular				
1	persona	tercera												
	nombre	singular												
		lex	+											
semàntica		contingut	<table border="1"> <tr> <td>relació</td> <td>like</td> </tr> <tr> <td>arg1</td> <td>1</td> </tr> <tr> <td>arg2</td> <td>2</td> </tr> <tr> <td>localització</td> <td>3</td> </tr> </table>	relació	like	arg1	1	arg2	2	localització	3			
	relació	like												
arg1	1													
arg2	2													
localització	3													
		índexs	<table border="1"> <tr> <td>variable</td> <td>3</td> </tr> <tr> <td>restricció</td> <td> <table border="1"> <tr> <td>relació</td> <td>coincideix-temporalment</td> </tr> <tr> <td>arg1</td> <td>3</td> </tr> <tr> <td>arg2</td> <td>1<sub>j</sub></td> </tr> </table> </td> </tr> </table>	variable	3	restricció	<table border="1"> <tr> <td>relació</td> <td>coincideix-temporalment</td> </tr> <tr> <td>arg1</td> <td>3</td> </tr> <tr> <td>arg2</td> <td>1<sub>j</sub></td> </tr> </table>	relació	coincideix-temporalment	arg1	3	arg2	1 <sub>j</sub>	
variable	3													
restricció	<table border="1"> <tr> <td>relació</td> <td>coincideix-temporalment</td> </tr> <tr> <td>arg1</td> <td>3</td> </tr> <tr> <td>arg2</td> <td>1<sub>j</sub></td> </tr> </table>	relació	coincideix-temporalment	arg1	3	arg2	1 <sub>j</sub>							
relació	coincideix-temporalment													
arg1	3													
arg2	1 <sub>j</sub>													

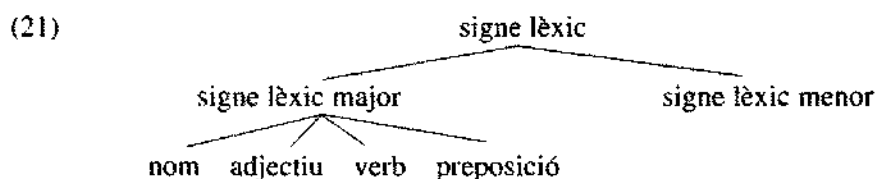
Si ens enfrontem amb entrades lèxiques d'aquesta complexitat (i les entrades lèxiques definitives encara haurien de ser més complexes), resulta evident que calen mecanismes per a estructurar la informació lèxica: ni el model teòric seria explicatiu (hi hauria un nombre molt elevat de relacions que no podrien ser explicades), ni podríem pretendre cap mena d'adequació psicològica per al model (seria impensable que els parlants d'una llengua tinguessin interioritzat un lèxic així, ja que és impossible de memoritzar tota aquesta informació per a cada paraula), ni es podria tractar d'un model computacionalment viable (a causa de la redundància en la informació lèxica, es gastarien quantitats impensables d'espai en la memòria dels programes). La informació lèxica, doncs, ha d'estar convenientment estructurada, de manera que la totalitat de



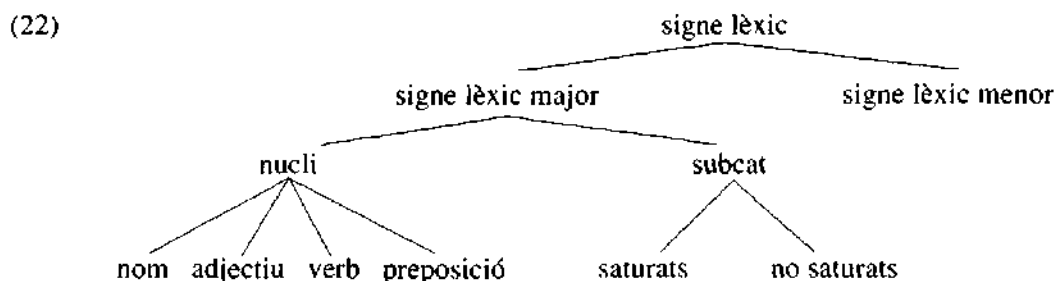
la informació que la teoria pressuposa per a entrades lèxiques com la de (20) pugui ser deduïda per mecanismes precisos a partir d'unitats més petites d'informació. El lèxic que proposen a HPSG està estructurat a partir de dos grans tipus de relacions: les que són determinades per la jerarquia de tipus lèxics i les que queden explicitades en regles lèxiques. A continuació les descriurem breument tal com queden exposades, per exemple, a Pollard i Sag (1987) i Flickinger (1987); vegeu Shieber (1985) per a una descripció paral·lela més elemental.

Les estructures de trets estan tipificades; és a dir, hi ha diversos tipus d'estructures de trets i les de cada un representen classes diferents d'objectes lingüístics. Així, per exemple, s'estableix una diferència clara entre els signes lèxics i els sintagmàtics: representen dues classes d'objectes lingüístics i, conseqüentment, són de dos tipus diferents; d'aquesta manera, es pot establir la diferència entre la paraula *nenes* i el sintagma nominal *nenes*: la primera és representada per un signe lingüístic de tipus lèxic, mentre que la segona ho és per un de tipus sintagmàtic. Com que cada tipus de signe representa uns objectes lingüístics diferents, és natural que les seves estructures continguin informació diferent; per tant una caracterització dels diferents tipus d'estructures de trets comportarà una definició dels atributs que hi poden aparèixer i dels tipus que aquests poden tenir com a valor. Així, entre les estructures de trets que representen les dues interpretacions de *nenes* hi haurà, almenys, una diferència en el tret LEX: quan *nenes* és interpretat com a paraula el seu valor serà '+', mentre que quan ho és com a sintagma, serà '-'.

En un primer pas en l'organització del lèxic, podem establir una jerarquia en els nostres tipus, de manera que els més generals (que són els que tenen menys informació) apareguin en els punts superiors de la jerarquia. Així, aquesta estructura jeràrquica podria començar com s'indica a (21):

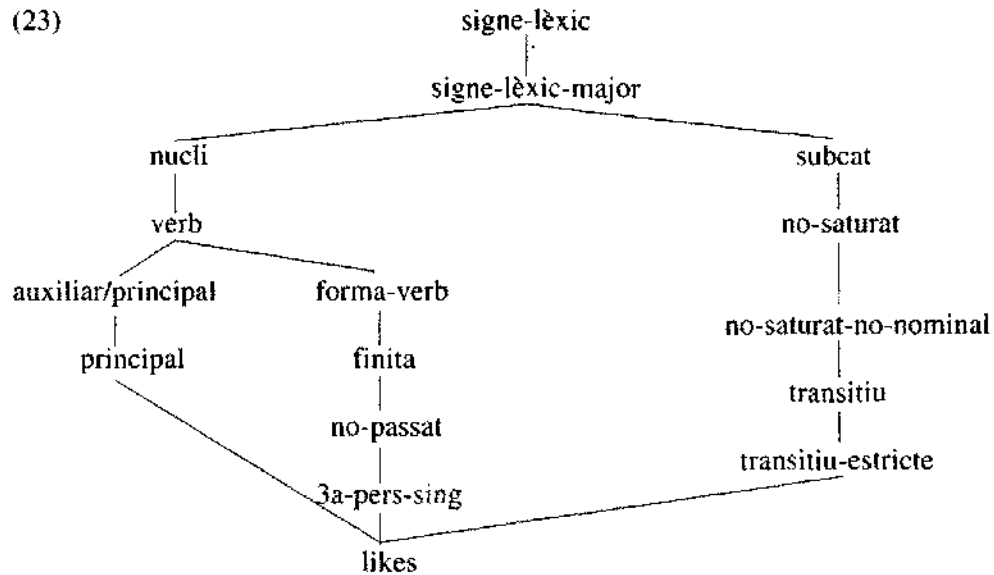


Cada tipus en la jerarquia conté una informació determinada. Per exemple, el tipus *signe lèxic* conté la informació  $\text{SINTAXILOCAL/LEX} = '+'$ , i el *signe lèxic major* nominal conté, entre altres, la informació  $\text{SINTAXILOCAL/NUCLI/MAJOR} = \text{N}$ . Ara bé, l'esquema de (21) és una simplificació que convertiria aquest tractament en inviable: és evident que els signes lèxics majors no es classifiquen només d'acord amb els seus trets nuclears, sinó que, almenys, s'han de tenir en compte també les seves propietats de subcategorització. Per tant, cal modificar l'esquema de (21) i convertir-lo en el de (22), que té en compte els dos criteris de classificació dels signes majors:



Entre els diferents tipus de la jerarquia s'hi estableix una relació anomenada de *subsumpció*, de manera que qualsevol element en la jerarquia subsumeix els tipus en posició inferior (per exemple, el tipus *signe lèxic major* subsumeix el tipus *nom*). La relació de subsumpció pot ser definida informalment com la inversa de la relació de *la seva informació és inclosa en*. La relació de subsumpció determina una estructura d'ordre per a tots els elements relacionats, que és la que representem amb la jerarquia. Cada element en la jerarquia rep o hereta la informació dels elements en posició superior (és a dir, dels elements que el subsumeixen) i, a més, com hem vist, pot contenir informació pròpia. Quan ens trobem amb una classificació creuada, com la de (22), els elements inferiors en la jerarquia, reben la informació que els prové dels dos braços de la jerarquia. Així, un verb com *likes* heretarà la informació pròpia del tipus *verb*, però també la del tipus *no saturat* (ja que es tracta d'un verb que té exigències de subcategorització). Com noten Pollard i Sag (1987:193), aquest tractament de la informació lèxica és molt habitual a les llengües amb propietats flexives importants: la forma concreta d'un temps verbal depèn del temps i la persona, per una banda, i del paradigma de conjugació, per l'altra.

D'aquesta manera es pot arribar a simplificar enormement la informació codificada per a cada entrada lèxica particular: de fet en la majoria dels casos la codificació consistirà en la determinació de la forma fonològica de la paraula en qüestió i de la seva caracterització semàntica (exactament, notem, els dos vessants del signe lingüístic segons Saussure); a més, només caldrà determinar en quin punt (o en quins punts) de la jerarquia s'hi ha d'integrar el signe, de manera que hereti tota la informació pertinent dels tipus que el subsumeixen. Vegem-ho amb un exemple. La posició de la forma verbal *likes* en la jerarquia és la indicada per l'esquema de (23), adaptat de Pollard i Sag (1987:205):



Un plantejament així s'anomena d'herència múltiple, perquè la informació que conté un element lèxic determinat prové de diversos camins en la jerarquia, com queda expressat gràficament a (23) en relació a la forma verbal *likes*.

No obstant, un sistema així, estructurat només a partir de l'herència, no és encara adequat. En primer lloc, no pot tractar les excepcions, és a dir, quan convé que l'element en qüestió hereti quasi tota la informació (però no tota) que li prové d'un camí determinat. A la vegada, no permet tampoc d'establir les relacions que hi ha entre moltes entrades lèxiques: per exemple, les que resulten de la flexió d'una única paraula (la base), o les que són conseqüència dels processos productius de derivació. És indubtable que hi ha una relació entre les formes verbals *cantar* i *canten* i, a més, que aquesta relació és general a molts parells de formes verbals; o que hi ha una relació general entre les formes actives i les passives dels verbs transitius, que és independent de les formes verbals concretes que s'estiguin considerant. Així doncs, de la mateixa manera que consideràvem natural que hi hagués uns mecanismes generals en el lèxic que permetessin evitar la redundància implícita en una descripció completa de moltes entrades lèxiques (les dels noms comuns, per exemple), convé que ens equipem també amb unes altres operacions que ens permetin de relacionar entrades lèxiques que tenen elements comuns. Aquestes relacions són expressades per les regles lèxiques. En concret a Pollard i Sag (1987) s'exemplifiquen les regles lèxiques per a la variació de formes provocades per la flexió i per a la passiva.

La idea general és la següent. La jerarquia de tipus és completa, de manera que caracteritza totes les classes de paraules. Per exemple, la part de la jerarquia de tipus lèxics corresponent al verb català tindrà en compte (adequadament estructurades) la totalitat de les formes verbals catalanes (en totes les seves variants de temps, persona, conjugació, etc). Ara bé, no tindria sentit haver d'afegir porcions de la jerarquia

específiques per a cada verb. El diccionari català contindrà les formes de base dels verbs, d'una manera molt semblant a com són llistats en un diccionari tradicional, amb la informació específica seva (almenys la fonològica i la semàntica). I seran les regles lèxiques les encarregades de generar les diverses formes flexionades, tot dotant-los de la informació mínima perquè el tipus resultant pugui heretar la informació pertinent a través de la jerarquia. Així per exemple, Pollard i Sag (1987:208) proposen l'estructura de trets de (24) com a entrada lèxica (simplificada) del verb anglès *like*:

(24) específica	<table style="border-collapse: collapse; width: 100%;"> <tr> <td style="padding: 2px;">fonologia:</td> <td style="padding: 2px;">like</td> </tr> <tr> <td style="padding: 2px;">sintaxillocallsubcat</td> <td style="padding: 2px;">&lt;[ ] 2, [ ] 1 &gt;</td> </tr> <tr> <td style="padding: 2px;">semànticalcontingut</td> <td style="padding: 2px;"> <table style="border-collapse: collapse;"> <tr> <td style="border-right: 1px solid black; padding: 2px;">relació</td> <td style="padding: 2px;">like</td> </tr> <tr> <td style="border-right: 1px solid black; padding: 2px;">arg1</td> <td style="padding: 2px;">1</td> </tr> <tr> <td style="border-right: 1px solid black; padding: 2px;">arg2</td> <td style="padding: 2px;">2</td> </tr> </table> </td> </tr> </table>	fonologia:	like	sintaxillocallsubcat	<[ ] 2, [ ] 1 >	semànticalcontingut	<table style="border-collapse: collapse;"> <tr> <td style="border-right: 1px solid black; padding: 2px;">relació</td> <td style="padding: 2px;">like</td> </tr> <tr> <td style="border-right: 1px solid black; padding: 2px;">arg1</td> <td style="padding: 2px;">1</td> </tr> <tr> <td style="border-right: 1px solid black; padding: 2px;">arg2</td> <td style="padding: 2px;">2</td> </tr> </table>	relació	like	arg1	1	arg2	2
fonologia:	like												
sintaxillocallsubcat	<[ ] 2, [ ] 1 >												
semànticalcontingut	<table style="border-collapse: collapse;"> <tr> <td style="border-right: 1px solid black; padding: 2px;">relació</td> <td style="padding: 2px;">like</td> </tr> <tr> <td style="border-right: 1px solid black; padding: 2px;">arg1</td> <td style="padding: 2px;">1</td> </tr> <tr> <td style="border-right: 1px solid black; padding: 2px;">arg2</td> <td style="padding: 2px;">2</td> </tr> </table>	relació	like	arg1	1	arg2	2						
relació	like												
arg1	1												
arg2	2												
	tipus per a heretar: {base, principal, transitiu-estricte}												

Aquesta entrada lèxica conté la informació específica de *like* i la indicació de quins són els tipus de la jerarquia dels quals ha d'heretar la resta de la informació (és un verb estrictament transitiu, principal, i de la forma que la gramàtica anglesa anomena la base). A la vegada, proposen una regla lèxica com la de (25) per a obtenir-ne la forma de la 3a persona del singular del present:

(25)	<table style="border-collapse: collapse; width: 100%;"> <tr> <td style="padding: 2px;">fonologia 1</td> </tr> <tr> <td style="padding: 2px;">sintaxillocallsubcat 2</td> </tr> <tr> <td style="padding: 2px;">semànticalcontingut 3</td> </tr> </table>	fonologia 1	sintaxillocallsubcat 2	semànticalcontingut 3	⇒	3a-pers-sing	<table style="border-collapse: collapse; width: 100%;"> <tr> <td style="padding: 2px;">fonologia f3A_SING (1)</td> </tr> <tr> <td style="padding: 2px;">sintaxillocallsubcat 2</td> </tr> <tr> <td style="padding: 2px;">semànticalcontingut 3</td> </tr> </table>	fonologia f3A_SING (1)	sintaxillocallsubcat 2	semànticalcontingut 3
fonologia 1										
sintaxillocallsubcat 2										
semànticalcontingut 3										
fonologia f3A_SING (1)										
sintaxillocallsubcat 2										
semànticalcontingut 3										

En aquesta regla, f3A\_SING és una funció que pren les formes fonològiques de base dels verbs anglesos i les retorna amb el fonema /s/ al final. Evidentment, la definició de la funció ha de ser més complexa per a tractar totes les variants fonològiques que es poden donar en la 3a persona del singular del present del verb anglès.<sup>13</sup> No obstant, la caracterització general és prou clara: la regla pren una estructura de trets d'un tipus determinat (*base*) i en retorna una d'un altre (*3a-pers-sing*) en la qual ha modificat una part de la informació original (en aquest cas, la fonològica). En ser el resultat de la regla una estructura de trets de tipus diferent de l'originària, hereta a través de la jerarquia d'herència múltiple tota la informació que li correspon com a exemplar del nou tipus. Així, en aquest exemple, el fet que el tipus resultant sigui *3a-pers-sing*, que és un subtipus de *no-passat*, fa que hereti, entre altres, la informació semàntica corresponent a la simultaneïtat temporal amb el moment de parla, com es pot veure a l'entrada lèxica expandida de (20). Així doncs, el conjunt de les entrades lèxiques de què disposa la teoria (com a material d'entrada per a l'aplicació de les regles o principis de la sintaxi i la semàntica) és el resultat del poder descriptiu i generatiu de les regles lèxiques i de les relacions d'herència múltiple entre tipus.

(13) A més, cal tenir en compte que la regla ha estat simplificada, ja que tal com està formulada no pot tractar les excepcions (com, per exemple, que la tercera persona del verb *have* no és *haves*).

Finalment, inclourem, simplificada també, la regla per a la passiva que proposen:

$$(26) \quad \left[ \begin{array}{l} \text{fonologia } 1 \\ \text{syntaxillocallsubcat} \\ \langle [ ] 3, [ ] 4 \rangle \\ \text{base \& transitiu} \\ \text{semànticalcontingut } 5 \end{array} \right] \Rightarrow \text{passiva} \left[ \begin{array}{l} \text{fonologia } \text{fpart\_pas} ( 1 ) \\ \text{syntaxillocallsubcat} \\ \langle (\text{sp}[\text{by}] 4), \dots, [ ] 3 \rangle \\ \text{semànticalcontingut } 5 \end{array} \right]$$

Aquesta regla accepta com a entrada estructures de trets de tipus *base* i *transitiu* i dóna com a sortida una nova estructura de trets del tipus *passiva* (del qual hereta la informació de tipus semàntic, per exemple), amb una forma fonològica diferent (la pròpia del participi passiu) i amb un esquema de subcategorització diferent també (amb el complement agent –un sintagma preposicional amb la preposició *by*– enlloc del subjecte).

A HPSG, doncs, el lèxic és plenament estructurat. Amb una simplificació evident del diccionari, es poden produir totes les estructures lèxiques que són necessàries per a un bon funcionament de la gramàtica en conjunt. A més, de manera similar a com ocorria a LFG, gràcies al mecanisme de les regles lèxiques, moltes relacions que en altres models gramaticals són tractades amb transformacions aquí poden ser resoltes de manera directa. En aquesta teoria els principis de projecció de les entrades lèxiques cap a les estructures sintàctiques són molt similars als de les teories anteriors; de fet, HPSG recull el tractament de la projecció de la informació lèxica que es va desenvolupar originàriament a partir de la teoria estàndard, formulat habitualment en el principi de la X-barrada. Com és natural, aquest marc teòric ha contribuït enormement a la revifalla actual dels estudis lexicogràfics, de manera que, en part, els plantejaments generals conflueixen amb els derivats de GB. Una bona mostra d'aquest fet la trobem, per exemple, en un volum com Sag i Szabolcsi (1992), en el qual conviuen autors sorgits de les dues tendències que es plantegem problemes gairebé idèntics, entre els quals destaquen l'estructura argumental i la naturalesa dels rols temàtics, els límits del lèxic, la naturalesa del significat lèxic, i la proposta d'alternatives lèxiques a anàlisis sintàctiques.

## 10. CONCLUSIÓ

Hem fet un llarg recorregut; la lingüística moderna ha fet un llarg recorregut. El paper central que tenia el lèxic en els primers models gramaticals que hem analitzat no ha estat retrobat fins a l'últim model que hem considerat. Ara bé, les diferències són tan grans que difícilment es pot parlar de paral·lisme: el que llavors eren consideracions informals, sense estructuració, s'ha convertit en un model estrictament formalitzat, que ha estat informatitzat; de fet, aquestes estructures no haurien sorgit si no hi hagués hagut una gran interacció entre la lingüística teòrica i la computacional.<sup>14</sup>

(14) Per a alguns no hi ha tal diferència: almenys parcialment, la lingüística computacional té com una de les seves finalitats el treballar cap a la consecució d'un tractament formalment adequat del llenguatge natural (que és, sens dubte, un dels objectius de la lingüística teòrica general).

Entre un i altre extrem, hi ha hagut un llarg període de sequera teòrica sobre el lèxic. Hi ha hagut tota una sèrie d'anys durant els quals semblava que la tasca lexicogràfica era una tasca menor, en oposició a l'activitat «reina» de la lingüística, la sintaxi. A la lexicografia li quedava la tasca de documentar la diversitat de formes i significats de què està format el lèxic d'una llengua, sense pretendre buscar-hi cap sistematicitat (la teoria lingüística mateixa ja s'encarregava de recordar la futilitat d'una tal recerca: el lèxic és el lloc on es consignen les idiosincràsies de la llengua).

Ara bé, vist des de la perspectiva actual, sembla que podem afirmar sense recança que aquesta mena de pas pel desert no ha estat debades. La sintactització de la lingüística ha comportat una rigorositat que no tenia. En la mesura que s'ha anat veient que les propostes estrictament sintàctiques podien ser millorades en incorporar-hi aspectes semàntics i en proposar-ne un tractament lèxic, la rigorositat que la sintaxi ha comportat no ha estat abandonada: no tindria sentit rebutjar per inadequada una proposta formal, sintàctica, per a adoptar-ne una altra, lèxica o semàntica, però informal; això no s'adiria gens als estàndards a què la lingüística generativa moderna ens ha anat acostumant i que són indispensables per a la lingüística computacional. Així doncs, el retorn al lèxic s'ha produït de la millor manera possible: s'ha simplificat molt el component sintàctic de la gramàtica, sense perdre la generalitat que aquesta havia aportat a la lingüística; les explicacions lingüístiques han guanyat en simplicitat (ja que les relacions lèxiques són en general més simples que les que es poden establir entre estructures sintàctiques); i s'ha retornat a una idea natural des del punt de vista psicològic, que el coneixement que els parlants tenim de la llengua és més un coneixement de les seves peces lèxiques que no pas de relacions entre estructures sintàctiques abstractes.

A la vegada, aquest canvi d'orientació està permetent un replantejament del paper de la semàntica en la lingüística. Un tractament amb estructures de trets implica l'eliminació de la separació dràstica entre la informació sintàctica i la semàntica. Ara bé, una perspectiva com aquesta pot servir també per a millorar el tractament de la semàntica lèxica dins la lingüística. La concepció de les jerarquies de tipus permet d'expressar les relacions semàntiques entre les peces lèxiques d'una manera força flexible, exactament igual que s'hi han pogut expressar les relacions sintàctiques que hem comentat. Com és natural aquest és un dels camps d'investigació actuals, especialment en l'àrea de la lingüística o lexicografia computacionals.

Des d'aquesta perspectiva, doncs, la lexicografia està cridada a esdevenir una branca essencial de la lingüística. A més d'haver-se de preocupar per la determinació de les formes d'organització del lèxic i per la seva estructura en general, molts aspectes morfològics, sintàctics i semàntics entren directament en el seu camp d'estudi, en tant que poden ser objecte de representació lèxica.

TONI BADIA

*Universitat Pompeu Fabra*

REFERÈNCIES BIBLIOGRÀFIQUES

- BÁEZ SAN JOSÉ, V. (1988) *Fundamentos de las gramáticas de dependencias*. Madrid, Síntesis.
- BARWISE, J.; PERRY, J. (1983) *Situations and attitudes*. Cambridge, Mass. MIT Press.
- BIERWISCH, M. (1966) *Der Strukturalismus. Kursbuch*. [traducció en castellà a Tusquets].
- BOTHA, R. P. (1968) *The function of the lexicon in transformational generative grammar*. La Haia, Mouton.
- BRESNAN, J. (1982) *The mental representation of grammatical relations*. Cambridge. MIT Press.
- CHOMSKY, N. (1957) *Syntactic Structures*. La Haia, Mouton.
- (1965) *Aspects of the Theory of Syntax*. Cambridge, Mass. The MIT Press.
- (1970) «Remarks on Nominalization». A: Jacobs, R.; P. Rosenbaum (eds.) (1970) *Readings in English Transformational Grammar*. Waltham, Mass., Ginn.
- (1981) *Lectures on Government and Binding*. Dordrecht, Foris.
- CORBIN, D. (1981) «Le monde étrange des dictionnaires (4): La créativité lexicale, le lexicographe et le linguiste». A: *Le dictionnaire*. Actes du Colloque Franco-Néerlandais. *Lexique 2*. PUF.
- DROSTE, F.; JOSEPH, J. (1991) *Linguistic Theory & Grammatical Description*. Amsterdam, Benjamin.
- FENSTAD, J. i al. (1987) *Situations, Language and Logic*. Dordrecht, Kluwer.
- FILLMORE, Ch. (1971) «Types of lexical information». A: Steinberg, D.; Jakobovits, L. (eds.) (1971) *Semantics*. Cambridge, Cambridge University Press. pàgs. 370-392.
- FLICKINGER, D. (1987) *Lexical rules in the hierarchical lexicon*. PhD diss. Stanford University.
- GEEST, W. de; JASPERS, D. (1991) «Government and Binding Theory». A: DROSTE, F.; JOSEPH, J. (1991).
- GRIMSHAW, J. (1990) *Argument structure*. Cambridge, Mass. MIT Press.
- JACKENDOFF, R. (1977) *X-bar syntax: A study of phrase structure*. Cambridge, Mass. MIT Press.
- LYONS, J. (1977) *Semantics*. 2 vols. Cambridge, Cambridge University Press.
- MONTAGUE, R. (1974) *Formal philosophy*. New Haven, Yale University Press.
- NEWMAYER, F. (1986) (2a ed.) *Linguistic Theory in America*. Nova York, Academic Press.
- PEREIRA, F. C. N.; WARREN, D. H. D. (1980) «Definite clause grammars for language analysis». A: *Artificial Intelligence*, 13, pàgs. 231-78.
- POLLARD, C.; SAG, I. (1988) *An information-based approach to syntax and semantics: Volume I. Fundamentals*. CSLI Lecture Notes No. 13. Stanford, Califòrnia.

- POLLARD, C.; SAG, I. (en premsa) *Topics in constraint-based syntactic theory*.
- SAG, I.; SZABOLCSI, A. (1992) *Lexical matters*. CSLI. Stanford, Califòrnia.
- SAUSSURE, F. de (1915) *Cours de linguistique générale*. [traducció en català: Barcelona. Ed. 62. 1990.].
- SCIULLO, A.-M. di; WILLIAMS, E. (1987) *On the definition of word*. Cambridge, Mass. MIT Press.
- SELLS, P (1985) *Lectures on Contemporary Syntactic Theories*. CSLI.
- SHIEBER, S. M. (1986) *An introduction to unification-based approaches to grammar*. CSLI.
- SOMMERFELDT, K.-E.; SCHREIBER, H. (1974) *Wörterbuch zur Valenz und Distribution deutscher Adjektive*. Leipzig.
- TESNIÈRE, L. (1959) *Éléments de syntaxe structurale*. París, Klincksieck.
- TUSÓN, J. (1980) *Teorías sintácticas y análisis sintáctico*. Barcelona, Teide.
- WESCOAT, M. T.; ZAENEN, A. (1991) «Lexical Functional Grammar». A: DROSTE, F.; JOSEPH, J. (1991).
- WILLIAMS, E. (1981) «Argument structure and morphology». A: *Linguistic Review*, 1, pàgs. 81-114.
- ZUBIZARRETA, M.-L. (1986) *Levels of representation in the lexicon and syntax*. Dordrecht, Foris.