

Etiquetatge sintacticosemàntic en diccionaris electrònics multilingües

Xènia Escolano Marín
Universitat d'Alacant

1. Introducció

L'interès pels sistemes de traducció automàtica comença a augmentar a partir dels anys seixanta a Estats Units i a Europa, especialment a França. Arran d'aquest fet, el lingüista francès Maurice Gross (1934-2001) funda el 1968 el Laboratoire d'Automatique Documentaire et Linguistique (LADL) i emprén una descripció sistemàtica de les propietats sintàctiques i semàntiques de tots els elements del lèxic francès d'acord amb les formulacions de la primera gramàtica distribucional i transformacional de Zellig Harris (1909-1992), precursor dels fonaments teòrics del lèxic-gramàtica. Anys després, el gramàtic Gaston Gross (1939), antic director del Laboratoire de Linguistique Informatique (LLI) –actual LDI (Lexiques, Dictionnaires, Informatique) de la Universitat París 13–, desenvolupa la noció de les classes d'objectes, model lingüístic essencial per al tractament automàtic del llenguatge (TAL) que incorporen els diccionaris electrònics del LDI.

El LADL ha realitzat aportacions fonamentals en l'àmbit de la lingüística informàtica mitjançant la creació d'un sistema complet de diccionaris electrònics i classificacions exhaustives de les unitats lèxiques del francès. Es tracta del sistema DELA, que comprén: el DELAS/ DELAF, diccionari de les formes simples i diccionari de les formes flexionades, amb més de 90.000 entrades (Blondine Courtois); el DELAC/ DELACF, diccionari de les formes compostes i diccionari de les formes compostes flexionades, amb més de 100.000 noms compostos, i el DELAP/ DELAPF, diccionari fonèmic on figuren, entre altres aspectes, les transcripcions fonètiques i les possibilitats d'elisió (Eric Laporte). D'aquesta manera, tot i que el camp de la lingüística informàtica es troba en auge i cada vegada són més les llengües que s'implementen en els sistemes TAL, podem afirmar que el francès és la llengua que compta amb una de les descripcions lingüístiques empíriques més rigoroses.

En aquesta línia, cal destacar que diversos grups d'investigadors que formen part de la xarxa Relex (<http://infolingu.univ-mlv.fr/Relex/dictionaries.html>) es dediquen actualment al desenvolupament de diccionaris electrònics que segueixen el sistema DELA del LADL¹. La construcció d'aquests diccionaris electrònics que compten amb un etiquetatge sintacticosemàntic es revela especialment pertinent per als sistemes de traducció automàtica del llenguatge de modalitat híbrida, els quals combinen estadística amb sintaxi i semàntica. En l'actualitat, podem dir que els dos sistemes TAL més representatius són Systran i Sadaw.

En el present article presentem en què consisteix l'etiquetatge sintacticosemàntic en diccionaris electrònics multilingües a partir del disseny del mòdul d'un diccionari electrònic monolingüe coordinat en espanyol amb equivalents en anglès² que reuneix la descripció del conjunt de propietats lingüístiques necessàries per al processament automàtic

¹ Segons assenyala Mogorrón (2004, 390-391), trobem a Portugal, el grup dirigit per E. Ranchood en la Universitat de Lisboa, a Itàlia, el dirigit per A. Elia en la Universitat de Salerno, a Espanya, dirigit per Xavier Blanco, el grup de lingüística aplicada a les llengües romàniques de la Univesitat Autònoma de Barcelona, i en alemany, el sistema CISLEX, diccionari electrònic coordinat pel CIS Centrum für Informations und Sprachverarbeitung (LMU-München), que també ha seguit el model del DELA

² Aquest tipus de diccionaris electrònics no es consideren bilingües, sinó monolingües coordinats, atès que el contingut dels camps corresponents a cada unitat lèxica de la llengua origen no resulta directament transferible a la llengua meta (Blanco 2001).

del llenguatge natural. Per a aquesta fi, hem procedit a l'elaboració de la classe d'objectes <prendas de vestir> en espanyol mitjançant la delimitació lingüística dels seus elements segons el marc teòric del lèxic-gramàtica.

2. El model de les classes d'objectes

Per tal d'explicar en què ha consistit l'elaboració de la classe d'objectes <prendas de vestir> i com hem etiquetat els elements que en formen part, en aquest apartat introduïm com a nocions preliminars una síntesi dels principals aspectes d'aquest model teòric desenvolupat per G. Gross (1992). El model de les classes d'objectes es basa en els treballs de Z. Harris (1971) i en la teoria del lèxic-gramàtica de M. Gross (1975) i consisteix en la descripció de la llengua a partir de grups semàntics que es regeixen per un mateix comportament sintàctic amb el propòsit principal de facilitar el processament automàtic del llenguatge natural.

2.1. Estructura de la frase simple: predicats i arguments

La primera etapa de l'anàlisi automàtica d'un text consisteix a reconèixer les frases que el constitueixen. En lingüística formal, la frase simple es divideix en predicats i en arguments³. El predicat, de forma variable, introdueix arguments, representats per substantius o frases. En una frase com *Joan llig el periòdic*, *lligir* representa un predicat verbal acompanyat a la seua esquerra per l'argument *Joan*, en posició de subjecte, i a la seua dreta, per l'argument *el periòdic*, en posició d'objecte. Per a Harris (1976, *apud* Gross 2012, 37), un predicat és “una palabra que opera una selecció determinada entre las restantes palabras del léxico para formar un esquema que sea la base de una aserción”. Els arguments els defineix com a “los elementos léxicos seleccionados por los predicados en el marco de una frase simple”. Així doncs, segons Harris (1976), mentre que el predicat selecciona el grup semàntic al qual pertanyen els substantius que l'acompanyen, en aquest cas, la classe dels humans (*Joan*) i la dels <textos> (*el periòdic*), els arguments no delimiten la classe de predicats que els introdueixen, ja que un periòdic, a més de *lligir-se*, també es pot *escriure*, *publicar*, *criticar*, *comprar*, *vendre*, etc.

2.2. L'ús d'un predicat

En aquest context resulta fonamental la noció d'ús. L'ús d'un predicat representa un determinat esquema predicatiu que porta associat un conjunt de propietats i que es defineix, sobretot, pel seu àmbit d'arguments (Gross 2012, 57). De fet, per a Gross (2012, 67), les entrades d'un diccionari haurien de correspondre a un ús específic i no a un lexema –com ocorre en els diccionaris tradicionals–, ja que cada esquema predicatiu presenta un significat concret “estrechamente relacionado con un conjunto de propiedades sintácticas, morfológicas y pragmáticas que forman parte de su definición lingüística global” (2012, 70). Vegem un exemple dels usos d'un verb polisèmic, indicats per mitjà de classes, juntament amb els seus equivalents de traducció:

1. *abatir*/N0:<acontecimiento>/N1:hum⁴/E:to demoralize
2. *abatir*/N0:hum/N1:hum, animal/E:to shoot down
3. *abatir*/N0:hum/N1:inanimado, concreto/E:to knock down
4. *abatir*/N0:hum/N1:<cartas>/E:to lay down
5. *abatir*/N0:<medio de transporte náutico>/E:to turn off
6. *abatir*/N0:hum/N1:<figura geométrica>/E:to project
7. *abatir*/N0:hum/N1:<artefacto>/E:to dismantle

(Gross 2012: 58)

³ En lexicografia electrònica també es coneixen com a *operadors* i *objectes* respectivament.

⁴ N0 correspon al subjecte i N1 a l'objecte.

La selecció d'un ús predicatiu especifica els trets lingüístics que el predicat en qüestió porta associats en contraposició a la resta d'esquemes que pugui tenir. Així, el fet que els diccionaris bilingües o multilingües en suport electrònic disposen d'aquestes dades permet monosemitzar les formes polisèmiques i que el processament automàtic de frases d'un text genere les traduccions correctes en una determinada llengua meta.

2.3. De les hiperclasses a les classes d'objectes

Per reconèixer els diversos usos d'un predicat resulta fonamental classificar els arguments que l'acompanyen tenint en compte els seus trets sintacticosemàntics, també coneguts com a *hiperclasses* (Gross 2012, 98). D'acord amb Gross (2012, 100), són sis les classes que caracteritzen els arguments elementals: noms humans, noms animals, noms de vegetals, noms concrets, noms locatius i noms de temps. A aquestes sis classes, Le Pesant i Mathieu Colas (10) afegeixen quatre més relatives als predicats: *humà predicatiu*, *acció*, *estat* i *esdeveniment*. No obstant això, aquests trets sintacticosemàntics resulten insuficients. A mode d'il·lustració, presentem un exemple de Man-Ghyu (155-156): si bé els trets mínims de *concret* i *abstracte* permeten traduir correctament a l'anglès dos dels usos del verb francès *porter*, aportant com a equivalent *to bear* quan introdueix un objecte abstracte (*Luc porte un nom français = Luc bears a French name*) i *to carry* quan es tracta d'un objecte concret (*Luc porte une valise = Luc carries a case*), no resulten esclaridors en altres usos en què l'objecte també és concret. En cas de referir-nos a una peça de vestir, el verb *porter*, que és l'apropiat en francès, ja no es tradueix per *to carry*, sinó per *to wear* (*Luc porte un chapeau = Luc wears a hat*).

D'aquesta manera, amb la finalitat d'obtenir la traducció correcta de l'últim cas, cal afegir el tret *peça de vestir*. Gross (2012, 101) denomina aquests trets semàntics, establits en funció de la coherència entre predicats i arguments, *classes d'objectes*. Perquè en aquest exemple l'operació de traducció tinga èxit és necessari, en primer lloc, la codificació de *barret* amb l'etiqueta <*peça de vestir*>, i, en segon lloc, la introducció en el sistema d'una regla contextual, segons la qual, quan el verb *portar* tinga com a complement un substantiu relatiu a una peça de vestir, es tradueix per *to wear*, i no per *to carry* (Le Pesant i Mathieu-Colas 28).

2.4. La descripció dels operadors apropiats

Una classe d'objectes queda definida per les relacions sintàctiques que manté amb una o diverses classes de verbs, anomenats *operadors apropiats* (Gross 2012, 101). Per a identificar amb precisió una classe d'objectes, es sol recórrer a diversos predicats, normalment dos o tres verbs. En el cas que ens ocupa, segons Le Pesant i Mathieu-Colas (13), la conjunció dels següents tres predicats permet delimitar la classe <*vêtements*> (<*peces de vestir*>): *mettre/N0:hum/N1:vêtement*, *être en/N0:hum/N1:vêtement* y *aller bien à/N0:vêtement/N1:hum*. Tot i que no necessàriament han de ser aquests verbs els únics delimitadors de la classe, el que sí és evident és que el verb *porter* (*portar*) per si mateix, resulta insuficient per a confeccionar la classe <*vêtements*>, ja que com que accepta diversos tipus d'objectes concrets, cal afegir altres predicats que ajuden a delimitar la classe.

Així mateix, hem de destacar que de la mateixa manera que l'anàlisi de les classes d'arguments permet desambiguar la polisèmia eventual dels predicats, com hem comprovat en l'exemple de *to wear* en contraposició a *to carry*, els predicats apropiats possibiliten igualment la desambiguació dels substantius polisèmics que actuen com a arguments (Le Pesant i Mathieu-Colas 28). Per exemple, si emetem dos enunciats com *aquest coll és còmode* i *aquest banc és nacional*, els predicats adjectivals *còmode*

i *nacional* permeten seleccionar l'ús pertinent dels seus respectius arguments, atès que *còmode* es diu d'una peça de vestir, no d'una part del cos, i *nacional* s'aplica a una entitat financera, i no a un seient.

2.5. La jerarquia de les classes i les relacions d'herència

Com és sabut, els hiperònims presenten trets definitoris inherents a una classe semàntica, i les paraules que s'adscriuen a aquesta classe constitueixen els seus hipònims. Segons aquesta definició, podem dir que *margarida*, *rosa* i *tulipa*, per exemple, constitueixen hipònims de l'hiperònim *flor*. Tanmateix, l'absència de criteris sintàctics en les xarxes semàntiques que representen els conceptes tradicionals d'aquests termes pot impedir que un hipònim es subordine a més d'un hiperònim, problema que no afecta a les classes d'objectes. Un exemple seria el cas d'una professió, com ara *metge* o *advocat*: si considerem la sintaxi que regeixen aquests objectes, resultarien subordinables a més d'una classe, no sols a les professions, sinó també als noms humans predicatius relacionals, ja siguin de parentesc (*pare*), espacials (*veí*) o professionals (*col·lega*). Comprovem, doncs, que aquesta nova lectura de la *hiperonímia* permet que els hipònims hereten els predicats que s'atribueixen als seus diversos hiperònims. Aquest fenomen es coneix com a "herència múltiple" (Le Pesant i Mathieu-Colas 16) i resulta essencial per una descripció precisa de la llengua i un bon processament automàtic d'aquesta.

La idea d'"herència múltiple" implica la constitució de subclasses dins d'una classe d'objectes, la qual cosa també comporta la identificació del tipus de relació que s'estableix entre les classes de predicats i les classes d'arguments d'una determinada classe d'objectes. Gross (2012, 105-110) divideix les classes d'objectes en classes de concrets, en classes de locatius i en classes de noms de temps. La classe dels concrets, a la qual pertanyen les <peces de vestir>, parteix d'una primera classificació basada en la dicotomia d'objectes naturals enfront d'artefactes o objectes fabricats per l'home. Al contrari dels objectes naturals, els artefactes es caracteritzen per comptar amb una finalitat, amb la qual cosa accepten predicats de creació (Gross 2012, 105), com *fabricar*, *produir* o *realitzar*. Una vegada establerta aquesta distinció, s'elaboren subclasses, com la dels <aliments>, els <mitjans de transport> o la <roba>, que poden disposar de diverses subclasses que, al seu torn, poden contenir altres subclasses, que hereten les propietats de totes les classes a les quals es subordinen.

Si apliquem aquesta descripció a la classe de les <peces de vestir>, arribem a la mateixa constatació. Per exemple, el conjunt d'adjectius relatius a la mida (*ajustat*, *gran*, *petit*, *estret*, etc.), manté una relació genèrica respecte al conjunt dels substantius de les <peces de vestir>, ja que són extensibles a totes les subclasses que dominen. Dins d'aquesta associació comú, s'estableixen relacions més específiques, per exemple, i seguint amb predicats adjectivals, *embetunat* i *de punta quadrada* constituïrien operadors apropiats a la subclasse del <calçat>, i *de pines* i *acampanat* serien exclusius dels <pantalons>, pertanyents a la subclasse de peces de vestir <inferiors>. Igualment, es poden trobar predicats verbals únicament operatius per a una subclasse determinada, com ara *encasquetar-se* per als <barrets>, *calçar* per al <calçat> o *quedarse en* per a la <roba interior>, i, tot i així, tots hereten els operadors generals de la classe, com *portar*, *llevar-se*, *vestir* o *canviar de*, així com les propietats generals de les classes dels concrets, com les relatives al pes o al color.

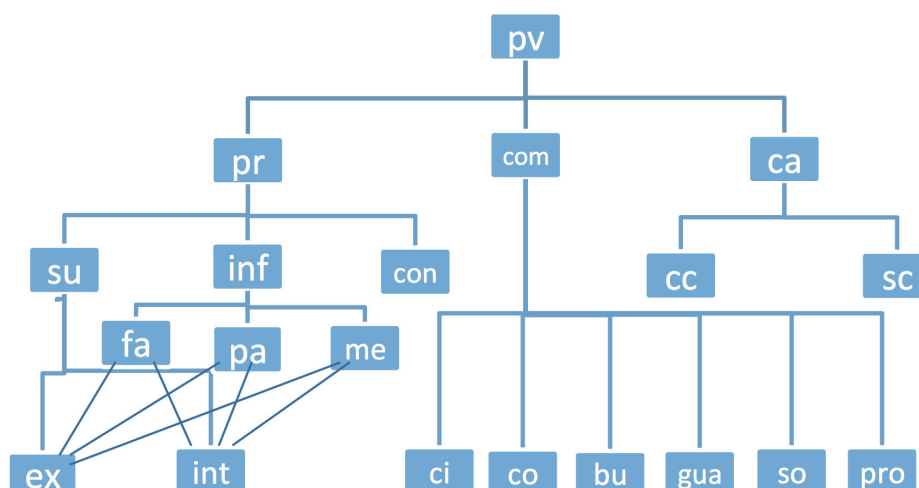
3. La classe d'objectes <prendas de vestir> en espanyol

3.1. Generalitats

La classe d'objectes <prendas de vestir> en espanyol s'inscriu dins de la dels artefactes, que pertanyen al seu torn a la hiperclasse dels concrets, que són inanimats. Com tots

els artefactes, compten amb una finalitat, que en aquest cas és, principalment, la de cobrir el cos humà. Tal com hem comentat abans (cf. 2.4.), aquesta classe pot quedar delimitada lingüísticament mitjançant la intersecció dels operadors *ponerse/N0:prenda/N1:hum*, *estar en/N0:prenda/N1:hum* i *sentar bien/N0:prenda/N1:hum*.

Una vegada hem delimitat la classe i establert la jerarquia general –representada com a *inanimado>concreto>artefacto>prenda de vestir*–, podem procedir a l’elaboració de diverses subclasses que deriven d’aquesta classe, que, al seu torn, poden comptar amb altres subclasses que hereten les propietats de totes les subclasses a les quals es subordinen (cf. 2.5). A cada subclassa se li ha de poder atribuir una definició sintàctica mínima (Le Pesant i Mathieu-Colas 13), és a dir, tota subclassa ha de disposar, almenys, d’un predicat apropiat als objectes que s’hi adscriuen, la qual cosa permet delimitar-la lingüísticament i, per tant, justificar-ne la creació, tal com veurem en l’apartat 3.3.2. L’elaboració d’aquestes subclasses dóna lloc a la jerarquia específica de la classe *<prendas de vestir>*, que podem descriure mitjançant el següent arbre construït en funció de criteris sintacticosemàntics (*vid. infra* llegenda):



pv = peces de vestir
 pr = peces de roba
 su = peces superiors
 inf = peces inferiors
 fa = faldes i faldes pantaló
 pa = pantalons i granotes
 me = mitges, leggings, malles, calcetins i calces
 ex = peces d'ús exterior
 int = peces d'ús interior
 com = complements

ci = cinturons i tirants
 co = corbates i corbatins
 bu = bufandes, mocadors i xals
 gua = guants
 so = barrets, cascs, casquets, gorres, cintes i corones
 pro = protectors (ex. genollera, colzera)
 ca = calçat
 cc = calçat amb cordons
 sc = calçat sense cordons

La descripció de la classe que delimitem comporta dos tipus de fitxers que representen la nostra base de dades de *<prendas de vestir>*, per a la creació de la qual seguim les convencions de Gross (1992). El primer fitxer està constituït pels arguments o objectes (noms simples i compostos) i el segon, pels predicats o operadors apropiats (adjectius, verbs, substantius predicatius, preposicions i locucions verbals).

3.2. Arguments

3.2.1. Camps del fitxer

D'acord amb Gross (1992, 255), els camps d'un fitxer d'un diccionari electrònic es poden classificar segons continguin informacions de naturalesa morfològica, sintacticosemàntica i sintàctica.

semàntica, diasistemàtica (de les quals fem abstracció en aquest cas) i referent a les traduccions. Analitzem a continuació un dels objectes inventariats en el nostre fitxer d'arguments:

bata de médico/G:nm/M:N de N/T:inc/C:<pv:su-ex>/D:uso profesional/SD:medicina/En:doctor's gown

El primer camp correspon al lema en espanyol, i el segon (G), a la seua categoria gramatical (*nm* per a substantius masculins, *nf* per als femenins i *nmp* o *nfp* per al caràcter obligatori del plural, p.ex. en *rompevientos*, *gafas de buceo* o *picardías*). El camp M dona compte dels tipus morfològic dels substantius compostos (NN, N de N, NAdj, Vn, N Prep N, AdjN, N Prep N, etc.). El quart camp (T) fa referència al tret sintacticosemàntic de l'entrada, en el cas de les peces de vestir, tots els objectes son inanimats concrets (*inc*). El cinqué camp (C) indica les classes i subclasses dels objectes, en el cas de *bata de médico*, *<pv:su-ex>* assenyala que es tracta d'una peça de vestir superior i d'ús exterior. Els camps sisé i seté defineixen l'àmbit (D) i subàmbit (SD) del lema. En aquest exemple, indiquem *de uso profesional* com a àmbit i *médico* com a subàmbit. El huité camp (En) correspon als equivalents de traducció del lema. En el nostre fitxer oferim sempre la traducció en anglés britànic en nom de la sistematicitat.

Així mateix, afegim un camp referent a la plausibilitat del lema, és a dir, el grau d'admissibilitat de l'objecte amb un rang d'1 a 3, en el qual 1 equival a una peça d'ús comú (p.ex. el cas mateix de *bata de médico* o *anorak*), 2 a una peça que coneixem però d'ús menys generalitzat (p.ex. *mañanita*), i 3, a una peça que rarament anomenariem a menys que ens trobarem en el respectiu àmbit d'especialitat (p.ex. *bashlik* o *carbatina*).

3.2.2. Noms simples i noms compostos

El fitxer d'arguments conté 714 entrades, 329 de noms simples i 385 de noms compostos amb 11 formes morfològiques distintes. Un nom simple seria *sujetador/G:nm/T:inc/C:<pv:su-int>/D:vestimenta/En:bra* i un nom compost *casco de boxeo/G:nm/M:N-deN/T:inc/C:<pv:so>/D:deporte/SD:boxeo/En:boxing helmet*

La qüestió que ací ens plantejem és fins a quin punt un lema és simple o compost, és a dir, si l'adjectiu que acompanya un substantiu forma part del lema i conforma, doncs, una unitat polilèxica, o si, en canvi, es tracta d'un operador que funciona com a adjectiu predicatiu i, per tant, no forma part de la unitat, que és simple. Tot i que ací no entrarem en detalls per no estendre'ns, cal remarcar que quan l'adjectiu es pot aplicar a diversos objectes i no assenyala cap característica inherent al substantiu que acompanya constitueix un adjectiu predicatiu (p.ex. *con vuelo, plisado*), mentre que quan només s'aplica a un únic nom sense ser predicatiu (**este zapato es ortopédico, *este vestido es de noche*⁵) haurà de figurar juntament amb el substantiu en el fitxer d'arguments en qualitat de nom compost.

3.3. Predicats

3.3.1. Camps del fitxer i característiques dels predicats

Aquest fitxer està format per una llista de 487 predicats apropiats als objectes inventariats en el fitxer d'arguments. Hi trobem verbs predicatius simples i compostos (també anomenats *locucions verbals*), adjectius predicatius simples i compostos i preposicions

⁵ Si generem construccions predicatives en els casos concrets de *falda escocesa* i *maillot amarillo* es produeix un canvi de significat: en *esta falda es escocesa*, *escocesa* fa referència a la procedència de la *falda*, però ja no es tractaria del nom compost que designa la falda de quadres típica d'Escòcia i Irlanda que vesteixen el homes en determinades celebracions. De la mateixa manera, en la frase *este maillot es amarillo*, *amarillo* deixa de ser l'adjectiu que forma part de la unitat polilèxica que descriu la peça de roba superior que porta el ciclista que lidera el Tour de França, i passa a ser un adjectiu predicatiu que designa un color, aplicable a qualsevol objecte inanimat concret.

predicatives. A continuació, oferim el verb *embutirse en* per il·lustrar la forma mínima que presenta l'entrada d'un predicat:

embutirse en/G:v/N0:hum/N1:<pv:fa-pa>/En:to stuff o.s. into

El primer camp correspon al predicat i el segon assenyala la seua categoria gramatical: *v, v_c, adj, adj_c, n, n_c* i *prep* (verb simple, verb compost, adjectiu simple, adjectiu compost, nom simple, nom compost i preposició respectivament). L'últim camp indica, com en el cas dels arguments, la traducció del predicat en anglés.

Els camps N0, N1 i N2 donen compte de la naturalesa sintacticosemàntica dels arguments del predicat en qüestió: N0 correspon al subjecte del predicat; N1, a l'objecte directe o al complement principal, i N2, a l'objecte indirecte; és a dir, assenyalen les restriccions de selecció que el predicat imposa sobre les diverses posicions argumentals. En la classe *<prendas de vestir>*, el subjecte N0 pot correspondre a un humà, com en l'exemple de *embutirse en*, i, per tant, el complement N1 es refereix a una peça de vestir, que codifiquem segons les abreviacions exposades en l'arbre sintacticosemàntic (cf. 3.1.); o bé, la peça de vestir actua com a N0 i no hi ha complement N1, com ocorre, per exemple, amb *arrastrar* (*arrastrar/G:v/N0:<pv:pr>/En:to sweep the ground*) i *desteñirse* (*desteñirse/G:v/N0:<pv>/En:to lose colour*). No obstant això, trobem cinc verbs en l'estructura argumental dels quals la peça de vestir actua com a N0 i un humà com a N1: *favorecer, venir, sentar bien, sentar mal* i *rozar*.

Així mateix, pot donar-se el cas que un mateix predicat presente diverses construccions en les quals les posicions argumentals varien, tal com ocorre amb el verb *encoger*: *encoger/G:v/N0:hum/N1:<pv:pr>/En:to shrink* (p.ex. *María ha encogido la camiseta al meterla en la secadora*) i *encoger/G:v/N0:<pv:pr>/En:to shrink* (p.ex. *La camiseta ha encogido por haberla metido en la secadora*). Així, quan un predicat compta amb dues construccions, cadascuna equivaldrà a dos usos diferents del predicat, que, per tant, figuraran en dues entrades distintes del fitxer de predicats.

Per finalitzar amb les característiques relatives a la representació dels verbs en el fitxer de predicats, cal remarcar dos aspectes més. D'una banda, quan un verb puga adoptar la forma reflexiva i la no reflexiva, també correspondrà a dos usos distints, per exemple: *abrocharse/G:v/N0:hum/N1:<pv:pr-ci-ca>/En:to fasten up*, on únicament trobem el complement N1 com a objecte directe (*un humano se abrocha una prenda de vestir*), i *abrochar/G:v/N0:hum/N1:<pv:pr-ci-ca>/N2:hum/En:to fasten up*, on assenyalen que la construcció regeix un complement N2 humà que actua com a objecte indirecte (*un humano le abrocha una prenda a otro humano*). I de l'altra, observem que un nombre considerable de verbs es caracteritzen per presentar construccions converses que regeixen entre si els mateixos arguments en les mateixes posicions argumentals. Aquests verbs poden formar l'antònim bé mitjançant un recurs morfològic, afegint el prefix *-des*, com en *abotonar(se)* i *desabotonar(se)*; bé mitjançant un recurs lèxic, com en *apretar(se)* i *aflojar(se)* o en *ponerse de moda* i *pasarse de moda*.

Pel que fa als noms predicatius, remarquem que compten amb un camp propi (W), on indiquem el verb suport⁶ amb què es combinen, acompanyat de *DET* en cas que s'empren amb un determinant⁷. En aquestes construccions, la peça de vestir pot actuar com a subjecte del predicat (p.ex. *agujero/G:n/N0:<pv>/W:hacerse DET/En:to get a hole in*), o

⁶ Els verbs suport són verbs buits semànticament (*fer, realitzar, tenir, ser, estar, donar, rebre*, etc.) que permeten actualitzar els noms predicatius, la qual cosa és fonamental per al TAL i, concretament, per a la traducció (Gross 1992, 257).

⁷ Per a major claredat, en el camp *En* incloem el verb suport, les preposicions i els arguments que regeix el nom predicatiu en anglés, tot i que formalment les informacions corresponents a l'equivalent de traducció les ha de proporcionar el diccionari de la L2 en qüestió (Blanco 1999: 9).

bé com a complement directe (p.ex. *confección/G:n/N0:hum/N1:<pv>/W:realizar DET/En:to do the dressmaking of*). Tanmateix, alguns noms predicatius presenten diverses construccions, on la peça de vestir pot ocupar tant la posició N0 com la N1 (p.ex. *diseño: diseño/G:n/N0:<pv>/W:ser DET/En:to be the design of* i *diseño/G:n/N0:hum/N1:<pv>/W:crear DET/En:to create de design of*).

Finalment, cal afegir que la meronímia també representa un recurs productiu a l'hora de delimitar lingüísticament determinades subclasses. D'aquesta manera, disposem de diversos merònims (parts d'una peça de vestir) –inventariats en el fitxer d'arguments– que resulten apropiats exclusivament a una determinada subclasse, per exemple, *bragueta* per a les peces de roba inferiors (*mer_<pv:inf>*).

3.3.2. Predicats apropiats a <prendas de vestir>

Els predicats apropiats a totes les peces de vestir en general (<pv>) i que corresponen, per tant, al primer nivell de la jerarquia de la classe (a la qual dóna nom), són molt nombrosos. Es tracta de predicats amb un ampli espectre, és a dir, que presenten poques restriccions de selecció perquè mantenen una relació privilegiada respecte al conjunt de tots els substantius que figuren en el fitxer d'arguments, tal com hem indicat anteriorment (cf. 2.5.). Trobem, sobretot, adjectius relatius al material de la peça de vestir, a la seua mida, al seu estil i aspecte o a l'ocasió en què s'utilitza; verbs simples genèrics que no fan referència a la part del cos que cobreix la peça de roba, verbs compostos (o locucions verbals) i noms predicatius aplicables a qualsevol tipus de peça de vestir, i la preposició *en*. Alguns exemples d'aquestes categories són els següents:

- Adjectius: *abombado/G:adj/N0:<pv:pr>/En:puffed; de niño/G:adj_c/N0:<pv>/En:children's*
- Verbs: *abrigarse con/G:v/N0:hum/N1:<pv:pr-bu-gua-so-ca>/En:to wrap o.s. up in; lucir/G:v/N0:hum/N1:<pv>/En:to dress up in*
- Locucions verbals: *sentar como un guante/G:v_c/N0:<pv>/En:to fit like a glove; ponerse de moda/G:v_c/N0:<pv>/En:to become in fashion*
- Noms predicatius: *pepunte/G:n/N0:<pv>/W:tener DET/En:to be backstitched// pepunte/G:n/N0:hum/N1:<pv>/W:dar DET/En:to do a backstitch on; agujero/G:n/N0:<pv>/W:hacerse DET/En:to get a hole in// agujero/G:n/N0:hum/N1:<pv>/W:hacerse DET/En:to make a hole in*
- Preposició: *en/G:prep/N0:hum/N1:<pv>/En:in*

3.3.3. Predicats apropiats a <prendas de ropa>

En aquest segon nivell de la jerarquia (<pr>), també trobem alguns predicats que poden ser, al seu torn, apropiats als objectes que pertanyen a altres subclasses:

- Adjectius: *abombado/G:adj/N0:<pv:pr>/En:puffed*
- Verbs: *abotonarse/G:v/N0:hum/N1:<pv:pr>/En:to button*
- Noms predicatius: *planchado/G:n/N0:hum/N1:<pv:pr-co-bu-gua-so>/W:realizar*

3.3.3.1. Predicats apropiats a peces de roba <superiores>

Les peces de roba superiors (<su>) constitueixen una subclasse lingüísticament ben delimitada. A mode d'exemple, ho podem observar en els següents casos:

- Adjectius: *escotado/G:adj/N0:<pv:su>/En:low-necked*
- Verbs: *echarse sobre los hombros/G:v_c/N0:hum/N1:<pv:su>/En:to drape Poss <pv:su> over Poss shoulders*

3.3.3.2. Predicats apropiats a peces de roba <inferiores>

Alguns dels predicats apropiats a la subclasse de les peces de roba inferiors (<inf>) resulten també adequats per a les peces de roba superiors:

- Adjectius: *por la rodilla*/G:adj_c/N0:<pv:su-inf>/En:half-length
- Verbs: *llegir hasta la rodilla*/G:v_c/N0:<pv:su-inf>/En:to go knee length

3.3.3.2.1. Peces de roba <faldas>

Els predicats apropiats específicament a la subclasse de les falces (<fa>) poden ser adjectius, verbs i noms predicatius. Per exemple:

- Adjectius: *con vuelo*/G:adj_c/N0:<pv:fa>/En:flared
- Verbs: *levantarse*/G:v/N0:<pv:su-fa>/En:to ripple
- Noms predicatius: *doblado*/G:n/N0:hum/N1:<pv:fa-pa>/W:coser DET/En:to sew the turn-up of

3.3.3.2.2. Peces de roba <pantalones>

La subclasse dels pantalons (<pa>) està molt ben definida lingüísticament, ja que, a més de comptar amb adjectius i verbs simples i compostos, també disposa d'un merònim que n'és exclusiu:

- Adjectius: *de tiro alto*/G:adj_c/N0:<pv:pa>/En:high-waisted
- Locucions: *coger el bajo*/G:v_c/N0:hum/N1:<pv:pa>/En:to shorten the hem of
- Merònims: *pernera*/G:nf/T:inc/C:mer_<pv:pa>/D:vestimenta/En:trouser leg

3.3.3.2.3. Peces de roba <medias>

Els següents predicats són exclusius de la subclasse de les mitges (<me>):

- Adjectius: *indesmallable*/G:adj_c/N0:<pv:me>/En:ladder-proof
- Noms predicatius: *carrera*/G:n/N0:hum/N1:<pv:me>/W:hacerse DET/En:to get a ladder in// *carrera*/G:n/N0:<pv:me>/W:tener DET/En:there to be a ladder in

3.3.3.3. Predicats apropiats a peces de roba que constitueixen <conjuntos>

La subclasse dels conjunts (<con>) disposa de diversos predicats que en són apropiats de forma exclusiva, amb la qual cosa ens trobem davant d'una subclasse perfectament delimitada. Alguns exemples en són:

- Adjectius: *de tres piezas*/G:adj_c/N0:<pv:con>/En:three-piece; *vestido de*/G:adj_c/N0:hum/N1:<pv:con>/En:dressed in
- Preposicions: *de*/G:prep/N0:hum/N1:<pv:con>/En:in

3.3.3.4. Predicats apropiats a peces de roba d'ús <exterior> i <interior>

L'únic predicat apropiat a les peces de roba situades en aquest nivell de la jerarquia és verbal:

- *quedarse en*/G:v/N0:hum/N1:<pv:int>/En:to stay in

Aquest verb predicatiu resulta apropiat exclusivament a la subclasse de les peces de roba d'ús interior, la qual cosa també implica la delimitació lingüística de la subclasse de les peces de roba d'ús exterior, que, si bé no compten amb un operador que en siga específicament apropiat, la falta d'adequació de *quedarse en* per a les peces de roba que se n'adscriuen permet delimitar-la per oposició a la subclasse de les peces de roba d'ús interior (*quedarse en bragas/sujetador/calzoncillos* vs. **quedarse en camisa*⁸/*pantalón/chanclas*).

3.3.4. Predicats apropiats a <complementos>

A continuació, presentem a mode d'exemple alguns dels predicats que són apropiats a cadascuna de les subclasse en les quals es divideixen els complements (<com>):

⁸ La locució adverbial *en mangas de camisa*, emprada generalment amb el verb *quedarse* (*quedarse en mangas de camisa*) i el significat de la qual és el de "vestido de medio cuerpo abajo, pero de la cintura arriba con sólo la camisa o con la camisa y el chaleco" (DRAE 2014), no l'incloem en la nostra llista de predicats, ja que no resulta apropiada a cap conjunt de peces de vestir que es puguin classificar en una subclasse determinada: com que constitueix una estructura completament fixa, no permet modificacions paradigmàtiques ni sintagmàtiques (**quedarse en mangas de camiseta/de jersey/de chaqueta/de chaleco/de abrigo*).

3.3.4.1. <cinturones>

- Verbs: *ajustarse*/G:v/N0:hum/N1:<pv:pr-ci-co>/En:to tighten up

3.3.4.2. <corbatas>

- Verbs: *anudar*/G:v/N0:hum/N1:<pv:co>/En:to tie
- Noms predicatius: *nudo*/G:n/N0:hum/N1:<pv:co>/W:hacerse DET/En:to make the knot of

3.3.4.3. <bufandas>

- Verbs: *ir envuelto en*/G:v_c/N0:hum/N1:<pv:su-bu>/En:to be tangled up in

3.3.4.4. <guantes>

- Adjectius: *desparejado*/G:adj/N0:<pv:me-con-gua-ca>/En:unpaired
- Verbs: *enguantarse*/G:v/N0:hum/N1:<pv:gua>/En:to glove in

3.3.4.5. <sombreros>

- Verbs: *ponérselo ladeado*/G:v_c/N0:hum/N1:<pv:so>/En:to put Poss <pv-so> askew; *descubrirse*/G:v/N0:hum/N1:<pv:so>/En:to take off

3.3.4.6. <protectores>

- Verbs: *protegerse con*/G:v/N0:hum/N1:<pv:su-pro>/En:to protect o.s. with

3.3.5. Predicats apropiats a <calzado>

Com a predicats apropiats a qualsevol tipus de calçat (amb i sense cordons), trobem adjectius, verbs i noms predicatius. Per exemple:

- Adjectius: *de punta cuadrada*/G:adj_c/N0:<pv:ca>/En:square-toed
- Verbs: *calzarse*/G:v/N0:hum/N1:<pv:ca>/En:to put on
- Noms predicatius: *número (Card)*/G:n/N0:hum/N1:<pv:ca>/W:probarse DET/En:-to try on the shoe size (Card) of

3.3.5.1. Calçat <con cordones>

La subclasse del calçat amb cordons (<cc>) queda delimitada, principalment, mitjançant verbs predicatius, però també hi trobem un merònim:

- Verbs: *acordonar*/G:v/N0:hum/N1:<pv:cc>/N2:hum/En:to lace up; *desacordonar*/G:v/N0:hum/N1:<pv:cc>/N2:hum/En:to unlace
- Merònim: *cordón*/G:nm/T:inc/C:mer_<pv:cc>/D:vestimenta/En:shoelace

3.3.5.2. Calçat <sin cordones>

Finalment, la subclasse del calçat sense cordons (<sc>) queda constituïda, d'una banda, per la manca d'adequació a les peces de vestir que se n'adscriuen dels predicats apropiats a la subclasse del calçat amb cordons (com ocorre amb la dicotomia de les peces de roba d'ús exterior i interior, cf. 3.3.3.4.), i, de l'altra, pels següents tres adjectius predicatius, respecte als quals el calçat sense cordons funciona com a subjecte:

- Adjectius: *con hebilla*/G:adj_c/N0:<pv:sc>/En:buckle; *de correa*/G:adj_c/N0:<pv:sc>/En:strap; *de velcro*/G:adj_c/N0:<pv:sc>/En:velcro

3.3.6. Predicats apropiats per meronímia⁹

A continuació, proposem alguns exemples de predicats apropiats a merònims. Trobem tant adjectius com verbs predicatius:

- Adjectius (tots amb N0= <pv>): *abombado*/G:adj/N0:<pv:pr>/En:puffed; *con vuelo*/G:adj_c/N0:<pv:fa>/En:flared; *plisado*/G:adj/N0:<pv:su-fa>/En:pleated
- Verbs (tots amb N1 = <pv>): *abotonarse*/G:v/N0:hum/N1:<pv:pr>/En:to but-

⁹ Recordem que la meronímia és la relació de significat d'un merònim [la part d'un tot] respecte del seu holònim [el tot respecte de les seues parts].

ton; abrocharse/G:v/N0:hum/N1:<pv:pr-ci-ca>/En:to fasten up; acordonar/G:v/N0:hum/N1:<pv:cc>/N2:hum/En:to lace up; aflojarse/G:v/N0:hum/N1:<pv:pr-ci-co-cc>/En:to loosen; arremangarse/G:v/N0:hum/N1:<pv:su-con>/En:to roll up

Observem que tots aquests predicats adjectivals i verbals apropiats a merònims resulten extensibles a la majoria dels seus holònims respectius, que “hereten” els predicats apropiats als primers: *abombado*= MÀNIGUES>CAMISA/ CAMAL>PANTALÓ; *convelo/plisado*= FALDA>VESTIT; *abotonarse/abrocharse*= BOTÓ, CREMALLERA>CAMISA/ BRAGUETA>PANTALÓ; *arremangarse*= MÀNIGUES>CAMISA; *acordonarse* o *aflojarse*= CORDONS>SABATES.

4. Conclusions

A tall de cloenda, podem afirmar que l’etiquetatge sintacticosemàntic dels diccionaris electrònics representa una tasca particularment laboriosa, atés que, en abstracte, les propietats de les unitats lèxiques i la seua combinatòria escapen qualsevol tipus de sistematicitat. Aquesta manera de descriure la llengua representa una eina per trobar aquesta sistematicitat i es torna, doncs, imprescindible per al tractament automàtic. Per aquesta raó, els sistemes de traducció no poden desvincular-se de la gramàtica i, per consegüent, no poden prescindir dels diccionaris electrònics, que aporten una gran efectivitat a l’hora de tractar la desambiguació per polisèmia. Així, si descrivim les paraules en el marc de la lexicogramàtica, és a dir, considerant la gramàtica, la semàntica i el lèxic en paral·lel, possibilitem que el traductor automàtic ja no treballi amb paraules, sinó amb usos, la qual cosa resulta en una eficàcia gens menyspreable respecte dels sistemes de traducció purament estadístics.

Per dur a terme aquesta tasca, comprovem que l’elaboració d’una classe d’objectes és un pas fonamental. L’associació dels fitxers d’arguments i predicats permet obtenir una descripció global de la classe, a més de possibilitar el reconeixement dels usos, no sols dels lliures, sinó també dels composicionals, sobretot, a través de verbs predicatius compostos. Aquest apunt ens porta a concloure que la localització i definició sintacticosemàntica dels elements constitutius d’una classe suposa una etapa decisiva per a tot tractament automàtic del llenguatge i, en concret, per a obtenir traduccions adequades a altres llengües.

Obres citades

- Blanco, X. “Dictionnaires électroniques et traduction automatique espagnol-français.” *Langages* 143 (2001): 49-70.
- . *Lexicographie bilingue français-espagnol et classes d’objets*. Barcelona: Servei de Publicacions de la Universitat Autònoma de Barcelona, 1999.
- Gross, G. “Forme d’un dictionnaire électronique.” En *L’environnement traductionnel*. Sillery-Montréal: Presses de l’Université du Québec, Aupelf-Uref, 1992. 255-271.
- . *Manual de anàlisi lingüística. Aproximació sintàctico-semàntica al lèxic*. Barcelona: UOC (Traducció de Xavier Blanco), 2012.
- Gross, M. *Méthodes en Syntaxe*. Paris: Hermann, 1975.
- Harris, Z. *Notes du cours de syntaxe*. Travaux linguistiques. Paris: Seuil, 1976.
- Le Pesant, D. & M. Mathieu-Colas. “Introduction aux classes d’objets.” *Langages* 131 (1998): 6-33.
- Man-Ghyu, P. “Traduction automatique et classes d’objets: le problème de ‘porter un vêtement’ en français et en coréen.” *Journal des traducteurs* 1 (1997): 155-167.
- Mogorrón, P. “Los diccionarios electrónicos fraseológicos, perspectivas para la lengua y la traducción.” *E.L.U.A.* 12 (2004): 381-400.