

Aditz-azpikategorizazioa hiztegintzarako lantzen

Izaskun Aldezabal Roteta
EHU/UPV (Ixa taldea)

Patxi Goenaga, Kepa Sarasola
EHU/UPV

SARRERA GISA

Hemen aurkezten dugun lana, duela hamabi bat urte Ixa taldeak hizkuntzalaritza konputazionalaren alorrean hasi zuen ibilbidearen barruan kokatu behar da.

Bestalde, garapenean dagoen saio baten berri emango dugula azpimarratu nahi dugu. Alegia, guk hemen aditzak lantzeko prozedura bat proposatzen dugu, horretara eraman gaituzten arrazoiak agerian jarritz, baina ez dugu emaitzarik ez eta ondoriorik aurkeztuko, horixe baita uneotan egiten ari garena. Honenbestez, hemen, aditzak lan.tzeko hautatutako metodologia bera eta horretarako baliatu ditugun tresna automatikoak nahiz beharrezkoa izan zaigun eskuzko lanaren nondik norakoak azaltzeak du garrantzia. Ondorengo lanetan etorriko dira, hemendik lortutako emaitzekin egin beharreko azterketak.

Hala, lehendabiziko puntuan, eta gure gaia kokatzearen, hizkuntzalaritza konputazionalaz mintzatuko gara labur. Segidan, azpikategorizazioaren azterketak izan dituen gorabeherak eta haietan topatu ditugun hutsuneak jarriko ditugu argitan. Ondoren, hutsune hauei irtenbide emateko, azken joera linguistikoek proposatzen duten bidea, sintaxia eta semantikaren arteko elkarreragina, alegia, izango dugu hizpide, bereziki B. Levin-en *English verb classes and alternations* (1993) lanaren planteamenduari erreparatuko diogularik. Lan horretan ikusi dugunetik ateratako ondorioetan oinarrituta, gure lan-prozeduraren berri emango dugu hurrengo atalean, horretarako hartu ditugun baliabide sintaktikoak, aurrerabakiak, eraikitako datu-basea, eta hautatutako aditz-zerrenda aurkeztuko ditugularik. Eginkizun horretan baliatuko ditugun baliabide automatikoen garrantzia ere azpimarratuko dugu. Bukatzeko, ondorioak eta etorkizunerako lanak aipatuko ditugu.

1 HIZKUNTZALARITZA KONPUTAZIONALAZ BI HITZ

Hizkuntzalaritza konputazionala, hizkuntza ulertu ahal izateko, ordenagailu batek beharrezko dituen baliabide linguistiko eta automatikoak uztartzean diharduen hizkuntzalaritza-mota da; alegia, diziplinartekoa. Hala, hizkuntzalaritza konputazio-

nalak, ahal duen neurrian, hizkuntzalaritza orokorretik edan eta gero hori guztia konputazionalki tratagarri izango den modu batean jartzen du. Ahal duen neurrian diogu, izan ere, hizkuntzalaritza konputazionalak *corpus errealen gainean* (hau da, gure eguneroko jardunean topatzen ditugun euskarazko testuen gainean) egiten du lan, eta corpus erreal hauek ezagutu eta analizatu ahal izateko beharrezkoa den sistematizazioaren bila jotzen duenean, behin baino gehiagotan topatzen du agor hizkuntzalaritza orokorraren iturria, askotan baliabide automatikoen eskaerei erantzuteko nahikoa suertatzen ez delarik. Puntu honetan, automatizazio-prozesua, hizkuntzalaritza orokorraren lagungarri bihurtzen da, eta fenomeno linguistiko jakin bat definitzerakoan beharrezko diren ezaugarri eta prozedurak, gizakiok gure kabuz beharbada erreparatuko ez genizkienak, argitan jartzen dizkigu. Hizkuntzalaritza orokorra, beraz, iturria da, baina automatizazioak bere garapenean laguntzen du askotan. Hortxe ageri zaigu diziplinarteko lana; azken batean, baliabide linguistikoak eta automatikoak konbinatzekoan, batak zein besteak eskaintzen dituzten ahalmen eta mugak ondo neurtu beharra. Horrela jardun dugu hizkuntzaren atal desberdinak lantzeko garaian, hala nola, ortografia, morfologia, morfosintaxia, sintaxia eta semantikaren aurrean (horiek guztiak ukitzen baititugu gure ikerlerroetan).

Ordea, helburu behinena corpora aztertzea izaki, gure lan egiteko ikuspegia eza-gutzara alderatu da, hori izanik garatu ditugun tresnen helburu nagusienetakoa. Bestela esanda, azaleko eta sakoneko kontzeptuak hartzen baditugu erreferentziatza, guk azalekotik abiaturiko analisia izan dugu aztergai. Zentzu horretan, perpausa, hitz kate jarraitu bat baino ez da, eta hala izanik, lehendabiziko eginkizuna hitzen osaera aztertzea bilakatzen da. Bata bestearen jarraian doazen hitz hauek, ordea, beti testuinguru jakinetan agertzen dira, eta hortaz, hitz bakoitzak besteekin harremanetan izan ditzakeen interpretazio guztiak aurreikusi beharra dago. Adibide batera joaz, “iritzi” hitza, izena zein aditza izan daiteke, baina horretaz gain, kontuan hartu behar da *zenbait iritzi* edo *iritzi bat* topatzen dugunean, lehenengoak mugagabearen marka (ez-ageria) daramala, eta, aldiz, bigarrenak, ez; edota *iritziak* hitza, *Alkatearen iritziak herritarrek harritu ditu* edo *Egunkariak herritarren iritziak plazaratu dituzte* gisako esaldietan ager daitekeela –hau da, ergatibo singular gisa nahiz absolutibo plural gisa. Orobat “iritzi” aditza denerako; *iritzi dio* bezalako segida batean, aditza dela esateaz gain, partizipio burutua dela ere esan behar dugu. Horrek esan nahi du, jarri ditugun “iritzi” eta “iritziak” hitzek, modu isolatuan analizatzen baditugu, hainbat interpretazio izango dituztela; bestela esanda, ambiguoak izango direla.

Anbiguitasun-mota hau oinarrikoa da hizkuntzalaritza konputazionalan. Azterketa teorikoetan, berriz, planteatu ere ez da egiten, sorkuntzatik abiaturik, ez zaiolako testuinguruari erreferentziarik egiten, eta, besterik gabe, sortzetik, delako interpretazioarekin datorrela suposatzen delako. Aztertzen den anbiguitasun-mota bakarrak, egitura sintaktikoen barruan gertatzen direnak dira, esaterako, bi izenlagunez osaturiko sintagma batean izenlagun bakoitza zein zatiri laguntzen dion zehazterakoan (adib., *gure munduaren ezagutza* [gure munduaren] [ezagutza] edo [gure] [munduaren ezagutza] izan daiteke). Hizkuntzalaritza konputazionalako zenbait sistematatan, berriz, horrelako anbiguitasun-kasuetan baten alde egin ohi da, hauen desanbiguoazioa ez delako testuinguru linguistikoetan bakarrik oinarritzen, askotan. Horra hor, hizkuntzalaritza konputazional eta teorikoaren arteko beste alde nabari bat.

Ordea, sintaxian aurrera egiteko, hitz batek izan ditzakeen interpretazio hauetatik guztietatik delako testuinguruari dagokiona hautatu behar da; alegia, deskribatutako interpretazio guztiek eragindako anbiguitasuna desanbiguitatu behar da.

Desanbiguatzea, batzuetan testuinguru hurbileko hitzetan eta hauek dakarten informazio morfosintaktikoan oinarrituz egin daiteke, baina beste batzuetan, ezinezkoa gertatzen da aditzei dagokien informazio lexikala edo azpikategorizazioari dagokiona izan gabe.

Alabaina, informazio lexikalaren beharra ukatu gabe baina hori izateari itxaron gabe, sintaxian aurrera egin da taldean, lexikoko informazio, eta sintaxi partziala deritzona landu da. Hitz bitan esateko, “sintaxi partziala” perpausaren “analisi osoa”ri kontrajarria ulertu behar da, perpausaren ulerkuntza osoa lortu gabe, zenbait mekanismo eta tekniken bidez (Abney 1991, 1997), analisi sintaktiko tradizionalaren informazio-zati erabilgarriak lortzean baitatza; hau da, analisi osoa lortzerik ez dagoenean, gutxienez, tarteko emaitzak aurkezten dira, eta analisi-aukera anbiguoak modu xeheago eta erabilgarriagoan erakusten dira. Gure taldean, horrelako mekanismoei esker, gai gara perpausako sintagmak dagokien informazio morfosintaktikoarekin identifikatzeko; eta horiek baliatu dira, hain zuzen ere, besteak beste, azpikategorizazioaren azterketan lagungarri izateko (Arriola & Soroa, 1996; Aldezabal et al. 1999; 2000; eta biltzar honexetan aurkeztu diren Gojenola & Sarasola eta Aduriz & Arriolaren lanak).

Lan hau, orduan, taldeak garatutako bide horiek erabilgarri ditugula, lexikoa sintaxiari begira lantzen hasiko garen unean kokatu behar da. Taldean garatu ditugun tresnek, beraz, batetik, guk azterkizun dugun gaiari buruzko informazioa dute beharrezko aurrerapausoak emateko, baina bestetik, baliagarri izango zaizkigu horixe bera lortzeko.

2 AZPIKATEGORIZAZIOAREN INGURUAN

Hizkuntzalaritza konputazionalen bezalaxe, hizkuntzalaritza orokorrean ere zalantzarik ez da aditzei buruzko informazio lexikala guztiz beharrezkoa dela, honek esaten baitigu esaldi bat zuzena eta ulergarria izateko aditz batek zein elementu behar dituen inguruan. Hau zehaztea baina, ez da lan erraza; eta horixe erakusten dute orain arte ildo horretan egin diren azterketa desberdinek.

Tradizionalki, azpikategorizazioa, aditzak objektua hartzearekin edo ez hartzearekin lotzen zen. Hau da, aditz batek objekturik erakusten bazuen, iragankorra zela esaten zen; objekturik ez bazuen, berriz, iragangaitza. Hori horrela izanik, aditzak iragankor eta iragangaitzetan “azpikategoriza” zitezkeela esaten zen.

Ikerketek aurrera egin ahala, ordea, agerian geratuko da sailkapen honek hutsuneak dituela aditzek erakusten duten kasuistika osoa azaltzeko garaian, aditz iragankortzat jotzen diren zenbait aditz, objekturik gabeko testuinguruetan agertzen direlako, *irakin, jardun, edota bazkaldu* (hegoaldeko euskalkietan) modukoak kasu, *urak irakin du, gogor jardun du lanean* eta *Arzaken bazkaldu du*g adibideek erakusten dutenez.

Horiek horrela, eta Gramatika Sortzailea Chomsky-ren *Government and Binding* (Chomsky, 1981) teoriari begira jarria zenean, aditzak sailkatzeko beste bide bat proposatu da: batetik, bi argumentuz osatutako aditzak egongo dira, azalean subjektu eta objektuaz gauzatzen direnak eta horrexegatik aditz iragankorrak deitua; eta bestetik, argumentu bakarreko aditzak. Hauen artean, batetik izango dira azalean subjektua eta aditz-morfologia zein markaketa iragankorra erakusten dutenak, hau da,

argumentu bakarreko iragankorrak, inergatiboak deiturikoak; eta bestetik, azalean subjektua eta aditz-morfologia zein markaketa iragangaitza erakusten dutenak, hau da, argumentu bakarreko iragangaitzak, inakusatiboak deituko dituztenak (sailkapen klasikoko iragangaitzak, alegia)¹.

Banaketa horrek, halere, ez ditu aukera guztiak asetuko, izan ere, aurreko aditz inergatiboaz gain, badira iragankortzat jotzen diren aditzak objekturik gabeko tesuinguruetan ere ager daitezkeenak; esaterako *idatzi* edota *irakurri* aditzak, *nik gaue-tan idazten dut* edota *nik ez dut inoiz ohean irakurtzen...* gisako erabileretan. Areago, badira arrunki iragangaitzat jotzen diren aditzak, objektudun erabilerak ere badituztenak; *ibili* aditzak, esaterako, *Munduko bazter guztiak ibili ditu* adibidean (Elhuyar hiztegitik hartua) ikus dezakegunez. Beraz, ohiz iragankortzat edota iragangaitzat jo izan diren aditzen “iragankortasuna” edota “iragangaitzasuna” ez da batere garbi ageri, eta horrek, ezta aditz bat sailkatzeko modua ere.

Gaur egun, onartzen da aditzak ezin direla besterik gabe iragankor edo iragangaitz gisa sailkatu, eta gauza gehiago daudela aztertu beharra. *Euskal Gramatika laburra: Perpaus bakuna* (Euskaltzaindia, 1993) liburukian, aditz-motak azaltzerakotan, hasieratik ohartarazten da aditzen azterketaren inguruan dagoen mugaz:

Aditz sintagmaren ardatz elementua aditza da eta honek hartzen ditu bere inguruan hainbat osagarri, osagarri hauek aditzarekin batera osatzen duten egituraren konplexutasuna behar bezala azterteke dagoela oraindik

Euskal Gramatika laburra: Perpaus bakuna (1993), 199-200. orr.

Hala eta guztiz ere, aipatu liburukian markatzen dira ildo batzuk aditzak aztertzeko, eta kasu gramatikaletan oinarrituriko sailkapen orokorretik abiatuta hainbat azpisail proposatzen dira:

- nor (adib. *etorri*, *berdur izan*)
- nor-nork (*erosi*, *bazkatu...*) eta bere azpimotak:
 - nor sintagmaren ordeztzesiboa hartzen dutenak (adib. *lanean dihardu*);
 - nor sintagma balio adberbialekoa denean (adib. *filmak hiru ordu iraun ditu*);
 - nor izuna denean (adib. *amari deitu diot*; *jakin-minez begiratu dizu*);
 - objekturik hartzen ez dutenean (adib. *urak irakin*; *bazkaldu*);
 - nork sintagma hartzen ez dutenean (adib. *euria hasi du*, *hotz handia egiten du*);
 - aditz esapide batzuk (adib. *merezi izan*, *balio izan*);
 - *-arazi* bidezko aditz arazleak.
- nor-nori (adib. *ekin*, *eutsi*)
- nor-nori (adib. *bururatu*, *jarraiki*)

Sailak proposatzeaz gain, hauek aurkezten dituzten sail-aldaketak definitzen dituzte, nahiz eta azken batean nor eta nor-nork saileko azpimota batzuetara mugatu:

- Inpertsionala, *nor* eta *nor-nork* sailekoentzat:

¹ Argumentu bakarreko aditz iragangaitzak izendatzeko, termino desberdinak erabili izan dituzte autoreek. Guk hemen Perlmutter-ek (1978) proposaturikoa erabili dugu, zabalduena delakoan eta egokitasunari gehiegi erreparatu gabe. Aipa dezagun, halere, Burziok (1981), “inakusatibo” terminoaren ordeztzes “ergatibo” erabiltzen duela. Gracia Solé-k (1987), bestalde, “inakusatibo” zein “ergatibo” terminoak mantentzearen alde egiten du, argumentu bakarreko aditz iragangaitz hauen artean bi motakoa topa daitezkeelako: batetik, iragankor nahiz iragangaitz modura joka dezake tenak (*sartu da / sartu dute*), eta bestetik, iragangaitz hutsak direnak (*etorri da/*etorri du*). Lehenengoen kasuan, hortaz, zilegi da “ergatibo” terminoa erabiltzea, beren erabilera trantsitiboan gai direlako subjektua ergatiboaz eta paper tematikoaz hornitzeko; bigarrenek, ordea, ez dute horretarako gaitasunik, eta beraz “inakusatibo” terminoa ongi ego kitzen zaie.

nor-nork -> inperstonala

Adib. *saltzaileak sagarrak garesti saldu ditu* -> *sagarrak garesti saltzen dira*

nor -> inperstonala

Adib. *hemendik joaten da autobusa egunero Donostiara* -> *hemendik joaten da Donostiara*

□ Arazlea, *nor* inkoatiboentzat

nor inkoatiboa -> arazlea

Adib. *aitona zaharra hil da* -> *aitona zaharra hil dute*

Aditz-motak definitzea, ordea, hiztegiaren ardura dela azpimarratzen dute gramatikariak:

Argi dago ez dagokiola gramatikari, baizik hiztegiari, aditz bakoitza zein sailetakoa den zehaztea,(...)

Euskal Gramatika laburra: Perpau bakuna (1993), 205. or.

Aditz-mota horiek bereizteko, gainera, aditzen semantikari erreferentzia egiten diote, zehazki aditz inkoatiboak azaltzen dituztenean:

Aditzen semantika kontuan hartzen badugu, mota askotako aditzak ditugu (informazio hori, hain zuzen, hiztegiak eman behar digunez, ez dugu guk hemen kontuan hartuko, aditzen forma edo itxuran eraginik ez duen bitartean). Horietako batzuk inkoatibo deitzen ditugu. Aditz inkoatiboak dira egoera aldaketa bat adierazten dutenak edo ekintzaren hasiera adierazten dutenak: *ontzia bete da* diogunean edo *aitona hil da* diogunean, argi ikusten da hor egoera aldaketa bat suertatu dela: hutsik zegoen ontzia lehenik, eta orain beteago; aitona bizirik zegoen, baina orain ez da bizi.

Euskal Gramatika laburra: Perpau bakuna (1993), 210. or.

Gramatikak, beraz, hiztegiari uzten dizkio aditz-sailak eta hauen izaera semantikoa zehaztearen ardura.

Euskarazko hiztegiatara jotzen badugu, ordea, aditz-laguntzailearen informazioa baliatzearen aldeko joera da bakarrik nabarmentzen dena.

Sarasolaren *Euskal Hiztegian* (Sarasola, 1997), esaterako, aditzen erabilera desberdinetan, laguntzailea zehazten da. Sarasolak bost laguntzaile-mota bereizten ditu: DA, DU, DA-DU, DIO eta ZAIO; non, DAK aditz iragangaitz arrunta adierazten duen, DUK aditz iragankor arrunta nahiz aipatu ditugun nor-nork azpimotak, DA-DUK erabilera iragankorra nahiz iragangaitza dutenak, DIOk objektu ageririk gabeko NOR-NORI-NORK aditzak, eta ZAIOk, azkenik, NOR-NORI aditzak. Horiekien batera, batzuetan zenbait zehaztapen morfosintaktiko ere adierazten dira. Halere, nahiz eta aditz bakoitzaren adiera desberdinei zein laguntzaile-mota dagokion espezifikatu, adiera zeren arabera bereizten den ez da esplizituki ageri.

Euskaltzaindiak argitaratu berri duen *Hiztegi Batuan* (Euskaltzaindia, 2000) ere antzeko joera ikus dezakegu.

Elhuyar hiztegian (1996) ere ildo bera jarraitzen da euskara-gaztelania atalean, aditz bakoitzeko laguntzaile-motak apur bat gehiago zehazten direlarik. Gaztelania-euskara atalean, berriz, gaztelaniazko hiztegi eta gramatiketan agertu ohi diren sailak (v.tr., v.pnrl., v.intr.) bereizten dituzte.

Gauzak horrela, hiztegiotan aditz bakoitzaren erabilera bakoitzerako laguntzailer-mota eta adiera sorta bat dugu, baina ez dugu inolako aditz-sail semantiko espliziturik topatzen.

Horrek guztiak, aukeran dauden bideak aztertzeraz eta horietako bati lehentasuna ematera eraman gaituzte.

3 SEMANTIKA ETA SINTAXIAREN ARTEKO ELKARRERAGINA

Euskaltzaindiaren gramatikan aipatu da aditzak sailkatzeko garaian semantikak duen garrantzia. Ordea, hori onartuta ere, zertan datza semantika horren izaera? Zein semantikaren arabera izan behar da sailkapen hori?

Hain zuzen ere, hizkuntzalaritzaren azken joeretan, semantika horrek aditzak erakusten duen jokaera sintaktikoarekin lotura handia duela defendatzen da. Hale & Keyser (1986, 1991) izan ziren semantika eta sintaxiaren artean lotura planteatu zuten lehenengoetariko autoreak. Eta horien jarraian etorri dira Jackendoff (1990b) eta Pustejovsky (1995) bezalako autoreen proposamenak. Autore hauek sistema konplexuak lantzen dituzte erlazio honen berri emateko; oro har, primitibo semantikoen deskonposaketan oinarritutako sistemak dira, gero primitibo horiek mekanismo sintaktiko konplexuen bidez azalerraten direlarik.

Guk, ordea, ez dugu hemen horietan sakonduko, batetik, konplexutasuna handiegia delako corpus errealean analisisa bideratzeko, eta, bestetik, guk une honetan dauzkagun baliabide automatikoak baliatzeko bideak mugatzen dituelako. Horren ordez, Levin-ek bere *English Verb Classes and Alternations* (1993) liburuan planteatzen duen antzeko bidea jorratuko dugu, autore honek planteatzen duen ikuspegiak metodologia interesgarri aurkezten duelako, eta metodologia horretan guk eskuragarri ditugun baliabideak erabiltzea posible egiten duelako. Autore honen ustez, goikoen ideia berari jarraituz, jokaera sintaktiko bera erakusten duten aditzek antzeko semantika dute, eta hortaz, jokaera sintaktiko horretatik abiatuta sail semantikoak egitea posible da. Jokaera sintaktiko hau berak “alternantzia” delakoen bidez azaltzen du:

“Verbs, as argument-taking elements, show especially complex sets of properties. As shown in B. Levin (1985b, in prep.) and other works, native speakers can make extremely subtle judgements concerning the occurrence of verbs with a range of possible combinations of arguments and adjuncts in various syntactic expressions. For instance, speakers of English know which *diathesis alternations* –alternations in the expressions of arguments, sometimes accompanied by changes of meaning- verbs may participate”.

Levin, B., *English verb classes and alternations*, 1993, 2. or.

Hala, alternantzia berak konpartitzen dituzten aditzek gutxienez beren semantikaren baitako osagairen bat batera dutela defendatzen du:

“If the distinctive behavior of verb classes with respect to the diathesis alternations arises from their meaning, any class of verbs whose members pattern together with respect to diathesis alternations should be a semantically coherent class: its members should share at least some aspect of meaning”.

Levin, B. *English verb classes and alternations* 1993, 14. or.

Horretarako, Levinek bere liburuaren lehen zatian ingelesez topatu dituen alternantzia-sorta aurkezten du (80 alternantzia), eta ondoren, bigarren zatian, hainbat sail

semantiko proposatzen ditu, teoriar, proposatutako alternantzia horietan oinarrituta (48 sail semantiko).

Ordea, alternantziak eta sailak erkatuta, ez da hain garbi ageri saio hori, izan ere, orotariko sail semantikoak ikus baititzakegu: azpisail batzuk prozesu morfologikoetan oinarriturikoak dira; esaterako, 10 saileko (*Verbs of removing*) 10.7 (*Pit verbs*) eta 10.8 (*Debone verbs*) azpisailek alternantzia berak konpartitzen dituzte, eta oinarritzko egitura beretik abiatzen dira: *The cook boned the fish* ('Sukaldariak arraina hezurgabetu zuen') eta *The cook deboned the fish*. ('Sukaldariak arraina hezurgabetu zuen'); desberdintasun bakarra, aditzen forma eratorria da. Beste zenbait sail eta azpisail, banaketa sintagmatiko hutsean oinarritzen dira. Esaterako, 29 saileko (*Verbs with Predicative Complements*) zenbait azpisail, onartzen dituzten sintagma predikatibo desberdinetan oinarritzen dira: esaterako, 29.2 (*Characterize verbs*) eta 29.3 (*Dub verbs*) azpisailak elkarren artean bereizten dira lehenengokoek as bidezko sintagma predikatiboa onartzen duelako eta bigarrenek ez. Beste batzuetan, berriz, ez dakigu zeren arabera bereizten dituen sailak; edota, hobeto esan, sailak zein alternantziak bereizten dituen; esaterako, 9.1 (*Put verbs*) eta 10.1 (*Remove verbs*) aditzek onartzen dituzten eta onartzen ez dituzten alternantziei begiratzen badiegu, bi aditz sailek alternantzia berak erakusten dituzte; ordea, 9.1 sailekoek "put an entity at some location" adierazten dute, eta 10.1 sailekoek, berriz, "remove an entity from a location". Eta hori, azken batean, aditzok hartzen dituzten sintagmek hautatzen duten preposizio desberdinek adierazten digute.

Azken batean, badirudi sail nagusiak egiteko garaian gehiago erreparatzen diola aditz bakoitzak inguruan onartzen dituen osagaiak zehazteari, aurkeztutako alternantziak konpartitzeari –edo ez konpartitzeari– baino. Ondorengoari, alegia: "the occurrence of verbs with a range of possible combinations of arguments and adjuncts in various syntactic expressions".

Autoreak berak onartzen du bere lanaren baliagarritasunaren muina:

"In this book, I have chosen a level of classification characterized by clustering of verbs that should further the isolation of meaning components".

Levin, B. *English verb classes and alternations* (1993) 18. or.

"Finally, this book does not assess the implications of the material it includes for the identification of meaning components, nor does it move beyond their identification to the formulation of a lexical semantic representation. Rather, it is intended to set the stage for these necessary next steps".

Levin, B. *English verb classes and alternations* (1993), 19. or.

Alegia, egiten diren multzoak gorabehera, gerora aztertu behar da multzo horietan zein osagai semantikok duen esangura. Azken batean, berak teoria bat planteatzen du baina gero ez dirudi azkeneraino iristen denik. Autoreak eginkizun hau etorkizunerako uzten du, hizkuntza gehiagotan egiten diren azterketen zain. Bere lanean, orduan, dirudienez garrantzia handiagoa eman dio oinarritzko egitura eta horiei loturiko egitura sintaktikoak aurkezteari. Honenbestez, Levinen gonbitea jarraiki, euskaraz ere saio bat egitea erabaki dugu.

Azpimarratu behar da, halere, Levinen lanean ez dela esplizituki ageri oinarritzko egitura horiek zehazteko zehatz-mehatz zertan oinarritzen den. Besterik gabe, oinarritzko egitura hauek zedarritutzat ematen ditu, eta sail bakoitza IS eta PS jakinez

definitzen du. Ikusi dugun bezala, gainera, proposatzen dituen klaseak ez daude beti irizpide berari jarraituz osatuak, eta ez dira hortaz guztiz oinarri fidagarriak. Horren aurrean, gure hautua izan da, ez oinarritzko egiturarik jakintzat ematea eta euskaran bertan topatzen ditugun zenbait baliabide sintaktiko baliatuz, 110 aditzen gainean azterlan bat egitea. Zehaztu dezagun hurrengo puntuan, bada, zein baliabide hautatu ditugun eta nola baliatuko ditugun gure lanean.

4 GURE AUKERA: EUSKARAREN BALIABIDE MORFOLOGIKO-SINTAKTIKO AGERIEN ERABILERA

Aipatu baliabide sintaktikoak hautatzeko, oso funtsezkoa izan da gaur egun gure baliabide automatikoen eskain diezaguketena kontuan hartzea. Sarreran esan dugun bezala, egun eskura ditugun tresna automatikoen, corpusean ageri diren esaldietako sintagma analizatzen dituzte, hauei dagokien informazio morfosintaktikoa –alegia, numeroa mugatasuna eta kasua– eskainiz. Bestetik, esaldi bakoitzean erabiltzen den laguntzaile-mota errazki lor dezakete. Beraz, hor baditugu hainbat ezaugarri gure erabilpenerako eskuragarri.

Azterketa linguistikoari begira, berriz, jakina da, aditzen osagaiak sintaktikoki gauzaterakoan agertzen duten sintagma-mota (azpikategorizazio-informazioa, alegia) zehaztea funtsezkoa dela lexikoa eta sintaxiaren arteko loturaren berri emateko. Zehaztu dezagun, ordea, euskaraz, sintagma-motari baino kasuari erreparatzen zaiola gehiago, besteak beste, sintagma-motek, ingelesez ez bezala, ez dutelako posizio desberdinetan funtzio desberdinen berri ematen. Esan beharrik ez da, laguntzaile-mota zehaztea ere garrantzitsua dela oso, hauek aditzaren oinarritzko osagaien berri ematen ditugulako euren presentzia beharrezkoa izan gabe esaldian.

Horiek horrela, gure azterketan, aipatutako bi ezaugarriez, alegia, laguntzaile-mota eta kasuez, baliatu gara. Halere, horiek kontuan hartzerakoan, zenbait erabaki hartu behar izan ditugu, aurkezten dituzten hainbat konplexutasun kontuan hartuta. Horiek izango ditugu aztergai 4.1 puntuan. Behin baliabideak hautatu ditugula, aditzen erabilera bereizteko, bi ezaugarri hauen aldi bereko konbinazioa erabiliko dugu. Hau da, aditz batek aditz laguntzaile-mota jakin batekin aldi berean onar ditzakeen kasuak markatuko ditugu, eta aldi bereko laguntzaile eta kasu-konbinazio desberdin bakoitzeko erabilera desberdin bat markatuko dugu. Saio honetan ere zenbait erabaki hartu ditugu, lana errazte aldera. Horiei buruzkoak 4.2 puntuan ikusiko ditugu. Azkenik, 4.3, 4.4 eta 4.5 puntuetan datu hauek guztiak gordetzeko datu-basea, hautatutako aditzak eta horretarako baliabide automatikoen eskain dezaketen laguntza ikusiko dugu, hurrenez hurren.

4.1 Aditz laguntzaile motak eta kasuak

4.1.1 Aditz laguntzaile-motak

Aditz-formak erabilera desberdin guztietan hautatzen duen aditz laguntzaile-mota baliatzeko garaian, ohizko lau motak izango ditugu kontuan; adierazpide eza-guna erabiliz: DA, DU, DIO eta ZAIO aditz laguntzaileak. Ez ditugu, halere, agerpen guztietan ageri zaizkigun erabilera guztiak markatuko. Azterketa errazteko eta lexikoko balioetara bakarrik mugatze aldera, aurrerabaki batzuk hartu ditugu.

Inkoatiboa, inpersonala, pasiboa, antipasiboa, erreflexiboa eta erreziprokoa

Gauza jakina da, laguntzaileak euskaraz, batez ere, DA motakoak, orotariko informazioa ematen duela. Ikus ditzagun segidan.

- Alde batetik, hain ezagun den *balio inkoatiboak* ageri duen laguntzaile mota dugu:
(*Haurrak lorontzia hautsi du*) *lorontzia hautsi da*
- Bestetik, *balio inpersonala* emateko baliatzen da:
(*Langileek lan handia egin dute*) *lan handia egin da*
- Bestetik, *erreflexibotasunaren adierazle* izan daiteke:
(*Amak haurra jantzi du*) *ama jantzi da; haurra jantzi da*
- *Erreziprokotasuna* ere adieraz dezake, beti ere pluralean ageri delarik:
(*Pellok Txomin topatu du*) *Pello eta Txomin topatu dira*

Eta azkenik, euskarazko pasiboetan, antipasiboetan, eta hizkuntza erromaniko-etako egituren itxuraz egin ohi diren pasiboetan ere erabiltzen den laguntzaile-mota da:

- Pasiboa:
(*106*)² *Anttonek etxe hori egin du* (*107*) *etxe hori, Anttonek egina da*
- Antipasiboa:
(*110*) *nik gauza asko ikusi dut* (*111*) *ni, gauza asko ikusia naiz*
- “Pasibo erromantzea”:
Deabruak Jainkoa tentatu du (*103*) *Jainkoa deabruaz tentatua izan da*

Erabilera hauetan guztietan, ordea, laguntzailearen itxura iragangaitzaz gain, topatzen diren fenomeno linguistikoak ez dira ez hain agerikoak ez eta hain errazak azaltzen ere.

Esaterako, inkoatiboan nahiz inpersonalean, aditz-laguntzailea iragangaitz bihurtzen da, iragankortasuna adierazten duen elementua –kasu ergatiboduna– desagertu egiten delako. Pasibo eta antipasiboetan, berriz, iragankortasuna adierazten duen elementua ez da desagertzen baizik eta beste itxura sintaktikoz ageri da. Erreflexibo eta erreziprokoetan, azkenik, zaila da iragangaiztasuna iragankortasuna adierazten duen elementuaren desagertzearen ondorio gisa ulertzea. Erreflexibotasuna, izatez, ekintza bat, norberak eginez, norberari zuzentzean datza; erreziprokotasuna, berriz, ekintza bera batak besteari egitean. Orduan, lehenengo kasuan, ez du dirudi inongo desagertzerik dagoenik, baizik eta bi funtzioak elementu beraren baitan gertzen direla esan daiteke. Bigarrenaren kasuan, berriz, bi elementuren presentzia derri-gorrezkoa da, eta hori pluraltasunaren bidez adierazten da; hau da, adizkia plurala da, erreziprokotasunean parte hartzen duten entitateak sintagma berean daudelako nola-baiteko pluraltasuna adierazten (numero plurala edota kopulazioa, esaterako). Hortaz, hauetan ere bi funtzioak bi elementu horien bidez adierazten direla dirudi.

Fenomeno hauek guztiak, orduan, iragankortasunaren kontzeptua are gehiago nahasten dute, morfologikoki “iragangaitzak” izan arren, ekintzan egiletasuna adierazten duen elementua, beti ez desagertzeaz gainera, egiletasuna ez delako beti modu berean adierazten. Guk, honen aurrean, zenbait erabaki hartu behar izan ditugu.

Pasibo, antipasibo eta inpersonalei dagokienez, aspaldi onartua da hirurak, lexikoko egitura jakinetatik abiatuta, operazio lexikalen bidez sintaxian gauzatzen

2 Aurretik zenbaki bat daramaten adibideak Euskaltzaindiaren *Euskal gramatika laburra*. *Perpaus bakuna* (1993) liburukitik hartu ditugu.

diren egiturak direla. Guk ere hauxe bera onartuko dugu, eta aditzen agerpenetan horrelako erabilerak topatzen ditugunean ez ditugu aintzat hartuko.

Besterik dugu inkoatiboaren kasuan. Hau, egun, eztabaidagai da oraindik auto-reen artean; batzuen iritiz, goiko kasuak bezala, operazio lexikal batez –alegia, ekintzaren egilea edo kausa oinarrizko egituratik syntaxirako bidean txertatuz, edota alderantziz, ezabatuz, abiapuntutzat hartzen den egituraren arabera– lortzen da (Hale and Keyser, 1991). Beste batzuek, aldiz, inkoatibotasuna lexikoan bertan dagoen baliotzat hartzen dute, hau da, lexikoko egituran zehaztu beharreko egitura gisa (Levin & Rappaport, 1992b). Gu, Levin eta Rappaport-i jarraituz, inkoatibotasuna lexikoko balio gisa ikusiko dugu, horrek ez baitigu behartzen egitura bata edo bestea hautatzera oinarrizko egitura gisa. Izan ere, badira aditzak berezko inkoatibotzat jotzen direnak, esaterako, *erori*, ez dutenak pareko kausatiborik (*erori du; *jaio du); eta baita berezko kausatibotzat jotzen direnak, *eragin*, esaterako, ez dutenak pareko inkoatiborik (*eragin da). Operazio lexikalen bidea hartuz gero (hau da, oinarrizko egitura bat hautatzera joko bagenu), hautatuko genukeen oinarrizko egituraren arabera, bi kasuotan aditzaren forma bera ere aldatu beharko genuke. Oinarrizkotzat kausatiboa hartuz gero, *erori*, agian *bota*-tik sortu beharko genuke, eta inkoatiboa hartuz gero, *eragin*-en inkoatiboa lortzeko *gertatu*?... beharko genuke. Jakin badakigu, aukera honen auzian, aldeko eta kontrako autoreak badirela (besteak beste, Halliday (1968), Fodor (1970), Cruse (1972, 1973)), baina gure jomugarik lehena aditz-forma beraren balioak zehaztea denez, albo batera utzi ditugu horrelako aukerak.

Erreflexibo eta erreziprokoei dagokienez, berriz, gauzak ez ditugu hain garbi ikusten. Etxeparek, *A grammar of Basque* (2001) liburuan hauen inguruko azterketa interesgarria aurkezten du. Bertan irakur dezakegunez, erreflexibotasuna adierazteko euskarak bi baliabide erakusten ditu: bata egitura iragankor arrunt batean *-ren burua* izenordaina erabiliaz, eta bestea hemen aipatu dugun adizki-mota iragankorraren bidez. Baina aditz guztiek ez dute joera bera erakusten bi baliabideekiko: aditz batzuek *-ren burua*-z bakarrik egin dezakete erreflexiboa; beste batzuek aditza iragangaitz bihurtzen soilik, eta badira biak onartzen dituztenak:

“The alternation is not entirely free, however. Many predicates can only take the reflexive argument, and cannot reflexivize via detransivitation, whereas other predicates can only form reflexive by detransitivizing. Still, a third class of predicates can be reflexivized in either of two ways”.

R. Etxepare, (2001) “Reflexives and Reciprocals” Hualde J.I. & Ortiz de Urbina, J. (eds.), *A grammar of Basque*, 293. or.

Beraz, ez dirudi erreflexibotasuna fenomeno guztiz erregularra denik, eta bade-nik oinarrizko egitura bat nondik bestea eratortzen den.

Gainera, erreflexibotasuna iragangaitzez bakarrik egin dezaketen aditzen artean, asko dira erabilera inkoatiboa ere erakusten dutenak:

“So-called inherent reflexives only admit the intransitive reflexive form: *gorde* ‘hide’, *jantzi* ‘dress’, *apaindu* ‘make up’, *garbitu* ‘wash’, *zikindu* ‘soil’, *idortu*, *lehortu* ‘dry’. Intransitive forms here not always have a reflexive meaning, they can also have an inchoative meaning: a. Jantziak lehortu dira; b. Leihoak zikindu dira; c. Eguzkia gorde da; d. Zelaia lorez jantzi da”.

R. Etxepare, (2001) “Reflexives and Reciprocals” Hualde J.I. & Ortiz de Urbina, J. (eds.), *A grammar of Basque* (2001), 295. or.

Horrek esan nahi du, kasuotan laguntzaile iragangaitzak balio anitz izan ditza-keela eta gure ikuspegitik aztertuta, anbiguotasun handia sorraziko lukeela. Hortaz, momentuz ez ditugu inongo modu bereziz markatuko erreflexibotasunari dagozkio- nak: laguntzaile iragangaitzez eginiko erreflexiboak DA laguntzaile-motakoak izan- go dira –inkoatiboak bezalaxe– eta *-ren burua* izenordainez osaturikoak, DU lagun- tzaile motakoak.

Erreziprokotasuna adierazteko ere bi baliabide dauzkagu euskaraz. Batetik, *elkar* izenordaina egitura iragankor arrunt batean erabiliz; eta bestea adizki iragan- gaitzez. Halere, bi jokamolde hauekiko joera ez da aditzek erreflexibotasuna adieraz- teko erabiltzen dutenaren parekoa. Alegia, erreflexiboa adizki iragangaitzez egin dezaketen aditz orok ez du erreziprokotasuna adizki iragangaitzez egiten. Esaterako, erreflexiboetan ez bezala, erreziprokotasuna adierazteko adizki iragangaitza baliatzen duten aditzek ez dute inoiz balio inkoatiborik:

“Reciprocal reading also arise either though a specific reciprocal argument *elkar* ‘each other’ or by detransitivizing the verb. The distribution of the two strategies however, is somewhat different when compared to reflexivization. Detransitivi- zing is not available with verbs whose intransitive forms are inchoative (see above). It is possible with some of the stative verbs: *maite izan* ‘love’, *ezagutu* ‘know’, *miretsi* ‘admire’; but not others: *atsegin izan* ‘like’, or *hobetsi* ‘prefer’. Intransitive reciprocals are also available with verbs expressing representation, such as *irudikatu* ‘picture’, *errepresentatu* ‘represent’ or *margotu* ‘paint’. (...)”

Hualde, J.I & Ortiz de Urbina, J. (2001), *A grammar of Basque*, 296. or.

Beraz, alde batetik, erreziprokotasuna ez dirudi ere guztiz erregularra denik, eta bestalde, ez dator erreflexibotasuna erakusten duten aditzekin bat.

Horren aurrean, gure erabakia izan da, erreflexibotasunarekin bezala, errezipro- kotasuna ez inola ere markatzea; honenbestez, laguntzaile iragangaitzez eginiko erre- ziprokoak DA laguntzaile-motakoak izango dira eta *elkar* izenordainez osaturikoak, DU laguntzaile motakoak. Ohar gaitezen, jadanik aipatua dugun bezala, erabilera erreziprokoetan adizki plurala erabiliko dela, eta hor bagenukeela nolabaiteko marka desberdintzailea; halere, aukera hori bazter batera utziko dugu eta hauek ere, besterik gabe, DA laguntzaile-motakoen artean kokatuko ditugu.

Bukatzeko, eta ZAIO motako aditz laguntzailera iraganez, DAren antzera, aditz iragankor datibodunen inpersonalaren adierazle dela aipatu behar dugu (*Behartsuei arropa eman zaie, Kazetariei egia esan zaie, liburua kendu zaio...*). Jakina, honela- koak ere ez ditugu kontuak hartuko.

4.1.2 Kasuak

Inguruan aditz batek onartzen dituen kasuak dira laguntzaile-motarekin batera erabiliko dugun beste baliabidea.

Argitu dezagun, halere, guk kasutzat deklinabide-tauletako atzizki oro jotzen ditugula; hau da, ez dugu bereizketarik egingo kasu gramatikal (absolutibo, ergatibo eta datibo) eta bestelakoen artean; guretzat denak izango dira kasuak: kasu ergatiboa, kasu absolutiboa, kasu adlatiboa, kasu ablatiboa... Horietaz gain, kasutzat hartuko ditugu ere mendeko perpausak eratzen dituzten erlazio-atzizkiak (*-ela*, esaterako).

Ordea, kasuen artean hautaketa bat egin dugu, batetik azterketaren konplexutasuna arintzeko, eta bestetik, azterketa xeheegia ez gertatzeko. Hautaketa egiteko, berriz ere gure baliabide automatikoetara jo dugu, eta taldean eskura dugun *Euskaldunon Egunkariako* corpusetik³, maiztasun handienarekin agertu diren lehenengo 9 kasuak hartu ditugu, hortik aurrerako maiztasunak txiki samarrak direlako. Hona maiztasunari dagozkion emaitzak eta guk hautatutako 9 kasuak marra batez bereizita:

• absolutiboa	% 41,79
• ergatiboa	% 36,37
• inesiboa	% 6,38
• datiboa	% 3,61
• - <i>ela</i>	% 2,70
• instrumentala	% 1,77
• soziatiboa	% 1,52
• ablatiboa	% 1,34

• adlatiboa	% 1,28
• - <i>teko</i>	% 0,52
• motibatiboa	% 0,18
• - <i>lako</i>	% 0,13
• destinatiboa	% 0,06
• prolatiboa	% 0,05
• adl. bukatuzkoa	% 0,04
• adl. norabide zuzenezkoa	% 0,0084

...

4.2 *Aldi berean ager daitezkeen aditz laguntzaile-moten eta kasuen konbinazioa markatzen*

Aditz baten balioak definitzeko, esan dugun bezala, hautatutako bi baliabide sintaktikoen aldi bereko konbinazio-aukerak zeintzuk diren aztertuko dugu, eta horretarako markaketa-sistema bat erabiliko dugu. Sistema horretan ‘X’ batez markatuko dira aldi berean ager daitezkeen aditz laguntzaile eta kasuak. Laguntzaileen kasuan, ‘X’ agertzeak esan nahiko du aditz horrek laguntzaile hori hartzen duela delako erabilera horretan. Berriz, kasu bati ‘X’ marka jartzen diogunean, oro har, balio bat izango du: “ager daiteke”. Hau da, corpusean agertu ala ez, gure ustez kasu horrek erabilera jakin horretan agertzeko gaitasuna duela adierazten dugu. Kasuak markatzeko, ‘X’z gain, ‘-’ balioa ere erabiliko dugu, delako kasu hori konbinazio horretan ezinezkoa dela adierazteko.

Alabaina, horiek izan arren besterik ezeko balioak, markaketan ere aurrerabaki batzuk hartu behar izan ditugu, eta zenbaitetan ‘X’ren balioa gehiago zehaztu. Ikus ditzagun segidan.

³ Taldean eskura ditugu 1999ko urtarriletik 2000ko maiatza bitarteko *Euskaldunon Egunkariako* zenbaki guztiak euskarri elektronikoa

4.2.1 Datiboa

Jakina denez, kasu datiboak balio asko baina oraindik erabat zehaztugabeak ditu.

Datiboa aukerazko den kasu askotan, “afektibotasun” gisa ezagutzen den balio berezia adierazten da; horrexegatik, “datibo afektibo/etiko” izendatu izan da. Afektibotasun hau, normalean posesio-erlazio batean datzala esan ohi da: *semea Ameriketatik etorri zaio* (bere semea), *alaba bahitu diote* (bere alaba), *liburua kendu diote* (bere liburua) *ama erotu zaio* (bere ama), *irakaslea haserretu zaigu* (gure irakaslea). Eta posesio-erlazio hau askotan parte-oso erlazioa izaten da: *besoa apurtu zaio* (gorputz osotik besoa); *mahaiari hanka kendu diote* (mahairi zati bat)...

Beste batzuetan, hautazkoa izan arren, afektibotasuna ez beste zerbaitek adierazten du: *Jokalariak egia esan du / Jokalariak egia esan die kazetariei; Margolariak koadro bat pintatu du / Margolariak koadro bat pintatu dio emakumeari; Irakasleak liburua ekarri du / Irakasleak liburua ekarri digu; Behartsuek arropa eskatu dute / Behartsuek arropa eskatu digute...*

Bi balio hauetan, datiboa, besterik gabe, egitura batetik bestera erantsi egiten dela dirudi, DA motakoa dena ZAIO bilakatuz, eta DU motakoa dena DIO. Alegia, DA erabilerako kasu-konbinazio berean datiboa txertatzen da eta oro bat DU motako konbinazio berean datiboa txertatzen da.

Baina, badira beste zenbait erabilera ez dutenak modu horretako eransketarik erakusten.

Batetik badira eransketa hutsa erakusten ez dutenak (nahiz eta hori ere erakuts dezaketen); hau da, kasu konbinazio berean bat etor ezin daitezkeenak. Adib.

Pello adiskideez ahaztu da

-> *Pellori *adiskideez/adiskideak ahaztu *zaio/zaizkio*

Pellok motxila ahaztu du

-> **Pellok Anderri motxila ahaztu dio*

Krisi ekonomikoak kalte handiak eragin ditu

-> *Krisi ekonomikoak kalte handiak eragin dizkio herriari (arrunta)*

-> *Pellok (*bizikleta) pedaleei eragin die*

Pello etxera heldu da

-> *Pello etxera heldu zaigu (arrunta)*

-> *Pellok (*etxera) lanari heldu dio*

Pellok oihala gerrian lotu du

-> *Pellok haurrari oihala gerrian lotu dio (arrunta)*

-> *Pello (*gerrian) lanari lotu zaio*

Bestetik, badira datiboa baino ez dutenak onartzen. Adib.: *jario, otu, ekin...*

Ur biziko ibaiak jarioko zaizkio barnetik (Sarasolaren *Euskal Hiztegi*ko adibidea)

Otutzen zaionean egingo du (Sarasolaren *Euskal Hiztegi*ko adibidea)

Edota, DU ere onartu arren, ia beti DIO motaz agertzen diren aditzak; hau da, ia goikoen pareko, datiboa berezkoa dutenak:

Lehen saria eman zioten (Sarasolaren Euskal Hiztegiko adibidea)

Balio hauen guztien aurrean, gure hautua izan da datiboa markatzea bakarrik, batetik, eransketa hutsa ez denean, bestetik, datiboa onartzen denean bakarrik, eta azkenik, ia beti datiboa erabiltzen denean; azken hauetarako corpusari begiratuko diogularik bereziki, hau da, corpusean bakarrik datiboaz agertzen bada, markatu egingo dugu; aldiz, ez bada hala, ez dugu markatuko.

Horrek esan nahi du, DIO eta ZAIO laguntzaile-motatan markatuta ageri ez diren datiboak, aukerazkoak edo ezinezkoak izango direla. Hala, datiboa aukerazkoa denean, datiboaren laukitxoa bete gabe utziko dugu, eta, aldiz, ezinezkoa denean, ‘-’ markaz adieraziko dugu. DIO eta ZAIO laguntzaile-mota markatuta ageri denean, berriz, datiboak beti ‘X’ marka izango du.

4.2.2 Absolutiboa

Absolutiboari ‘X’ marka jartzen diogunean, automatikoki ulertu behar da delako aditzak kasu hori beharrezkoa duela, baina askotan ez dela agertzen elipsi-prozedurak direla tarteko. Hau da, ‘X’ horrek markatzen duen “ager daiteke” zentzua, kasu hauetan oso jakina da: agertzen ez denean, elipsiaren ondorioztat ulertu behar da (pro-drop fenomenotzat). Aldiz, elipsia ez den bestelako fenomenoak direnean (Euskaltzaindiak nor-nork sailean zerrendatzen dituen zenbait azpimotetan topatzen ditugunak, hala nola, absolutiboa barik inesiboa agertzen dutenak, nor izuna dutenak, nor agertzen ez dutenak eta zenbait aditz esapideen berezitasunak) ‘-’ ikurrarekin markatuko dugu esplizituki, ezin ager daitezkeela markatzeko.

Ondorioz, DA eta ZAIO aditz laguntzailea markatzen ditugun unetik, absolutiboa beti agertuko da ‘X’ markarekin, kasu horietan agertzen ez denean beti pro-drop fenomenoaren ondorioz izango baita. DU eta DIO aditz laguntzailea markatzen ditugunean, ordea, absolutibo markatzeko bi aukera izango ditugu: ‘X’ edo ‘-’, non lehenak esan nahi duen “beti ageri da, pro-drop fenomenoak barne”, eta bigarrenak “balio jakin horretan ezin da agertu”.

4.2.3 Ergatiboa

Ergatiboari dagokionez, eta nor-nork azpimoten artean nork sintagma agertzen ez dutenak gogoan, DU aditz-laguntzailea markatzeak ez du sistematikoki ergatibo kasua markatzea ekarriko. Hemen, absolutiboarekin bezalaxe, bi aukera izango ditugu: ‘X’ edo ‘-’, non lehenak esan nahi duen “beti ageri da, pro-drop fenomenoak barne”, eta bigarrenak “balio jakin horretan ezin da agertu”. DIO laguntzailea markatzen denean, berriz, ergatiboa sistematikoki ‘X’ batez markatuko dugu, mota horretakoetan beti ageri delako.

4.2.4 Kasuek denborazko erreferentzia dutenean: -n, -z, -tik, -ra, -kin...

Jakina da ere, aditzek gehien-gehienetan denborari erreferentzia egiten dioten osagai sintagmatikoak onar ditzaketela, eta zentzu hori hainbat kasuk eman dezaketela aditzera, hala nola, inesiboak, ablatiboak, adlatiboak, instrumentalak eta baita soziatiiboak ere (*gauean, igandetik, igandera, arratsalde, igandearekin...*). Jakin badakigu denborazko erreferentzia hauek normalean ez dutela aditzari dagokion

informazio berezirik ematen, ekintza denboran kokatzeaz gain, eta aditz guztiek, oro har, horrelakoak onartzeko gaitasuna dutela.

Hori dela eta, gure erabakia izan da, denborazko erreferentzia duen kasurik ez kontuan hartzea aldi bereko kasuak markatzeko garaian. Alegia, agerpen batean inesibo, ablatibo, adlatibo edo instrumentalak denborazko zentzua badu, ez dugu kontuan hartuko.

4.2.5 Kasurik gabeko sintagmak

Badira sintagma batzuk ez direnak kasu bidez osatzen; adberbio sintagmak alegia. Adberbio sintagmak beti lotu izan dira –goian aipatu ditugun denbora-baliozko sintagmak bezalaxe– aditzaren izaerari ezer gutxi aportatzen dioten elementu gisa. Jakina da, halere, adberbioak mota desberdinetakoak izaki, esaldian funtzio desberdina izan ditzaketela, eta ez direla beti hain esangura gutxikoak aditzari begira. Har dezagun, esaterako, *jokatu* aditza ondoko esaldian: *Pellok [ondo/oker...] jokatu zuen aldi hartan*. Esaldi honetan *ondo/oker* kenduko bagenu, *Pellok jokatu zuen aldi hartan*, guztiz bestelako irakurketa egingo genuke, seguruenez, elipsi bat ulerturik ‘zer-baitetan jokatu zuela’ pentsatuko genukeelarik; adberbioidunaren balio dezentez desberdina, hortaz. Guk, ordea, horrelakoak albo batera utzi behar ditugu, ez dutelako agerian kasurik erakusten. Horrek esan nahi du, aditz batzuen zilegizko balio batzuk kanpoan geratuko direla.

Adjektiboekin ere batzuetan antzekoa gertatzen dela esan dezakegu, esaterako *jarri*-ren ondoko esaldia hartzen badugu: *Pello bileran gorri jarri zen*. *Gorri* kenduko bagenu, ez genuke, seguruenez, esaldia ulertuko (testuingururik gabe). Adberbioen kasuan ez bezala, ordea, adjektibo hauek kasu absolutiboz osatuak direla esan ohi da (mugagabea askotan), izenki predikazioa osatzen dutelarik; hala izanik, kasudun sintagmak genituzke. Zabalak bere tesian (*Predikazioaren teoriak Gramatika Sortzailean (Euskararen kasua)*, 1993) sakon aztertzen ditu predikazio harreman oro, eta hainbat proposamen egiten ditu erlazio hori gauzatzen duten elementuei dagokienez. Hemen aipatu dugun adibidean (alegia, adjektiboaren kasuan), eta baita izenenen kasuan ere, erlazio hori zenbaki komunztaduraren bidez edota komunztadura ageririk gabeko \emptyset atzizkiak osatzen dela esaten digu. Guk, ordea, komunztadura hauek kasura mugatuko ditugu, numero ageria (singularra eta plurala) nahiz ageria ez dena (guk mugagabe gisa ulertuko duguna), absolutibotzat hartuz. Hala, gure sisteman bi absolutibo aldi barean agertzea onartuko dugu horrelako erlazioen berri ematen duten neurrian. Bestelako kasurik aldi berean errepikatzea, ordea, ez dugu momentuz kontuan hartuko. Hala, horrelako kasuetan bi ‘X’ jarriko dizkiogu absolutiboari.

4.3 Aditz bakoitzeko balioak gordetzeko datu-basearen diseinua

Datu hauek guztiak modu koherente batean gordetzeko, datu-base bat antolatuko dugu (Aldezabal et al., 1998). Hala, aditz bakoitzaren balioak lantzeko bi pausu eman beharko ditugu.

Lehen go pausuan, aditz bakoitzak lau laguntzaile-motei dagozkion aukerak izango ditu, eta bertan aditzari dagokiona markatuko dugu 'X' batez; *joan* aditza erabiliz adibide gisa:

ADITZA: joan 1	
LAGUNTZAILE-MOTA	X/-
DA	X
DU	
ZAIO	
DIO	

Bigarren pausuan, markatutako aditz laguntzaileko taulara pasatuko gara. Laguntzaile-mota bakoitzeko kasu-aukera berak izango ditugu. Hala, adibidearekin jarraituz, DA laguntzailea izan bada markatu duguna, honelakoxe taula agertuko zaigu:

ADITZA: joan 1	
LAGUNTZAILE-MOTA DA	
KASUAK	X / - / HUTSIK
ABSOLUTIBOA	X
ERGATIBOA	-
DATIBOA	
ADLATIBOA	X
ABLATIBOA	X
SOZIATIBOA	X
INESIBOA	X
INSTRUMENTALA	X
-ELA	-
Oharrak: Markatutako kasu guztiak batera agertu badaitezke ere, ez dugu adibiderik topatu ez ablatiboarekin, ez soziatiboarekin ez instrumentalarekin. Guztietan adlatiboa da azaldu zaiguna.	
Adibideak: @ 185869,2634,0028,05: Bere langelara <i>joan</i> eta begiratu dezakegu. @ 185916,2634,0066,13: Hoteletik deitu dugutenean berehala <i>joan</i> naiz hara eta gela egoera penagarrian ikusi dut.. @ 185999,2634,0142,17: Orduan, kotxera <i>joan</i> naiz ibilgailuarekin inguruak azkarrago aztertzeko.	

Bertan, delako laguntzaile-mota bakoitzarekin aldi berean topa daitezkeen kasuak markatuko ditugu arestiko puntuetan adierazitako gida-lerroei jarraituz. Hala, aldi bereko konbinazio bakoitzeko fitxa berri bat irekiko dugu; bestela esanda, 'laguntzaile + kasu-konbinazio' desberdin bakoitzeko, sarrera bat izango dugu, eta horretarako homografo desberdinak erabiliko ditugu: 1, 2, 3... Hauetako taula bakoitzak, iturri bezala erabili ditugun adibideentzako lekua izango du, bai eta interesgarri

iruditzen zaigun edozein oharrentzako ere. Ikus dezagun, esaterako *joan* aditzaren beste balioetat jo dugun fitxaren adibidea:

ADITZA: joan 2	
LAGUNTZAILE-MOTA	X/-
DA	X
DU	
ZAIO	
DIO	

ADITZA: joan 2	
LAGUNTZAILE-MOTA DA	
KASUAK	X / - / HUTSIK
ABSOLUTIBOA	X
ERGATIBOA	-
DATIBOA	-
ADLATIBOA	-
ABLATIBOA	X
SOZIATIBOA	-
INESIBOA	-
INSTRUMENTALA	-
-ELA	-
Oharrak: Erabilera honetan ez dirudi helbururik defini daitekeenik; horregatik jarri dugu beste erabilera gisa.	
Adibideak: @205731,3803,0002,02: Hodiak etxeen fatxadetatik <i>joan</i> beharko dira.	

Bi markaketa hauetatik, hortaz, bi sarrera izango genituzke: bata 'DA + abs abl' markatua eta bestea 'DA + abs abl ala soz ins' gisa.

Hautatu dugun datubase-sistema, Macintosh ordenagailuetan baliatzen den FileMiker delakoa da, honek modu erraz eta txukun batean uzten baitugu bi taulen arteko lotura egiten, eremu bakoitzeko balioak definitzen, eta eremu hauei dagozkien balioak errazki aukeratzen.

4.4 Baliabide automatikoen erabilera

Esan dugun bezala, markaketarako, gure intuizioak erabiliko ditugu, eta hala izaki, subjektibotasunari lotuta egongo dira ezinbestean. Horren alloan, corpusak objektibotasuna jartzen digu agerian; alegia, guk ontzat eman ditugun esaldietan agerian azaldu dena bederen, badakigu zilegi dela. Hori, hortaz, oso kontuan hartzeko izango da gure eskuzko lanarekiko konparagarri. Horrela, guk zedarritutako erabilera bakoitzean aurreikusi ditugun laguntzaile eta kasu-konbinazioak egiaztatzeko eta

osatzeko, gure analizatzaile automatikoak aplikatu osteko emaitzak baliatuko ditugu (Aldezabal et al. 2001).

Horretarako, esaldien iturri gisa erabili ez dugun *Euskaldunon Egunkaria*-ko corpusa erabiliko dugu. Dena den, aipatu behar dugu, datuek esangura izateko, 10 agerpen baino gehiago dituzten aditzen datuak bakarrik izango ditugu kontuan. Hala, esaldi horietan oinarrituz, guk erabakitako aldi bereko kasu-konbinazioak, aipatu corpusan agertu diren edo ez begiratu dugu. Agertu badira, kasu bakoitza zein maiztasunaz; eta agertu ez badira, zein kasu agertu diren, zein maiztasunez eta zein konbinaziotan (binaka, hiruak...). Horiek oso garrantzitsuak izango dira azterketa honetatik aterako ditugun ondorioetarako.

4.5 Hautatutako aditzak

Aditzak hautatzeko, taldean eskura dugun EEBS⁴ corpusetik 22.000 hitzeko lagin bat hartu eta 622 aditz zerrendatu ditugu, maiztasun gehienekotik gutxienera; zerrendatutako guztietatik 110 aditz hautatu ditugu irizpide desberdinen arabera: batetik, maiztasunari begiratu diogu, maiztasun handienekoak hartuz lehenik, baina horiekin batera, azpikategorizazio-jokamolde aldetik ustez interesgarri izan zitezkeen beste batzuk ere hautatu ditugu. Hona hemen osatutako zerrenda:

abestu, adierazi, afaldu, agertu, ahaztu, aldatu, amaitu, argitu, asmatu, atera, aurkeztu, aurkitu, banandu, barkatu, begiratu, besarkatu, bete, bihurtu, borrokatu, bukatu, bultzatu, deitu, edan, egin, egon, egotzi, ekarri, elkartu, eman, entzun, erabili, eragin, eraman, erantzun, erein, erosi, erre, erreparatu, esan, eskaini, eskatu, eskiatu, etorri, eutsi, ezkondu, eztabaidatu, gelditu, gertatu, gogoratu, hartu, hasi, hautatu, hautsi, hazi, hil, hornitu, hozkatu, hustu, ibili, igo, ikasi, ikusi, irakurri, irau, iritsi, isildu, isuri, izan, izendatu, izerditu, jaitsi, jakin, jan, jantzi, jarraitu, jarri, jaso, jo, joan, jokatu, josi, kezkatu, kokatu, konparatu, konturatu, korritu, loratu, lortu, lotu, marraztu, mintzatu, moztu, mugitu, musukatu, nahastu, nekatu, oroitu, otu, pasa(tu), poztu, saldu, saltatu, sartu, sinistu, topatu, ukitu, ulertu, zamatu, zeharkatu, zutitu

Ikus daitekeenez, ez dago osagai bat baino gehiagoko aditzik (*lo egin, zain egon, axola izan, ari izan, barre egin...*). Hau, alde batetik, aditzen artean hautaketa egiteko irizpidea ere izan da nolabait, baina batez ere gure tresna automatikoek, oraindik unitate lexikal gisa analizatzeko dituzten zailtasunek eraman gaituzte hauek bazterreratzea. Esan behar dugu, halere, taldean badirela hitz anitzeko unitate lexikalen inguruko azterketak, eta dagoeneko urrats batzuk ere emanak ditugula (Aduriz et al., 1996).

5 ONDORIOAK ETA ETORKIZUNEN LANAK

Hautatutako prozedura honen bitartez, hauxe da lortzen duguna, edo bederen lortu nahi genukeena: aditz-forma bakoitzak, guk definitutako irizpideetan oinarrituta izan ditzakeen balio desberdin guztiak banatuta. Hau da, modu honetan hain eza-gun diren balio kausatibo eta inkoatiboak, esaterako, bereizita agertuko zaizkigu, delako aditzaren bi balio gisa. Balio guztiak zehaztea, halere, oso puntu inportantea

4 EEBS (Egungo Euskararen Bilketa Sistemakoa), UZEI etengabe aberastuz doan corpusaren izena da, eta XX. mendean erditik hasita gaurdaino agertu diren euskarazko testuen bilketa sistemakoa egitea du helburu. (Urkia & Sagarna, 1991).

iruditzen zaigu aditz banakoak corpusean onar ditzakeen egiturak definitzen diren aldetik. Ondoren, esan dugun bezala, atera zaizkigun balio desberdin horietan, kasu bakoitzak zein izaera semantiko duen aztertuko dugu, eta horren arabera, balio semantiko berak edo antzekoak dituzten aditzekin multzo koherenteak osatzen saiatuko gara. Honen bidez, uste dugu, sintaxia eta semantikaren artean dagoen harreman horren berri ematen dugula neurri handi batean. Horretaz gain, aditzak multzokatzeak duen garrantzia ukaezina da, lexikoia (edota hiztegia) modu orokorrean antolatzeke, eta horrela informazioa alperrik ez errepikatzeke.

Bestetik, azpikategorizazioaren kontzeptua lauso samar gelditzen da, azpikategorizazioa derrigorrezkotasunarekin lotzen badugu, bederen. Izatez, derrigorrezkotasuna kontu irristakorra da, are gehiago euskara bezalako pro-drop hizkuntzetan, eta euskara bertan ageri zaizkigun hainbat fenomenoren aurrean. Guk hemen planteatzen dugun ikuspegian, derrigorrezkotasunari baino, aukerakotasuna eta ezinezkotasuna markatzeari eman diogu lehentasuna. Hain zuzen ere, aditz bakoitzak perpaus zuzen bat osatzeko "beharrezkotzat" jotzen diren elementuak zehaztea bezain garrantzitsua edo are garrantzitsuagoa da, gure ustez, zein elementu ez duen onartzen zehaztea. Eta hau guztiz bat dator aplikazio konputazionalen barruan egiten diren Gramatika Zuzentzaileek beharrezko duten informazioarekin.

Azkenik, esan behar dugu, azterketarako iturri gisa corpora hartzeak, ondorio garbi bat dakarrela: bertan agertu dena bakarrik aztertuko dugu; alegia, aditz batek corpusean erakutsi dituen erabilerak. Honenbestez, azterketa linguistiko sakona jartzen badugu helburu nagusi gisa, beharbada hankamotz geratuko gara; baina, aldiz, gure helburu behinena corpora aztertu ahal izatea bada, zentzu horretan asetu egingo ditugu gure asmoak. Eta esana dugun bezala, une honetan horixe da gure ikuspegi konputazionaletik lortu nahi duguna. Hortaz, hasteko oso baliagarri eta beharrezkotzat jotzen dugu azterketa hau.

ERREFERENTZIAK

- ABNEY, S. P. (1991). *Parsing by chunks*. R.C. Berwick, S.P. Abney, and C. Tenny eds, Principle-Based Parsing: Computation and Psycholinguistics, Dordrecht: Kluwer.
- _____ (1997). *Part-of-Speech Tagging and Partial Parsing*. S. Young, and G. Bloothoof ed, Corpus-Based Methods in Language and Speech Processing. Dordrecht: Kluwer.
- BURZIO, L. (1981). *Intransitive Verbs and Italian Auxiliaries*. [Doctoral dissertation, MIT], Cambridge, MA.
- ADURIZ I., ALDEZABAL I., ARTOLA X., EZEIZA N., URIZAR R.: "Multiword Lexical Units in EUSLEM: A Lemmatiser/Tagger for Basque", COMPLEX'96 Forth Conference on Computational Lexicography and Text Research. Budapest, Hungaria. Irailak 15-17, 1996.
- ALDEZABAL I., GOJENOLA K., SARASOLA K. "Subcategorización verbal vasca: propuesta inicial y herramienta de validación". SEPLN'98. Alicante. 23-25 de Septiembre, 1998.
- ALDEZABAL I., GOJENOLA K., OROÑOZ M. "Combining Chart-Parsing and Finite State Parsing" Proceedings of the European Summer School in Logic, Language and

- Information (ESSLLI). Student Session 99. Utrecht, the Netherlands, August 10-14, 1999.
- ALDEZABAL I., GOJENOLA K., SARASOLA K. "A Bootstrapping approach to parser development". International Workshop on Parsing Technologies (IWPT 2000). Trento (Italy), 2000.
- ALDEZABAL I., ARANZABE M., ATUTXA A., GOENAGA P., GOJENOLA K., SARASOLA K. "Extracción masiva de información sobre subcategorización verbal vasca a partir de corpus" SEPLN'2001, Septiembre 12-14, Jaén, 2001.
- ARRIOLA, J. M., SOROA A. (1996). "Lexical Information Extration for Basque" Student Conference in Computational Linguistics, 40-47, Montreal
- CHOMSKY, N. A. (1981). *Lectures on Government and Binding*. Dordrecht: Foris.
- CRUSE, D.A. (1972). "A note on English Causatives". *Linguistic Inquiry* 3, 522-528.
- (1973). "Some thoughts on Agentivity". *Journal of Linguistics* 9, 11-23.
- ELHUYAR KULTUR ELKARTEA (1996). *Elhuyar hiztegia*, Usurbil.
- ETXEPARE, R. (2001). Hualde, J.I & Ortiz de Urbina, J. (edits.), *A grammar of Basque*. Unpublished. University of Illinois and Deustuko Unibertsitatea.
- EUSKALTZAINDIA (1993). *Euskal Gramatika laburra: Perpaus bakuna*, Bilbo
- (2000). *Hiztegi Batua*. [*Euskera*, 45. Liburukia, Bilbo].
- FODOR, J.A. (1970). "Three Reasons for Not Deriving *Kill* from *Cause to Die*", *Linguistic Inquiry* 1, 429-438.
- GRACIA I SOLÉ, LL. (1987). *La Teoría Temática*. [Ph D. Publicacions de la Universitat Autònoma de Barcelona]. Bellaterra.
- HALE, K. L. & KEYSER, S. J. (1986). "Some transitivity Alternation in English". Lexicon Project Working Papers 10, Center for Cognitive Science, Cambridge, MA: MIT Press.
- (1991). "On the syntax of Argument Structure". Unpublished ms., MIT, Cambridge, MA.
- HALLIDAY, M. A. K. (1968). "Notes on Transitivity and Theme in English Part III". *Journal of Linguistics* 4, 179-215.
- JACKENDOFF, R.S. (1990b). *Semantic Structures*. Cambridge, MA: MIT Press.
- LEVIN B., (1993). *Verb classes and alternations*, Chicago: The University of Chicago Press.
- LEVIN, B. & RAPPAPORT, M. (1992b) *Unacusativity: At the Syntax-Semantics Interface*, unpublished ms., Northwestern University and Bar Ilan University, Evanston, IL an Ramat Gan
- OYHARÇABAL, B. (1993) "Verb Agreement with Nonargumentns: On Allocutive Agreement", Hualde J.I. & Ortiz de Urbina, J. (edits.), *Generative Studies in Basque Linguistics*, Amsterdam: John Benjamins, 89-114.
- PERLMUTTER, D. M. (1978). "Impersonal Passives and the Unaccusative Hypothesis", *BLS*, 4 157-189.
- PUSTEJOVSKY, J. (1995). *The Generative Lexicon*. Cambridge, MA: MIT Press.
- SARASOLA, I. (1997). *Euskal Hiztegia*. Donostia: Kutxa Fundazioa.
- URKIA, M. & SAGARNA, A. (1991). Terminología y Lexicografía asistida por ordenador. La experiencia de UZEI. SEPLN'91, Donostia.
- ZABALA, I. (1993). *Predikazioaren teoriak Gramatika Sortzailean (Euskararen kasua)*. Doktoretza-tesia. Euskal Filologia Saila, Euskal Herriko Unibertsitatea.