



IKER
GAZTE
NAZIOARTEKO
IKERKETA EUSKARAZ

I. IKERGAZTE

NAZIOARTEKO IKERKETA EUSKARAZ

2015eko maiatzaren 13, 14 eta 15
Durango, Euskal Herria

ANTOLATZAILEA:
Udako Euskal Unibertsitatea (UEU)

GIZA ZIENTZIAK

**Euskararako HPSG gramatikaren
lehen proposamena**

Jon Alkorta Agirrezabala

24-31 or.
<https://dx.doi.org/10.26876/ikergazte.i.02>

ANTOLATZAILEA:



BABESLEAK:



LAGUNTZAILEAK:



Euskararako HPSG gramatikaren lehen proposamena

Jon Alkorta Agirrezabala

IXA ikerketa-taldea (EHU/UPV). jon.alkorta@ehu.es

Laburpena

Sintaxi konputazionala Hizkuntzalaritza Konputazionala ikerlerroko adarretako bat da. Bere helburua sintaxiko ezagutza formalizatzea da. Sintaxi konputazionalan, sintaxia formalizatzeko hainbat formalismo daude, ezaugarri eta helburu ezberdinekin. Artikulu honetan, euskararako HPSG gramatikaren lehen urratsak azalduko ditugu eta bidean izandako arazoak eta bilatutako konponbideak kontatuko ditugu.

Hitz gakoak: HPSG gramatika, euskara, hitz-hurrenkera, aditz-sail ezberdinak

Abstract

Computational syntax is a field of Computational Linguistics related to the description of the phenomena involved in the combination of words into sentences, that is, syntactic knowledge. There are various formalisms for its description, most of them centered on the notion of a grammar of a language. In this article, we present the first formalization of a grammar of Basque in the framework of Head-Driven Phrase Structure Grammars (HPSG).

Keywords: HPSG grammar, Basque language, word order, types of verb

1 Sarrera eta motibazioa

Sintaxi konputazionala hasieratik izan da alor garrantzitsua Hizkuntzalaritza Konputazionalaren barruan. Sintaxi Konputazionalan, sintaxian bezala, osagai sintagmatikoen euren artean dituzten loturak ikertzen dira. Sintaxi konputazionalak, Hizkuntzalaritza Konputazionalako beste alorrek moduan, ohikoa den arazo bati egin behar dio aurre: anbiogutasuna eta bere ebazketari. Arazoak arazo, sintaxi konputazionalak aurrera egin du eta analizatzaile sintaktikoak horren adibide dira.

IXA taldean, sintaxi konputazionalaren ikuspegitik euskara landu da eta zenbait tresna eta produktu egin dira. Tresna eta produktu horien atzean beti egon da oinarri teoriko bat. Hala ere, oraindik zenbait oinarri teoriko ez dira landu, HPSG gramatika esaterako, eta horietan ere lehen urratsak egin nahi dira.

2 Arloko egoera eta ikerketaren helburuak

Sintaxi konputazionalan hainbat formalismo eredu erabili izan dira gaurdaino. Alde batetik, ezagutza linguistikoan oinarritutako sintaxia dago. Honen barruan, *Testuingururik gabeko gramatikak* (TGG), TGG sinpleak nahiz baterakuntzan oinarritutakoak (*Lexical functional grammar*, *Head-driven phrase structure grammar*, *Generalized phrase structure grammar*), daude. Bestalde, teknika probabilistikoetan oinarritutako sintaxia ere badago.

Euskararako sintaxi konputazionala lehena lantzen Joseba Abaitua izan zen. *Lexical functional grammar* (LFG) (Abaitua, 1988) gramatika oinarritutako analizatzaile sintaktikoa garatu zuen. Gerora, IXA ikerketa-taldean euskarazko hainbat gramatika eredu garatu dira. *PATR II* (Gojenola, 2000) eta *Euskarako Dependentzia Gramatika Konputazionala* (EDGK) (Aranzabe, 2008) ditugu besteak beste. *PATR II* TGG sinplea da eta EDGK egoera finituko mekanismoetan oinarritzen da. Gainera, teknika probabilistikoetan oinarritutako sintaxian, *Maltixa* izeneko gramatika estatistikoa (Bengoetxea, 2015) aurki dezakegu.

HPSG gramatikak 1990. hamarkadan hasi ziren garatzen beste hizkuntzetan, oinarri teorikoa sortu

eta gutxira. Ingeleserako *LinGO English Resource Grammar* (ERG) eta japonierarako *Jacy Japanese Grammar* (Jacy) dira nabarmentzekoak. Azken honetan aditza azpikategorizatu egiten da *SAT* eta *VAL* atributuen bidez.

Ikerketa honen helburua euskararako HPSG gramatikan lehen urratsak egitea da, gerora pixkanaka gramatika garatzen joan ahal izateko.

3 HPSGren deskribapen laburra

Buruak gidatutako egitura sintagmatikoen gramatika (ingelesez, *Head-driven phrase structure grammar*, *HPSG*) (Pollard eta Sag, 1994) gramatika sortzailea da eta sintagma-egituren gramatikan oinarritzen da. Ez dago eraldaketarik eta erregelen hurrenkerarik. Gramatika formalismo honek hainbat teoria sintaktiko eta semantikoen eragina izan du, tartean, LFG gramatika eta *Minimal recursion semantics* (MRS) semantika formalismoa. Gramatikan bi atal bereizten dira: tresna deskribatzaileak (*descriptive devices*) eta gramatika osagaiak (*grammar components*).

3.1 Tresna deskribatzaileak

- Ezaugarri-egiturak (*Feature Structures*, *FS*). Atributu-balio bikoteen zerrenda bat da.
- Motak (*Types*). HPSGk hizkuntza-objektuak mota ezberdinetan sailkatzen ditu.
- Baterakuntza (*Unification*). Baterakuntzan, bi FS edo gehiagoren baturak FS berri bat sortzen du, aurretik adierazi bezala.
Adibidez, $[AGR.GEN\ fem]$ eta $[AGR.NUM\ sing]$ ezaugarri-egituren batura zuzena da: $[AGR/GEN\ fem, NUM\ sing]$. Izan ere, generoa eta numeroa batzen dira. Aldiz, $[AGR.GEN\ fem]$ eta $[AGR.GEN\ masc]$ ezaugarri-egiturak ezin dira batu; bietan atributua generoa delako.
- Egitura partekatua (*Structure sharing*). FS barruko bi atributu ezberdin edo gehiagok balio bera dutela adierazi behar da. Hurrengo kasuan esaterako, komunztadurak eta subjektuak balio bera partekatzen dute eta beraz, egitura ondo eraikita dago.

(1) $[AGR\ \#agr, SUBJ\ [AGR\ \#agr]]$

3.2 Gramatika osagaiak

- Zeinu mota (*the type sign*). Zeinua bi modutakoa izan daiteke.
 - Hitza mota (*word type*). ARG-ST atributu gehigarri bat du: zerrenda bakarrean azpikategorizatutako argumentuak adierazteko.
 - Sintagma mota (*phrase type*). Hierarkia herentzia bat osatzen dute.
- Hatsarreak (*principles*). Lengoia baten adierazpen egokiak zehazten ditu. Ezaugarri-egitura bidez implementatuta dago eta AVM¹ erabiliz deskribatzen da.
- *Minimal Recursion Semantics* (MRS). HPSG gramatiketan erabili ohi den gramatika formalismoa da. Gramatika murriztapenetan eta Konposizio Hatsarrean (*Principle of compositionality*) oinarritzen da.

4 Hizkuntza gertakarien azterketa

Bi helburu nagusi jarri ditugu: gramatikak euskararen hitz-hurrenkera posible guztiak eta aditz-sail ezberdinak aztertu ahal izatea. "Control" egiturak eta galdegai-mintzagai partizioak lan honetatik at utzi ditugu (Engdahl eta Vallduv, 1995). Gure analisian, hitzak morfemetan zatituko ditugu.

4.1 Hitz-hurrenkera

HPSG gramatiken eredu tradizionalan, aditzaren argumentuak zerrenda batean doaz. Zerrendetan bi atal bereizten dira: FIRST eta REST.

¹AVM edo *Attribute value matrix* sintagma-egitura gramatiketan erabili ohi den ezaugarri-egitura da eta atributu-balio bikotea izaten da.

(2) *Daramakio*

ARGS	FIRST	ERG	ABS	
	REST	FIRST	FIRST	DAT
		REST	REST	null

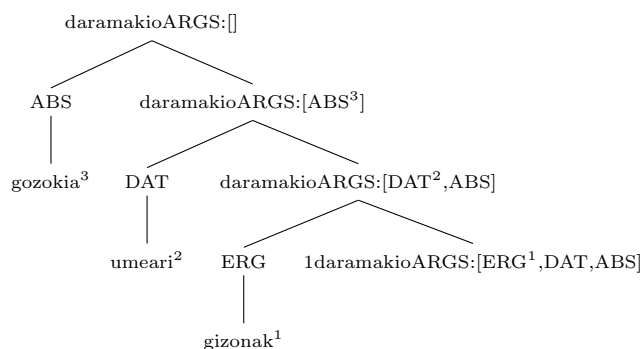
Goiko adibidean, zerrendako lehen ergatiboa da eta ondoren datoz absolutiboa eta datiboa. Osagai horiek erregela sintaktikoen bidez konbinatzen dira:

(3) *gizonak-ERG + dakar:*

ARGS	FIRST	ABS	
	REST	FIRST	DAT
		REST	null

Erregela sintaktiko honen bidez, ARGS zerrendako lehen elementua, ergatiboa, *jan* egiten da. Per-pausa erdi-osatuta dago eta guztiz osatzeko absolutiboa eta datiboa falta dira. Erregela honek *umeak ura gizonari eman dio* bezalako perpausak aztertuko lituzke baina ez *umeak eman dio ura gizonari* modukoak eta beste hurrenkera bat duten perpausak.

Esaterako, *daramakio* aditzeko argumentu zerrenda [ERG, DAT, ABS] da. HPSGk erabiltzen duen mekanismoan argumentu zerrendako osagai bakoitza *jan* egiten da ondoren agertzen den bezala:



Aditzaren lehen osagaia *gizonak* da. *Gizonak* ergatiboa da eta aditzaren argumentuen zerrendan ere ergatiboa lehenengo dago. Beraz, HPSGren mekanismoak osagai hau *jan* egiten du. Ondorioz, argumentu zerrendan datiboa eta absolutiboa geratzen dira. Ondoren, perpausean datibo kasuan dagoen *umea* dago eta mekanismoak zerrendan hurrengo osagaia den datiboa hartzen du. Beraz, aditzaren argumentu zerrendan absolutiboa geratzen da. Azkenik, perpausean absolutibo kasuan dagoen *gozokia* dago eta HPSGren mekanismoak argumentu zerrendako azken osagaia, absolutibo kasua, aurkitu egiten du.

4.1.1 Arazoa

Hurrenkera finkoa duten hizkuntzetan, ingelesean eta alemaneran esaterako, HPSGren mekanismo hau egokia da, argumentu zerrenda bakarra egongo delako. Baina hitz-hurrenkera askea duten hizkuntzetan, euskarari esaterako, hurrenkera ezberdin guztiak jarri beharko lirateke eta horrenbestez, sarrera lexiko asko egin beharko lirateke: [ERG, DAT, ABS], [DAT, ABS, ERG]... eta hau ez da oso egokia. Gainera, kontuan hartu behar da tartean adjuntuak ager daitezkeela edozein tokitan (*atzo, kotxean...*) egoera zailduz.

4.1.2 Konponbidea

Konponbideetako bat *Categorical Grammar* (Karttunen, 1989) izan daiteke. Gramatika honetan, sarrera lexikoek hitzak sintagmetan bihurtzeko konbinaketei buruzko informazioa kodetuta dute. Hitzaren ezaugarrien konbinaketa bere kategoria lexikoan adierazten da. Bi kategoria mota daude: oinarritzkoak (*basic*) eta deribatuak (*derived*), *functor* ere deituak. Izen arruntek eta izen-sintagmak oinarritzko kategoriak dira eta determinatzaileak eta aditz iragangaitzek kategoria deribatuen multzoa osatzen dute.

Oinarritzko Kategoriak (<i>Basic Categories</i>)	Kategoria Deribatuak (<i>Derived Categories</i>)
CN (izen arrunta)	NP/CN (determinatzailea)
NP (izen-sintagma)	S\NP (aditz iragangaitza)

Kategoria deribatuetako etiketek sintagma-egitura gramatiketako erregela sintaktikoen funtzio bera dute. NP/CN bidez, determinatzailea izen arruntarekin batu eta izen-sintagma eratzen dela adierazten da. Sintagma-egituraren erregela bera $NP \rightarrow Det\ CN$ bidez adieraziko litzateke.

Kategoria deribatuen kategoria etiketek honako itxura dute: (*result direction argument*). *Result* eta *argument* kategoria etiketak dira, sinpleak edo konplexuak eta *direction* / edo \ izan daiteke, kategoria deribatua argumentuaren aurretik (/) edo ondoren (\) baldin badator.

Aditz iragangaitzaren etiketa, $S \backslash NP$, da. Honen bidez, aurretik datorren izen-sintagmarekin sintagma perpausa osatzen duela adierazten da. Aditz iragankorrek $(S \backslash NP) / NP$ kategoria dute eta adberbioek $(S \backslash NP) \backslash (S \backslash NP)$. Aditz iragankorra izen-sintagmarekin eskuinetik batuko litzateke eta adberbioa aditz iragangaitzari ezkerretik.

Hala ere, ingeleserako eta euskararako *Categorial Grammar* ezberdinak dira. Ingelesean, hitzhurrenkerak finkoa da eta *eat* bezalako aditz iragankor batean $eat-(np/s) \backslash np$ sarrera dugu, hau da, np (objektua) eskuinetik bilatzen du: $the\ meat\ np = (np/s)$. Hori bilatzen badu, emaitza (np/s) izango da. Ondoren, np/s kategoriak, np (subjektua) bat bilatzen du eskuinetik: $the\ man\ np + np/s = s$. Ondorioz, bigarren elementua ere jan egiten du eta azken emaitza s da. Ingelesean nahitaez subjektuak eskuinetik etorri behar du.

Euskaran, ordea, beti ez da hala gertatzen: *umeak oparia lagunari eman dio, oparia lagunari eman dio umeak, lagunari eman dio oparia umeak...* Horregatik, Karttunen (1989) irizpidea jarraitzea erabaki dugu. Kategoria nagusia izen-sintagma da eta berak konbinatzerakoan aditz-mota jakin batzuk eskatzen ditu. Ondorioz, esaterako *umea-ABS* hitzaren kategoria vp/vp da gramatika kategorialean. Kategoria honek ezkerretik aditz bat datorrela eta hitzarekin konbinatzean beste aditz bat (zehazkiago, aditz-proiekzioa $dator-vp + umea-vp/vp = dator-umea-vp$) sortzen dela adierazten du.

4.2 Aditz-sail ezberdinak

Aditza da aditz-sintagmaren burua eta aditz-sintagmak aditzak eskatzen dituen zenbat osagai (argumentu) ere agertzen dira. Aditzak aztertzerakoan, argumentu kopuruari eta euren izaerari, hau da, duten kasuari begiratu behar zaio. Horrela, aditzak iragangaitzak edo iragankorrak (soilak eta datibodunak) dira. Beraz, aditz-sail ezberdinak ditugu.

- (4) Anaia gaztea da
Mikelek udare bat jan du
Josuk niri udare bat eman dit.

Adibide honetan, lehena iragangaitza da; aditzak argumenturik eskatzen ez duelako; bigarren iragankor soilak da objektu zuzena duelako eta azkena aditz iragankor datiboduna da, objektu zuzenaz gain zehar objektua ere eskatzen duelako.

4.2.1 Arazoa

Aditzak izaera ezberdinekoak izatea buruhauste handi bat da baterakuntza oinarritzat duen HPSG gramatikan. Aditzak modu berean tratatzea zailtzen du, izan ere, aditzaren arabera argumentu kopurua ezberdina da. Gainera, elipsiak egoera gehiago zailtzen du. Gerta liteke, argumentu bat perpausuan gauzatzea nahiz eta argumentu hori ezkutuan dagoen.

- (5) Mikelek $\emptyset$ jan du
Josuk $\emptyset$ eman dit.

Adibide hauek goikoak dira, baina nahiz eta argumentua gauzatzen den argumentua ez dago begibistan. Lehen perpausuan, absolutibo kasuko argumentua ez da ageri eta bigarrenetan, absolutibo eta datibo kasukoak.

4.2.2 Konponbidea

Aditz-sail ezberdinen arazoa saihesteko bi atributu sortzea erabaki dugu. Bat *SUBCAT* atributua da eta bestea *SAT* atributua, azken hau japonieraren HPSG gramatikatik hartutako ideia (Siegel eta Bender, 2002). Atributu honen bidez, aditzaren argumentu komunztadura asetuta dagoen edo ez adierazten da.

Euskararako garatu dugun HPSG gramatikan, *SUBCAT* balioa eta *SAT* atributua erabiliko ditugu. *SUBCAT* balioak aditzak eskatzen dituen argumentuak biltzen ditu eta *SAT* atributuak argumentu horiek asetzen diren edo ez adierazten du.

Esaterako, pentsa dezagun *daramakio-ERG-ABS-DAT* aditza dugula. Goran aipatu bezala, kotxea hitzarekin konbinatzen badugu berriro (kotxea) daramakio aditza izango dugu baina balentzian (*VAL*),

hau da, *SAT* atributuen zerrendan aldaketa bat gertatuko da: *ABS* den elementua sartu da.

Kotxea eta daramakio hitzak konbinatu baino lehen, *daramakio* aditzak honako balioak ditu:

- (6) *daramakio* = [SUBCAT [SUBJ ERG, OBJ ABS, IOBJ DAT]
VAL [SUBJ unsat, OBJ unsat, IOBJ unsat]]

Azpikategorizazioan (*SUBCAT*), subjektuarekin kasua ergatiboa dela, objektuarena absolutiboa eta zehar objektuarena datiboa dela adierazten da. Bestalde, balentzian (*VAL*) subjektu, objektu eta zehar objektu atributuek balio asegabeak (*unsat(urated)*) dituztela adierazten da.

Baina *kotxea* eta *daramakio* konbinatzean, balentziako balioetan aldaketa gertatzen da, *kotxea* objektua baita:

- (7) *daramakio* = [SUBCAT [SUBJ ERG, OBJ ABS, IOBJ DAT]
VAL [SUBJ unsat, OBJ sat, IOBJ unsat]]

Bi hitzak konbinatzean, *VAL.OBJ* asetu (*sat(urated)*) egiten da, kotxea objektua delako. Beraz, atributu honen bidez, aditz-mota ezberdinak erraz adieraz daitezke:

- (8) NOR: VAL [SUBJ unsat, OBJ sat, IOBJ sat] NOR-NORI-NORK: VAL [SUBJ unsat, OBJ unsat, IOBJ unsat]
NOR-NORI: VAL [SUBJ unsat, OBJ sat, IOBJ unsat]
NOR-NORK: VAL [SUBJ unsat, OBJ unsat, IOBJ sat]

NOR aditzetan, subjektua bakarrik asetu daiteke; NOR-NORI aditzetan subjektua eta zehar objektua; NOR-NORK aditzetan subjektua eta objektua eta azkenik, NOR-NORI-NORK aditzetan subjektua, objektua eta zehar objektua.

5 HPSG gramatika garatzen

5.1 HPSG gramatika garatzeko ingurunea

HPSG gramatikak lantzeko inguruneak hiru taldetan banatzen dira. Alde batetik, murritzapenean oinarritutako sistemak daude. Beste batzuk, berriz, erlazioetan oinarritutako sistemak dira eta hauen barruan, sintagma-egituretan oinarritzen den azpitaldea ere badago. Azkenik, bi sistemak konbinatzen dituztenak daude.

Guk bi sistemak konbinatzen dituen *LKB berria* (Copestake, 2002), *Linguistic Knowledge Builder*ren ondorengoa, erabiliko dugu gramatika garatzeko. Jarraian, helburutzat jarri ditugunak betearazteko sortu ditugun erregeletako batzuk aztertuko ditugu.

5.2 Hitz-hurrenkera

Ondorengo irudietan, hitz-hurrenkeraren arazoa konpontzeko erabili ditugun bi erregelak ageri dira.

- | | |
|--|---|
| <p>(9) backward-application := phrase-verb &
[CAT #cat,
SUBCAT #subcat,
VAL #valnew,
NORA #direction,
ARGS <phrase-verb & [CAT #cat,
SUBCAT #subcat,
VAL #val,
NORA #direction],
phrase-nom & [CAT [RESULT [CAT #cat,
SUBCAT #subcat,
VAL #valnew,
NORA #direction],
DIRECTION #direction & backward,
ACTIVE phrase-verb &
[CAT #cat,
SUBCAT #subcat,
VAL #val,
NORA #direction]]>].</p> | <p>forward-application := phrase-verb &
[CAT #cat,
SUBCAT #subcat,
VAL #valnew,
NORA #direction,
ARGS <phrase-nom & [CAT [RESULT [CAT #cat,
SUBCAT #subcat,
VAL #valnew,
NORA #direction],
DIRECTION forward,
ACTIVE phrase-verb &
[CAT #cat,
SUBCAT #subcat,
VAL #val,
NORA #direction]]],
phrase-verb & [CAT #cat,
SUBCAT #subcat,
VAL #val,
NORA #direction & backward]>].</p> |
|--|---|

Ezkerreko erregela sintaktikoak lehenik aditz-sintagma eta ondoren, izen-sintagma hurrenkera duten perpausak aztertzen ditu. Argumentu zerrendan (*ARGS*), lehenik eta behin aditz-sintagma (*phrase-verb*) dugu; kategoria (*CAT*), azpikategoria (*SUBCAT*), balentzia (*VAL*) eta norabidea (*NORA*) balioekin. Ondoren, izen-sintagma (*phrase-nom*) dago.

Izen-sintagma honen barruan dagoen aditz-sintagma (*phrase-verb*) aktibatu (*ACTIVE*) egiten da eta atzeranzko norabidea (*DIRECTION #direction & backward*) hartzen du; izen-sintagmarekin batzeko. Biak batzearen ondorioz (*RESULT*), beste aditz-sintagma bat (*phrase-verb*) sortzen da eta kategoria (*CAT*), azpikategoria (*SUBCAT*), balentzia (*VAL*) eta norabidea (*NORA*) balioak heredatzen ditu. Aipatzekoa da, baterakuntzaz sortu den aditz-sintagma honek balentzia berria (*#valnew*) duela.

Eskuineko erregela sintaktikoak berriz, alderantzizko hurrenkera aztertzen du, hau da, lehenengo izen-sintagma eta ondoren aditz-sintagma dutenak. Argumentu zerrendan, lehenengo izen-sintagma (*phrase-nom*) dugu eta ondoren, aditz-sintagma (*phrase-verb*). Aditz-sintagmak kategoria (*CAT*), azpikategoria (*SUBCAT*), balentzia (*VAL*) eta norabidea (*NORA*) balioak biltzen ditu.

Izen-sintagman, aditz-sintagma aktibatzen (*ACTIVE*) da. Aditz-sintagma honek kategoria (*CAT*), azpikategoria (*SUBCAT*), balentzia (*VAL*) eta norabidea (*NORA*) balioak ditu. Aditz-sintagma honek aurreranzko norabidea (*DIRECTION forward*) hartzen du eta izen-sintagmarekin egiten du topo. Bien arteko baterakuntzaren ondorioz, beste aditz-sintagma bat sortzen da kategoria (*CAT*), azpikategoria (*SUBCAT*), balentzia (*VAL*) eta norabidea (*NORA*) balioak heredatzen dituen. Aurreko erregela sintaktikoa bezala, hemen balentziak balio berria (*#valnew*) hartzen du, sortzen den aditz-sintagma berria delako.

5.3 Aditz-sail ezberdinak

Gramatikak aditz-sail ezberdinetako perpausak azter ditzan hainbat erregela mota sortu behar izan ditugu.

(10) *word-verb* := *word* & *nuclear-arguments* & [*CAT simple-cat*].

Motak definitzerakoan, aditza gure gramatikan hitz mota bakarra da. Aditz izaerako hitzez (*word*) eta kategoria simple batez (*CAT simple-cat*) osatuta dago.

(11) *phrase-verb* := *phrase* & *nuclear-arguments*.

Aditz-sintagma (*phrase-verb*), berriz, sintagma batek (*phrase*) eta argumentu nuklear batek (*nuclear-arguments*) osatzen dute.

Balentzia (*VAL*) eta azpikategoria (*SUBCAT*) hierarkiak ere sortu ditugu.

(12) <i>subcat</i> := *top* & [SUBJ case, OBJ case, IOBJ case].	<i>val</i> := *top* & [SUBJVAL saturation, OBJVAL saturation, IOBJVAL saturation].
---	--

Azpikategorizazioa hierarkia (*subcat*) hiru kasuk (subjektua [*SUBJ*], objektua [*OBJ*] eta zehar objektua [*IOBJ*]) osatzen dute. Balentzia hierarkia (*val*), berriz, hiru balentziek (subjektua [*SUBJVAL*], objektua [*OBJVAL*] eta zehar objektua [*IOBJVAL*]) osatzen dute.

Balentzia hauek asetu egin daitezke edo ez (*saturation*), aditzaren ezaugarrien arabera. Azpikategoriak aditzak eskatzen dituen osagaien gauzapean sintaktikoak irudikatzen ditu eta balentziak osagai horien gauzapean sintaktikoak asetzen diren edo ez zehazten du.

(13) <i>case</i> := *top*. <i>erg</i> := case. <i>abs</i> := case. <i>dat</i> := case.	<i>saturation</i> := *top*. <i>sat</i> := saturation. <i>unsat</i> := saturation.
--	--

Adibide honetan, euskarak dituen kasuak hierarkia eta saturazioa hierarkia daude. Kasua hierarkia bat (*top*) da eta bere barnean hiru kasu daude: ergatiboa (*erg*), absolutiboa (*abs*) eta datiboa (*dat*). Balentziak asetu egin daitezke aditzaren izaeraren arabera. Asetasuna (*saturation*) hierarkia bat (*top*)

da. Bi adar ditu: asetzea (*sat*) edo asegabetasuna (*unsat*). Lexikoan, aditz-sail ezberdinak osatu ditugu; bai aditz trinkoak bai perifrastikoak. Jarraian, aditz perifrastikoen adibide batzuk azalduko ditugu.

- | | |
|---|--|
| <p>(14) dio := word &
 [ORTH dio,
 CAT [RESULT phrase-verb &
 [CAT vp,
 DIRECTION backward,
 ACTIVE word-verb &
 [CAT adi,
 SUBCAT #subcat,
 VAL #val],
 SUBCAT #subcat,
 VAL #val]]].</p> | <p>etorri := word-verb &
 [ORTH etorri,
 CAT adi,
 SUBCAT [SUBJ abs],
 VAL [SUBJVAL unsat,
 OBJVAL sat,
 IOBJVAL sat]
].</p> |
|---|--|

Dio aditz laguntzailean, kategorian, hiru elementu daude: emaitza (*RESULT*), norabidea (*DIRECTION*) eta aktiboa (*ACTIVE*). Azken honetan, gorago deskribatutako aditz nagusia dago, *word-verb* hitz motakoa. Aditz nagusi honek atzerantz egiten du eta aditz laguntzailearekin bateratzen da. Ondorioz, aditz-sintagma kategoriako osagai bat sortzen da.

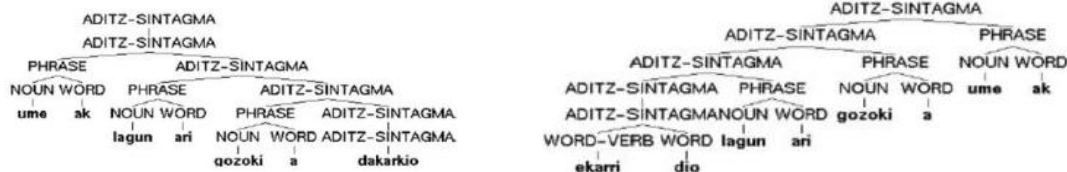
Etorri aditzean, bere kategoria aditz nagusia dela (*adi*) jarri dugu. Azpikategorian, subjektua absolutiboan duela jarri dugu. Aditzak argumentu bakarra eskatzen duenez azpikategorian absolutiboa besterik ez dago. Bestalde, balentzian zehar objektuaren eta objektuaren balentziak asetuta (*sat*) ditu baina subjektuarena ez (*unsat*).

6 HPSG gramatikaren adibideak

Atal honetan, LKB tresnak idatzi ditugun perpausak aztertu ondoren sortu dituen zuhaitzak erakutsiko ditugu.

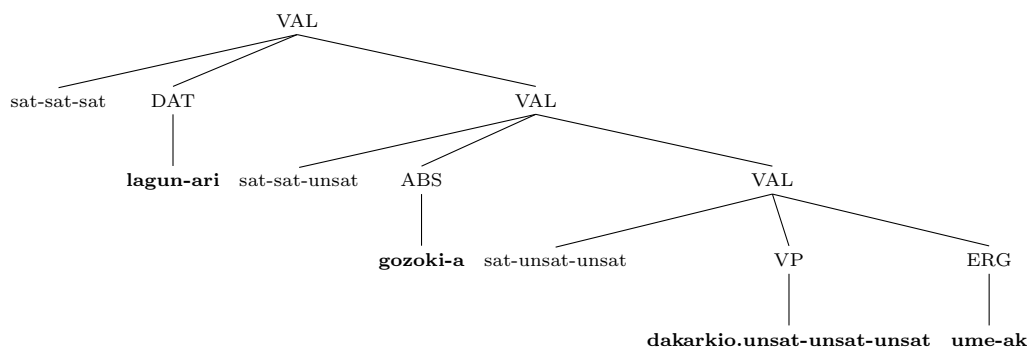
6.1 Hitz-hurrenkera

1 Irudia: *Ume-ak lagun-ari gozoki-a dakarkio* eta *Ekarri dio lagun-ari gozoki-a ume-ak* perpausak



Goiko bi irudiak alderatzen baditugu, euskararako HPSG gramatikak hitz-hurrenkera ezberdineko perpausak aztertzeke gaitasuna duela ikus daiteke. Ezkerreko adibideak NORK-NORI-NOR hurrenkerakoa da eta eskuinekoa, berriz, NORI-NOR-NORK. Gainera, ezkerrekoan aditza lehena da eta eskuinekoa, azkena.

6.2 Aditz-sail ezberdinak



Adibide honetan, berriz, HPSG gramatikako aditz-sail ezberdinak LKB tresnan nola irudikatzen diren ikusten da. Lehenik, *ume-ak* eta aditza bateratu egiten dira eta balentzietan, *SUBCAT* atributua asetu egiten da. Gero, *Ume-ak dakarkio* eta *gozoki-a* bateratu egiten dira eta *OBJSAT* atributua asetzen da. Azkenik, *Ume-ak dakarkio gozoki-a* eta *lagun-ari* bateratu egiten dira, *IOBJSAT* atributua asetuz.

7 Ondorioak eta etorkizunerako norabidea

Helburua euskararako HPSG gramatikaren lehen urratsak egitea izan da, hitz-hurrenkera askea eta aditz-sail ezberdinak landuz.

Gramatika baten hasiera izanik, lan asko dago egiteko. Gehiago sakondu eta fenomeno linguistiko zehatzagoak aztertu behar dira. Semantika (*Minimal Recursion Semantic*) eta adjektiboak teorikoki aztertu eta inplementatu egin behar dira. Beharrezkoa da inplementatu aurretik fenomenoari buruzko azterketa zehatz bat egitea. Gainera, ezezko perpausak eta izenlagunak inplementatzea ere komeni da.

Erreferentziak

- ABAITUA, JOSEBA, 1988. *Complex predicates in Basque: from lexical forms to functional structures*. Manchesterreko Unibertsitatean burututako tesia.
- ARANZABE, MAXUX, 2008. *Dependentzia-ereduan oinarritutako baliabide sintaktikoak: zuhaitz-bankua eta gramatika konputazionala*. Euskal Filologia Saila. UPV/EHU tesia.
- BENGOETXEA, KEPA, 2015. *Estaldura zabaleko euskararako analizatzaile sintaktiko estatistikoa*. Informatika Fakultatea. UPV/EHU tesia.
- COPESTAKE, A. 2002. *Implementing Typed Feature-Structure Grammars*. CSLI.
- ENGDAHL, E., eta E. VALLDUV. 1995. Information packaging in hpsg. *IATL 2: Proceedings of the Tenth Annual Conference of the Israeli Association for Theoretical Linguistics and the Workshop on the SyntaxSemantics Interface* 22–50.
- GOJENOLA, KOLDO, 2000. *Euskararen sintaxi konputazionalerantz. Oinarritzko baliabideak eta beren aplikazioa aditzen azpikategorizazio-informazioaren erauzketan eta erroreen tratamenduan*. Informatika Fakultatea, UPV-EHU tesia.
- KARTTUNNEN, L. 1989. Radical lexicalism. *University of Chicago Press* 43–65.
- POLLARD, C., eta I. SAG. 1994. Head-driven phrase structure grammar. *The University of Chicago Press*.
- SIEGEL, M., eta E. BENDER. 2002. Efficient deep processing of japanese. *Proceedings of the 19th International Conference on Computational Linguistics*, 43–65.

8 Eskerrak eta oharrak

Lan hau *Hizkuntza Azterketa eta Prozesamendua* masterreko amaierako lanetik eratorria dago. Master horretako amaierako lanean tutore izan dudana Koldo Gojenola irakasleari eskerrak eman nahi dizkiot izandako zalantzen aurrean erakutsi duen prestutasunagatik.