

EPEC-RS:
EPEC (Euskararen Prozesamendurako Erreferentzia Corpusa)
rol semantikoekin etiketatzeko eskuliburua

Aldezabal I.; Aranzabe, M.J.; Díaz de Ilarraza, A.; Estarrona, A.; Fernández, K.; Uria, L.

UPV/EHU / LSI / TR 02-2010

Laburpena

Ixa taldearen helburuetako bat da taldean erreferentziatzen dugun *Euskararen Prozesamendurako Erreferentzia Corpora* (EPEC) hainbat hizkuntza-mailatan eskuz etiketatzea, horrela, gero, tresna eta baliabide konputazionalak ebaluatu ahal izateko eta ikasketa automatikoa garatzeko.

Eskuliburu honetan EPEC corpora rol semantiko mailan etiketatzeko gidalerroak aurkezten ditugu. Nagusiki PropBank-en erabilitako eredua (Kingsbury & Palmer, 2002; Palmer *et al.*, 2005) jarraitu dugu.

Beti gertatzen den moduan, eskuliburu hau etengabe eguneratzen joango gara etiketatzea aurrera joan ahala. Horrek esan nahi du hemen jasotzen diren erabakiak argitaratze-data artekoak direla.

Abstract

One of the goals of the IXA research group is the manual annotation of the EPEC corpus (*the Reference Corpus for the Processing of Basque*) at different linguistic levels. This way, it will be possible to evaluate the developed computational tools and resources as well as to carry out machine learning.

In this internal report, we present the criteria fixed to annotate the EPEC corpus at the semantic role level. We have mainly based on the model used in PropBank (Kingsbury & Palmer, 2002; Palmer *et al.*, 2005).

As usual, the report will be continuously updated as we carry on the annotation process. This means that we here present the decisions we have made so far.

Donostian, 2010eko urtarrilaren 18an

AURKIBIDEA:

1. Sarrera	4
2. Helburua	4
3. Abiapuntua	5
4. Wikia eta etiketatze-kopuruak	5
5. Aldez aurretiko baliabideak eta erabakiak.....	6
6. Metodologia	15
7. Irizpideak.....	16
8. Interfazeari begira nahi genituzkeen hobekuntzak	25
Erreferentziak	27
A ERANSKINA: PropBank-eko gidalerroak.....	29
B ERASKINA: EADBko aditzak beren balio sintaktiko/semantikoekin.	62
C ERANSKINA: Landutako aditzentzat aukeratutako “ordain berriak”.....	80

1. Sarrera

Ixa taldearen helburuetako bat da taldean erreferentziatzen dugun EPEC corpusa (Aduriz *et al.*, 2006) hainbat hizkuntza-mailatan eskuz etiketatzea, horrela, gero, tresna eta baliabide konputazionalak ebaluatu ahal izateko eta ikasketa automatikoa garatzeko.

Bai eskuzko etiketatzerako bai baliabide konputazionalak egiteko, oinarri eta irizpide linguistikoak -maila bakoitzari dagozkionak- behar dira, eta eskuliburu honetan rol semantiko mailari dagozkionak aurkezten ditugu. Oraindaino, corpus osoa, eskuz, modu orokorrean hiru mailatan orraztu/etiketatu dugu: segmentazio mailan (Aldezabal *et al.*, 2007a); sintaxi mailan, dependentzien eredu jarraituz (Aldezabal *et al.*, 2007b.); eta semantika mailan, Euskal Wordneteko adierak jarraituz eta soilik izen kategoriakoei gagozkiela (Agirre *et al.*, 2005). Maila apalago batean eta oraindik ere hastapenetan, bidean dira anaforaren (Ceberio *et al.*, 2008) eta diskurtso markatzaileen (Iruskieta *et al.*, 2008) etiketatzea ere.

Maila guztietan bezala, arazo handienetako bat da zein eredu jarraitu erabakitzea, rol semantikoak eredutik eredura aldatu egiten baitira. Guk hemen ProBank-en erabilitako eredu jarraituko dugu (Palmer *et al.*, 2005), alde batetik, abiapuntutzat sintaktikoki etiketatutako corpusa dutelako eta bestetik masiboki, ingelesez gain, beste hizkuntza batzuetarako ere erabili delako (Martí *et al.*, 2007).

Horretaz gain, Ixa taldean bertan egindako azterketa batean proposatutako 100 aditzen lexikoa erabili da (Aldezabal, 2004), informazio osagarri eta konparagarri gisa.

2. Helburua

Helburua da EPEC corpusa rol tematikoekin etiketatzeko gidalerro nagusiak ematea, argi eta garbi utziz zeintzuk diren abiapuntuak, alde aurretiko baliabideak, metodologia eta etiketatzeko oinarri eta irizpideak, bai orokorrak, bai zehatzagoak. Lanean zehar, etiketatzeko tresna ere aipatuko dugu, oso laburki bada ere; izan ere, etiketatze sintaktikorako erabilitako Abarhitz bera (Díaz de Ilarraza *et al.*, 2004) baita. Azken atal gisa, interfazeari begira nahi genituzkeen hobekuntzak zerrendatu ditugu.

3. Abiapuntua

Abiapuntua dependentzia mailan etiketatutako corpora da. Hots, rolei dagokien informazioa dependentziekin etiketatuta datozen elementuei erantsiko zaie.

Han egiten zen bezalaxe, etiketatzea aditzez aditz egingo da eta bakarrik aditzaren mende dauden osagai sintaktikoak etiketatuko dira, hauek baitira argumentu eta adjunturako hautagaiak. Horrek esan nahi du buru gisa aditzik ez duen elementurik ez dela etiketatuko. Azken batean, aditzari loturiko honako etiketadun elementuei jarriko zaie rolei buruzko informazioa *arg_info* etiketaren bitartez:

- a. nsubj
- b. nobj
- c. ncobj
- d. nmod
- e. npred
- f. ccomp_obj
- g. ccomp_subj
- h. cmod
- i. xcomp_obj
- j. xcomp_subj
- k. xcomp_zobj
- l. xmod
- m. xpred

Adibidea: Artzain joango nintzateke Urbiara.

npred (abs, joango, artzain, artzain)
arg_info (go_01, joango, artzain, argM, MNR, - , -)
nmod (ala, joango, Urbiara, Urbiara)
arg_info (go_01, joango, Urbiara, arg4, Destination, helburuko kokapena, -)

Guztira etiketatu beharreko aditzak dependentzietan zerrendatutako berberak dira (C eranskina).

4. Wikia eta etiketatze-kopuruak

Rolei buruzko informazioa trukatzeko eta biltegiratzeko wiki bat sortu da: <http://ixa2.si.ehu.es>. Wikia izanik, izaera aldakorra du eta bertan zintzilikatzen diren zenbait dokumentu etengabe eguneratzen doaz; eskuliburu hauxe bera, esaterako.

Bestetik, etiketatu ahala kopuruak unean-unean ikus daitezke honako helbidean: <http://ixa2.si.ehu.es/abarhitz/rolak.php>

5. Aldez aurretiko baliabideak eta erabakiak

Atal honetan zerrendatuko ditugu, eta nahiko azaletik azaldu, lanari ekiteko eskuartean ditugun baliabideak eta erabakiak.

Esan dugun bezala, gauzarik zailenetakoa eredu bat aurkeratzea da, baina behin ereduak aukeratuta ahalik eta ongien ulertu behar da. Horretan saiatuko gara atal honetan. Esan behar da, hala ere, aukeratutako ereduaren irizpideak ez daudela guztiz argi azalduak eta zenbaitetan zail egiten dela aditzen arteko desberdintasunak ulertzea. 5.1. ataletan zehazki, aukeratutako ingelesezko ereduak aipatzen ditugu eta beroiei buruzko zertzelada batzuk ematen ditugu, eranskinetan jarriz bere horretan egileek egindako gidalerroen dokumentua.

Gainontzeko baliabide eta erabakiak guk geuk sortuak dira eta hobeto ulertzen direlakoan gaude.

5.1 Propbank/Verbnet

5.1.1 PropBank

PropBank-en web-orria honakoa da: (<http://verbs.colorado.edu/~mpalmer/projects/ace.html>):

Paul Kingsbury eta Martha Palmer jo ohi dira PropBank proiektuaren aintzindari (Kingsbury & Palmer, 2002; Palmer *et al.*, 2005))

Proiektuaren nondik norakoak:

- Ingelesko corpus bat semantikoki etiketatzeko proiektua da.
- Oinarrian *Penn Treebank*eko zuhaitz sintaktikoak ditu.
- Aditz-adiera bakoitzari rol-zerrenda bat egokitzen diote.
- Hipotesia:
 - Aditz-forma baten adiera ezberdinak, izaera sintaktikoak eraginda sor daitezke:
 - Adibidez:
 - Ni etxera itzuli naiz
 - *itzuli*: <agente, helmuga>
 - Nik testua euskarara itzuli dut.
 - *itzuli*: <agente, tema, helburuko egoera >

PropBank-eko rolak oso espezifikoak dira, aditz banakoari lotuak esango genuke (Adib.: *buyer*, *thing bought*, *speaker*...); horregatik ez dago rolen zerrenda estandar bat. Beharren arabera erabiltzen doaz. Hauxe da euren filosofia:

- Corpus bat etiketatzeko erabilgarriak diren rolak sortu dituzte.
- Aditz baten rolen kasuistika corpusean oinarrituta dago.

Hala ere, PropBank-en bertan Verbnet-eko rolak ere erabiltzen dira (beheraxeago aztertuko dugu) eta hauek orokorrak dira.

- Etiketatzeari: 2 maila bereizten dituzte:
 - Argumentu zenbakituak (0tik 4ra: Arg0, Arg1, ..., Arg4) eta adjuntuak (adjuntuek ez dute zenbakirik; besterik gabe Modifikatzailearen M-a (ArgM))
 - Rol tematikoak

Hots:

- Argumentu batek rol tematiko desberdinak har ditzake (adib., aditz batean Arg1 *theme* izango da, eta bestean *topic*; edota Arg0 *agent* edo *theme*).
 - Rol tematiko bat argumentu desberdinetan ager daiteke (modu berean, *theme* Arg1 izango da batean eta Arg0 bestean).
- Egitura sintaktikoekin parekatuz gero:

- Aditz iragankorrak: **Peiok liburua** ekarri du

Arg0 = SUBJ

Arg1 = OBJ

- Aditz iragangaitzak: **Leihoa** puskatu da

Arg1 = SUBJ

- Aditz inergatiboak: **Peiok** ondo abestu du

Arg0 = SUBJ

- Adibide bat: *to buy*:

John bought his mother **a dozen roses**

Arg0: John

REL: bought

Arg2: his mother

Arg1: a dozen roses

- Aditz-sarrerak eta esaldi etiketatuak: honako helbideak ikus daitezke
 - <http://verbs.colorado.edu/framesets/>

- Horretaz gain, PropBank-en aditz-adiera bakoitzak bere gauzapean sintaktiko desberdinak ditu (*frameak* deitzen dituzte; azken batean, alternantziak). Adibidez: *tell.01*-en rolset-a eta honi lotutako frame-ak honakoak dira:

rolset *tell.01*:

Arg 0 *speaker*

Arg 1 *utterance*

Arg 2 *hearer*

tell.01

ditransitive (-)

The score tell you what the characters are thinking and feeling.

Arg0: The score

REL: tell

Arg2: you

Arg1: what the characters are thinking and feeling

odd ditransitive (-)¹

prepositional arg2 (-)

fronted (-)

PropBank estiloen etiketatzeko xehetasun gehiago A eranskinean daude (“PropBank-eko gidalerroak”). Eranskin hau sareko “PBguidelines.pdf”etik aterea da.

5.1.2. Verbnet

Verbnet-en web-orria honakoa da: <http://verbs.colorado.edu/~mpalmer/projects/verbnet.html>

Verbnet berez lexikoi bat da, ez corpus bat, eta Levin-en (1993) aditz-sailkapenean dago oinarrituta; horrela izanda, aditz bakoitzari Levin-en klase semantikoetako bat esleitzen zaio (9.1, 9.2, 10.1...). Hala ere, esan behar da oraindik ere sailkapen xehakatuago bat egiten dutela zenbait kasutan. Adib., 37.1 klasean bi azpisail egiten dituzte: 37.1.1 / 37.1.2. Beraz, oinarrian Levin (1993) dute, baina beren moldaketak ere egin dituzte. Horretaz gain, hautapen-murriztapenak ere ageri dira zenbaitetan: $\pm animate$, $\pm organization$, $\pm communication$, $\pm concrete$, $\pm location$, $\pm region$, $\pm animal$...

Esan dugun bezala, erabiltzen dituzten rolak orokorragoak dira (Adib.: *agent*, *theme*, *beneficiary*, *topic*, *patient*...) eta klaseak biltzeko pentsatuak dira oro har (Levin-en klaseak oinarrian bederen).

Guk gure lanean, hain zuzen, Verbnet1.0 lexikoiko rolak erabiliko ditugu, eta rol horiek PropBank-eko aditz adierei lotuta ulertuko ditugu (aditz-adiera deitzen diogu PropBank-eko sarreretan agertzen diren 01, 02, 03... horiei).

Verbnet-eko xehetasun gehiagotarako jo bedi web-orrira.

5.1.3. Propbank-Verbnet datu-basea

Direktorio honetan (cd /sx09a1/jirhizts/Corpus/Propbank/data/frames) daude, PropBank eta Verbnet-eko datu-baseak lotu ondoren, datu-base hauetan ageri diren aditzen zerrendak, dagokien informazioarekin (rolset-zenbakia edo aditz-adiera, argumentuak, rolak...).

¹ Hemen frame edo alternantzia bakarria dago azalduta (ditransitive(-)).

5.2 Euskal Aditzen Datu Basea (EADB)

EADBren diseinu eta edukia Aldezabalen (2004) tesian azaltzen da sakonki. Hemen etiketatze-lan hau egiteko beharrezkoa iruditu zaiguna azalduko dugu.

Aldezabalek bere lana Levin-en (1993) eta Vázquez *et al.*-en (2000) teoria eta irizpideak kontuan hartuz egina du oro har, eta horretaz gain, corpuseko adibideak aztergai izanda. Levin-en teoria egokituz eta Vázquez *et al.*-en lanaren eskarmentua baliatuz, euskal aditzak lantzeko proposamen bat egiten da.

Proposamen horren oinarriak laburki hauek dira:

Aditz baten rol semantikoek aditzaren “esanahia” bereizten dute, esanahi hau predikatu-mota orokor gisa ulertu behar delarik; hots, aditzak sail orokorragoetan sailkatzeko bidea emango duten predikatu orokor gisa. Beraz, esanahia ez da hiztegietan agertzen den modu xehean ulertu behar, baizik eta balio kognitibo orokor moduan (mugimenduzko balio kognitiboa, egoera-aldaketazkoa, sortzezkoa...). Izan ere:

Aditz guztien erabilerak aztertzerakoan (eta ez aditz bakarrarenak edota erabilera bakarra hainbat aditzetan) argi ageri da aditz bakoitzak egitura *usu* bat (egitura prototipiko bat) edo batzuk erakutsi ohi dituela, aditz guztietan errepikatzen diren *balio semantiko orokorren* bat adierazteko. Balio semantiko hauek ez dira adierak, hiztegietan agertzen diren modukoak, baizik eta kategoria kognitibo oinarritzkoak, predikatu-mota orokorrak, *sailak* osatzeko baliagarri izan daitezkeenak. (Aldezabal, 2008: 38).

Ikusten den bezala, egitura *usu bat* edo *batzuk* ager ditzake aditzak balio bera adierazteko. Hain zuzen, balio bera adierazten duten egitura sintaktiko desberdin guztiak, elkarren artean *alternantzia* dira. Aldiz, balio desberdina adierazten duten egiturek elkarren artean ez dituzte alternantziak osatzen, baizik eta delako aditzari dagozkion egitura sintaktiko desberdinak dira, besterik gabe. Alternantzia osatzen duten egituretako bakoitza *aldaera* izango da; hain zuzen, alternantzia baten aurrean gaudenean nahitaez bi aldaera jarri beharko ditugu harremanetan.

Balio kognitibo orokor bakoitza definitzeko beharrezkoak diren rolak erabiltzen ditu Aldezabalek eta ondoren, rol horiek modu koherente batean konbinatuta predikatu-mota orokorrak planteatzen ditu.

Hauexek dira rolak:

- gaia
- helburuko kokapena
- esperimantatzailea
- jarduera
- gai ukitua
- helburuko egoera
- kausa
- neurria
- gai sortua
- abiapuntuko kokapena
- iturria
- modua
- egoera
- bidea
- edukitzailea
- kokapena
- abiapuntua
- edukia
- denbora
- helburua
- ezaugarria

Eta hauexek predikatu-mota orokorrak:

Entitate baten egoera-aldaketa
Entitate baten kokapen-aldaketa
Entitate baten aldaketa
Entitate baten sortzea
Entitate baten jarduera
Entitate baten elkarraldatzea
Entitate bat edukitzea
Entitate bati ezaugarria esleitzea
Entitate baten izatea, existitzea
Entitate baten kokapena
Entitate baten egoera
Entitate baten deskribapena
Ustezko adierazpena

Hala, predikatu-motetan hauexek dira rol-konbinazioen aukerak; beste modu batera esanda, predikatu-mota bakoitzean atera diren patroi semantikoak (betiere aztertutako aditzekin). Maiztasun handienekoak grisez nabarmenduta ageri dira:

Egoera-aldaketa:

kausa+gai ukitua (*hil*)
kausa+gai ukitua+gai ukitua (*banandu*)
gai ukitua (*irakin*)
kausa+gai ukitua+helburuko (*egokitu*)

Kokapen-aldaketa:

kausa+gai ukitua+abiapuntuko kokapena / bidea+helburuko kokapena (*ekarri*)
gai ukitua+abiapuntuko kokapena / bidea+helburuko kokapena (*joan*)
kausa+gai ukitua+helburuko kokapena (*jarri*)
gai ukitua+helburuko kokapena (*bisitatu*)
kausa+gai ukitua+abiapuntuko kokapena (*hartu*)

Aldaketa:

Kausa+gai ukitua+abiapuntua/bidea+helburua (*aldatu*)

Entitate baten sortzea:

iturria+gai sortua (*isuri*)
iturria+gai sortua+kokapena (*eragin*)

Entitate baten jarduera:

esperimentatzailea+gaia (*abestu*)
esperimentatzailea+esperimentatzailea (*besarkatu*)
esperimentatzailea+jarduera (*aritu*)
esperimentatzailea+modua (*jokatu*)
esperimentatzailea+jarduera+denbora (*eman*)

esperimentatzailea+jarduera+helburua (*deitu*)

esperimentatzailea+gaia+gaia (*aldatu*)

Entitate baten elkarraldatzea:

abiapuntua+gaia+helburua (*eman*)

Entitate bati ezaugarria esleitzea:

abiapuntua+ezaugarria+helburua (*esan*)

Entitate bat edukitzea:

edukitzailea+edukia (*izan*)

edukitzailea+edukia+kokapena (*ekarri*)

Entitate baten izatea, existitzea:

gaia (*existitu*)

Entitate baten kokapena:

gaia+kokapena (*egon*)

Entitate baten egoera:

gaia+egoera (*egon*)

Entitate baten deskribapena:

gaia+ezaugarria (*izan*)

Ustezko adierazpena:

gaia (*eman*)

Ikusten dugunez, predikatu-mota bakoitza hainbat patroi semantikoz gauza daiteke, baina delako predikatu-motakoa izateko delako gutxieneko osagai semantiko bat(zuk) izan behar ditu(zte):

- Egoera-aldaketetan *gai ukitua* izatea da nahitaezkoa:

Adib.: *hautsi, hazi, hil, isildu, kezkatu, haserretu, egokitu, aldatu, pasatu, igo, jaitsi...*

- Triste al hago [HURA] **hil** zelako?- galdetzen dio Juttak. (EusCor)

- Kokapen-aldaketetan, *gai ukitua* eta *osagai lokatiboren bat*:

Adib.: *kokatu, zintzilikatu, sartu, iritsi, ekarri, eraman, etorri, joan, jarri...*

- Non jarri da argazkilaria, gozotegiaren barruan ala kanpoan? (EusCor)

- Sortzezkoetan, *gai sortua* eta *iturria*²:

Adib.: *egin, etorri, ekarri, atera, jarri...*

- Horregatik hilekoa duen emakumeari maionesak ez zaizkio **ateratzen**. (EusCor)

- Edukitzeetan, *edukitzailea* eta *edukia*:

Adib.: *izan, ekarri, etorri, jaso, eraman, hartu...*

- Problema gutxiago izan zuen zentsura frankistak. (EusCor)

² Hemen salbuespena da *loratu* aditza; sortzekoa izanik ez duelako *iturria* behar.

- Bigarren liburu honetan badatoz, gainera, aurrekoaren zuzenketak. (EusCor)

- Eta jardueretan, azkenik, *esperimentatzailea* eta, normalean, beste osagaien bat:

Adib: *erabili, abestu, afaldu, ehizatu, grabatu, jan, ikusi, asmatu, oroitu, ahaztu...*

- Eskolan **bazkaltzen** du [HARK] eta gehien gustatzen zaion janaria saltxitxak dira. (EusCor)

Hala izanik, *entitate baten aldaketa*, entitate baten *egoera- edo kokapen-aldaketa* ere bada (*gai ukitu* bat dagoelako). Gertatzen dena da *aldaketatzat* jo dugun predikatu motan aditzak ez duela berez egoera edo kokapenaren informazioa kodetuta; informazio hori sintagmek hartzen duten izenaren baitan dago. Adib.:

- Karga negatibo, edo elektroiak, gorputz batetik bestera aldatzen direnean korrante elektrikoa sortzen dela esaten dugu. (EusCor)

gorputza: leku fisikoa -> kokapen aldaketa

- Izan ere, autonomi edo probintzia-mailara aldatu nahi baditugu, zatikaturiko inkestak ez dira lehen bezain adierazgarriak. (EusCor)

probintzia maila: ezaugarria -> egoera-aldaketa

Aditzek batzuetan adiera bakarra adierazten dute, eta horri loturiko alternantziak izan ditzakete; beste batzuek ez dute alternantziarik. Ondorioz, aditz batek adiera bakarra duenean, mota semantiko horretakotzat jotzen da (adib. *jan: jarduerazko* aditza). Ordea, aditz batzuek adiera bat baino gehiago adieraz dezakete, eta horiei loturiko alternantziak ere izan ditzakete. Horrelakoetan sortzen dira guztiak elkarren artean bereizteko zailtasunak (eta beharrak,aldi berean); eta aditz batzuk mota semantiko batekotzat jotzen badira ere (adib. *eman: elkarraldatzezko* aditza; *ekarri: kokapen-aldaketazkoa*), bestelako predikatu motetan ere parte hartzen dute (adib. *eman: jarduerazko* predikatu motan; *ekarri: sortzezko* eta *edutezko* predikatu motan). Horregatik da bidezkoagoa *predikatu-motez* hitz egitea aditz-motez hitz egitea baino.

EADBren xehetasunak B eranskinean daude. Landutako 100 aditzen azterketa ikusi nahi izanez gero, jo bedi Aldezabalen tesi-laneko (2004) C eranskinera. Wikian ere zintzilikatuta dago.

5.3 Levinen aditz guztiak beren euskal ordainekin

1998an, Marylan-eko Unibertsitateko egonaldian, Aldezabalek Levin-en liburuko aditz guztiei euskal ordainak jarri zizkien, horretarako jokaera sintaktikoari emanez garrantzia. Hots, ez zen itzulpen hutsa egin, baizik klase horretarako egitura sintaktiko antzekoen zuen/zuten aditzen ordainak hautatu ziren. Aditz-zerrenda eta ordainak ikusi nahi izanez gero, Rolen wikian daude (<http://ixa2.si.ehu.es/rolak/index.php/Trudi:C> ingelesezko aditza_Levin-klasea_euskarazko aditzak.doc).

5.4 Aditz baliokidetza horiek PropBankekin gurutzatuta eta corpusarekin lotuta

Levin-en aditzen itzulpen guztiak hartu eta itzulpen horietatik EPECen ageri diren aditzetara lotu dira. Jakina, lotura horretan PropBank-ekin egindako ebakidura hartu da kontuan. Hots: Levin-klasearen arabera mapatze egoki bat izan duten aditzen PropBank-informazioa (esaterako, *esani* tell 37.1.1 / 37.2 eta say 37.7ren PropBank-informazioa esleitu zaio).

Aldi berean, mapatze horren ondorioz agertu diren Vebnet-rol guztien zerrenda atera da. Honekin, aurrerantzean etiketatzeko baliatuko diren rolen zerrenda nahiko osoa lortu da.

PropBank-ekin egindako mapatze horretan honako kasuistika agertu zaigu:

EQUAL: Aldezabalen itzulpenen zerrendako aditzak PropBank-eko datuetan zegoen Levin-klase bera du. Aukera hau kasuen % 51n gertatzen da.

SUBCLASS: Aldezabalen itzulpenen zerrendako aditzak klase orokorrago bat du PropBank-eko datuekin alderatuta: PropBank-en azpiklase bat ageri da (adibidez, itzulpenean: esan 37.1, eta PropBank-en 37.1.1). (% 6)

CHANGED: Klase bakarra badago PropBank-ekin bat ez datorrena eta mapaketan eta PropBank-en klase-kopurua berdina bada, orduan mapaketako klasea PropBank-en dagoenarekin ordeztu dugu. (% 2)

MISSING: Aditza ez dago PropBank-en edo ez dauka Levin-klaserik. (% 41)

Kasuistika honen aurrean, guk printzipioz *equal* eta *subclass* motakoak baliatu ditugu EPECekin lotura egiteko; bestek errebisatzeko geratzen dira mementoz.

EADB eta corpusen agertzen diren aditzen arteko konparaketa eta maiztasunei buruzko datu batzuk Wikiko dokumentu honetan ikus daitezke: <http://ixa2.si.ehu.es/rolak/index.php/Irudi:EADBvsEPEC.xsl>

5.5 Aditzaren mende dagoen dependentzia orori jarriko zaion etiketa: *arg_info*

Corpusen rolei buruzko informazioa txertatzeko, “arg_info” etiketa erabiliko da, eta etiketa honek dituen eremuak honako bostak dira::

- **VNrs** (“Verbnet-PropBank rolset zenbakia”): PropBank/Verbnet-eko rolset zenbakiarentzako eremua. Gehienetan aditz bat baino gehiago dagoenez PropBank-ekiko ebakiduran, aditza ere zehaztea erabaki da. Adib.: tell_01/say_01, 02, 03, 04, 05, 06, 07...
- **BAL** (balentzia): argumentu edo adjuntuentzako eremua. Balioak ondokoak dira: arg0, arg1, arg2, arg3, arg4, argmod.
- **VNrol** (“Verbnet rola”): Verbnet-i dagokion rola. Balioak, printzipioz, izango liriteke 5.4 atalean aipatu dugun rolen zerrendan agertzen direnak.
- **EADBrol** (“Euskal Aditzen Datu Baseko rola”): EADBko rola ezartzeko eremua. Balioak ondorengo 21 rolak dira: *gaia, gai ukitua, gai sortua, egoera, kokapena, denbora, helburuko kokapena, helburuko egoera, abiapuntuko kokapena, bidea, abiapuntua, helburua, esperimantatzailea, kausa, iturria, edukitzailea, edukia, ezaugarria, jarduera, neurria, modua*.
- **HM** (“Hautapen-murritzapenak”): Hautapen-murritzapenak jartzeko eremua. Balioak, printzipioz, ondorengoak dira: [+biz], [-biz], [+konkr], [-konkr], [+giz], [-giz].

Nahiz eta rolen etiketatzearen oinarria sintaktikoki etiketatutako corpusa izan, *arg_info* etiketa independenteki erabili ahal izatea nahi dugu. Horregatik, aipatutako 5 eremuez gain, beste bi eremu gehiago zehaztuko dira etiketan: lantzen ari garen aditza (burua, alegia) eta aztertzen ari garen argumentu/adjuntua, bigarren eta hirugarren eremu izango direnak hurrenez hurren. Hala ere, bi balio hauek automatikoki kopiatuko dira etiketatze sintaktikotik, lan errepikakorra ekidinez. Horrela, halaxe geratuko zaigu *arg_ifo* etiketa::

arg_info: (VNrs, aditza, hitz aztergaia, BAL, VNrol, EADBrol, HM)

Adibide hipotetiko batez:

arg_info: (tell_01/say_01, esan, nik, Arg0, agent, esperimentatzailea, [+giz])

Esan bezala, etiketa hau 3. atalean adierazitako dependentzia-etiketa bakoitzaren ondoren erantsiko da.

Adib.: Hau esanik, gartzelak, bi modutan sailaktuko genituzke.

ncobj (abs, esanik[w305], Hau[w304], Hau[w304], obj)
arg_info (tell_01/say_01, esanik[w305], Hau[w304], arg1, topic, gaia, -konkr)

xmod (mod, sailkatuko[w311], esanik[w305], esanik[w305])
arg_info (classify_01, sailkatuko[w311], esanik[w305], argM, MNR, modua, -)

ncobj (abs, sailkatuko[w311], gartzelak[w307], gartzelak[w307], obj)
arg_info (classify_01, sailkatuko[w311], gartzelak[w307], arg1, Theme, gai, -biz)

detmod (-, modutan[w310], bi[w309])
ncmod (ine, sailkatuko[w311], modutan[w310], modutan[w310])
arg_info (classify_01, sailkatuko[w311], modutan[w310], argM, MNR, modua, -)

auxmod (-, sailkatuko[w311], genituzke[w312])

5.6 EADBn landuta dauden 100 aditzen lotura automatikoa kasu/rol mailan

EADBko 100 aditzak hartuko dira eta aditz bakoitzaren mende dauden osagaiak EADBn definitutako rolak esleituko zaizkie automatikoki. Horretarako osagai bakoitzak sintaktikoki gauzatzeko baliatzen duen postposizio-atzizkia hartuko da oinarritzat.

Horrela, aditz jakin batean postposizio-atzizkiak edo bere mendeko osagaiak anbiguoak ez direnean, postposizio-atzizki edo osagai horiei dagozkien ezaugarri semantikoak automatikoki esleituko zaizkie. Ondorioz, aditz horren osagaiak automatikoki etiketatuko eta desanbiguatuko dira.

Esaterako, ondoko adibideko “ez” osagaiari analisi bakarra esleituko zaio:

Adib.: Hau guztia osatu gabe geratuko litzateke (...) aldaketekin joango ez balitz.

(...)
ncmod (neg, joango[w2839], ez[w2840], ez[w2840]
arg_info (VNrs, joango[w2839], ez[w2840], argM, NEG, -, -)
(...)

Gainerakoak eskuz desanbiguatu beharko dira. Halakoetan, anbigua den rol bakoitzeko etiketatze-lerro bat sortuko da, eta beste eremuak eskuz osatu edo etiketatuko dira.

Esaterako, ondoko adibideko “andereñoa” hitzaren postposizio-atzizkiak anbiguitasuna sortzen duenez, postposizio-atzizki horri dagozkion ezaugarri semantiko desberdinak sortuko dira:

Adib.: Mitchell andereñoa Perth-era joan zenean, (...)

(...)
ncsubj (abs, joan[w1915], andereñoa[w1917], andereñoa[w1917], subj)

arg_info (VNrs, joan[w1915], andereñoa[w1917], BAL, VNrol, gai_ukitua, -)
arg_info (VNrs, joan[w1915], andereñoa[w1917], BAL, VNrol, edukia, -biz/-konkr)
arg_info (VNrs, joan[w1915], andereñoa[w1917], BAL, VNrol, ezaugarria, -)

(...)

6. Metodologia

Etiketatzelaian hau 3 fasetan gauzatuko da:

1. fasea: proba-fasea

Hiru hizkuntzalarik aditz beraren esaldi berak (50) etiketatuko dituzte tresna probatzeko eta dudak argitzeko (adib. *esan*). Horrekin eskuliburua osatzen joango gara.

2. fasea: Corpuseko aditz guztien eskuzko etiketatzea

a) EADBko aditzetatik hasiko gara (100 aditz).

b) Bakoitzetik 60 agerpen etiketatuko dira.

Etiketatzailerik bakoitzak aditz beraren agerpen desberdinak etiketatuko ditu. Hots, 20 etiketatzailerik batek, 20 beste etiketatzailerik eta beste 20 hirugarren etiketatzailerik. Printzipioz, 10 EEBS corpusekoak eta beste 10 Egunkaria corpusekoak³.

Agerpen gutxiago egonez gero, dauden agerpen guztiak etiketatuko dira.

³ Oso komenigarria litzateke, bururatzen zaizkigun ondorioak apuntatzen joatea (adib. *Joan* aditzarekin 20 agerpen gutxi egingen dira (edo ez))

Hau eginda, gutxi gorabehera 3.200 esaldi etiketatua izango ditugu. Batez beste esaldiko 2 aditz daudela emanda, 1.600 esaldi desberdin izango dira.

Bi motako rolak etiketatzeak lan errepikakorra izateaz gain, alferrikako lana dela ere irudi dezake, baina hasteko 100 aditzekin baino ez da hori egingo. Horrela lagin hau baliatu ahal izango dugu ondorio orokorrak ateratzeko: ProBank-Verbnet-en hartutako erabakiak rol horiek (eta ez beste batzuk) erabiltzeko, adierei, argumentuei, adjuntuei eta abarrei buruzko iritzi desberdinak... Azken batean, ondorioak ateratzea izango dugu helburu eginkizun hau definitzerakoan adierazi bezala.

Lehen fase hau bukatutakoan eta esaldi bakoitza hamabost bat hitzez osatuta dagoela kontuan harturik, 24.000 hitz etiketatuta izango genituzke.

- c) Gainontzeko aditzekin hasiko gara. Hauetan ere 60 agerpen hartuko dira.
- d) Hori bukatutakoan, corpusetik aditz bakoitzeko lexikoa automatikoki erauziko da, ondorengo informazioa kontuan hartuta:

Patroi desberdin gutiak, bakoitza ezaugarritzen duten ondoko eremu hauek kontuan hartuta: aditza, WNrs, BAL, WNrol_(+KAS), HM.

Horrela, aditzak euren argumentu-ezaugarriekin erauziko dira eta ondorioz, PropBank/Verbnet-en egiten den modura, aditz bakoitzeko sarrera lexiko bat lortuko da, euskarako rol bakoitzari dagokion postposizio-atzikia ere adieraziko da eta horrenbestez posible diren egitura sintaktikoak ere bai.

3. fasea: Aditz guztien gainerako agerpen guztien etiketatze automatikoa eta eskuzko errebisio-lana

Aurreko fasetik ateratako lexikoekin, antzeko ezaugarriak dituzten corpuseko aditzak inferitzeko ahalegina egingo da. Egin dira halako lanak Ixa ikerketa-taldean; Aldezabal *et al.*-en (2001) postposizio-atzikien konbinaketan oinarrituta egin zen eta Zafirain *et al.*-en (2008) ingeleserako etiketatutako roletan oinarrituta. Horrela, aditz berri bakoitzeko postposizio-kasuen eta rolen proposamena egingo da.

Ondoren eskuz gainbegiratu da. Horrekin zenbatean asmatzen duen analisi automatikoak aztertuko da.

7. Irizpideak

Etiketatzeko-lana gauzatu ahala hartu diren erabakiak dira jarraian datozenak:

7.1. Irizpide orokorrak

- *arg_info* etiketako lehen eremurako **PropBank-eko baliokideak hautatzeko garaian**, tokatzen zaigun esaldian baliokide bat (edo batzuk) jarriko ditugu gauzak argi daudenean. Zalantzak sortzen zaizkigunean, ordea, ja gehiago begiratu eta aztertu beharko dugu.

Adibidez, *esan* tell_01 eta say_01 da argi eta garbi adiera batean. Aldiz, *joanen* kasuan, askotan go_01 da, baina testuinguruaren arabera, besteren bat egokiagoa izan daiteke (*joanek* baliokide asko baititu lotuta).

- Era berean, **lehenengo eremuan** klaseari dagokion zenbakiaz gain, aditza ere jarri behar da (adib. (go_01,...) (tell_01/say_01,...; eslaxa erabiliko dugu aditzak banatzeko, baina ez dugu zuriunerik utziko). Bi aditz jarri behar ditugunean aditz horiek PropBank-en ageri diren hurrenkeran adieraziko ditugu (end_01/finish_01) eta ez (finish_01/end_finish). Aditza *phrasal verb* bat baldin bada, azpimarra erabiliko dugu aditza etiketaren lehen eremuan idazterakoan, adibidez: *take_out_01, come_out_01...*

- **Zer egin PropBank ereduarekin bat ez gatozenez?**

Lehenik eta behin ezadostasuna ondo arrazoituta egon behar da, eta gero, eskuliburuan zehaztu behar da, horretarako, irizpideetan espresuki sorturiko azpiatalean (7.3. PropBank-ekiko ezadostasunak) jarriz.

- **Zer egin aditza EADBN dagoenean, baina ez guk nahi dugun balioa?**

Horrelakoetan, balio hori nola landu baldin badakigu, landu egingo dugu, horretarako rol edo/eta predikatu berriak sortu behar badira sortuz . Ez badakigu, “-“ marka erabili berari dagozkion eremu guztietan.

- **Zer egin aditza ez dagoenean EADBN?**

Horrelakoetan, baldin eta gai bagara EADBN jarraitzen den ereduaren arabera lantzeko, landu. Ez badakigu zer jarri “-“ marka erabili berari dagozkion eremu guztietan.

- **Zer egin aditz baten informazio osoa topatzen ez dugunean?**

Horrelako kasuetan, iturri desberdinetara jo beharko dugu informazio horren bila. Adibide batekin hobe ulertuko dugu prozesua zein izan daitekeen. *Dirua jokatzan dizut* adibidea agertu zaigu, kasu horretan “jokatu” “apustu egin”-en parekoa da; beraz, ingelesez “to bet” izango litzateke, baina ingelesezko aditz hori ez dugu “jokatu”-ren ordainen artean; beraz, Verb-Index (<http://verbs.colorado.edu/verb-index/index.php>) orrialdera joko dugu aditz horren argumentu-egituraren eta rolen bila. Kasu honetan, PropBank-ek argumentu-egitura ematen digu, baina hor ez ditugu VerbNet-eko rolak ikusten:

Roleset id: *bet.01* , *wager*, vncls: 54.5 94

Roles:

Arg0: *better*

Arg1: *amount of bet*

Arg2: *basis, proposition, bet-on*

Arg3: *co-better*

Beraz, hemen argumentuak ditugu, baina orain rolak falta zaizkigu. Verb-Index orrialdean ikusten dugu “bet” aditza beste aditz batzuekin parekatzen dela eta beraz aditz horien rol berberak izango dituela. Verb-Index-en “bet” aditzaren pareko agertzen dira “bill, rely, risk” eta hauen rolak ikusita “bill” horrenak gure “apustu egin” aditzarekin bat datozela ikusten dugu:

Agent: [+animate / + organization]
Asset: [+currency]
Recipient: [+animate / +organization]
Cause:

Beraz, azken pausu gisa egingo duguna goiko argumentuak beheko rolekin bateratzea izango da:

Arg0: Agent
Arg1: Asset
Arg2: theme
Arg3: recipient

- **Arazoak ordainak aukeratzean**

Lan honetan aurkitzen dugun zailtasunik handienetariko bat euskarazko aditzari ingelesezko zein baliokide esleitu behar diogun aukeratzea da, zaila baita hain desberdinak diren bi hizkuntza parekatzea.

Esaterako, "aurkitu" aditzaren kasuan, 2 adiera edo balio semantiko bereizten ditugu:

- entitate baten _kokapena
- entitate baten _egoera

Ingelesezko ordainen artean "find" daukagu eta argi dago lehen adiera edo balio semantikoarentzat ezin hobeto datorrela ordain hau, baina zer gertatzen da bigarren adierarekin?

Ikus ditzagun bi adibide:

- Zer moduz aurkitu zara proban zehar?
- Jokalaria egoera berezian aurkitzen baitzen.

Bi adibide hauetan badirudi ingelesez egokiago litzatekeela "to feel" edo "to be" jartzea, baina orduan, gehiegi aldenduko ginateke gure hasierako aditzetik (aurkitu / find). Kontua da horrelako testuinguruetan "aurkitu" aditza erabiltzea gaztelerazko "no se encuentra bien" horren kalkoa dela; beraz, ingelesez ere aditzaren itzulpen literala mantentzea erabaki dugu.

- **Zer egin Verbnet-ek argumentu berari bi rol desberdin esleitzen dizkionenean?**

Ikusi aditzaren kasuan gertatu zaigu hau. EADBN *ikusi* aditzak bi argumentu ditu eta argumentu horietako bakoitzari rol bat esleitzen zaio:

Arg0: Esperimentatzailea
Arg1: Gaia

Baina PropBank eta Verbnet-era jotzerakoan, argumentu horientzat bi rol ematen dituztela ikusten dugu:

ID: see.01 NAME: view VERBNET: 29.2 30.1
ROLE(0): viewer VNTHETA (29.2): **Agent** VNTHETA (30.1): **Experiencer**
ROLE(1): thing viewed VNTHETA (29.2): **Theme** VNTHETA (30.1): **Stimulus**

Hau da, Arg0-ri "agent" eta "experiencer" jartzen dizkio eta Arg1-i "theme" eta "stimulus".

Kasu honetan, EADBN oinarritzea erabaki dugu eta beraz, horrekin bat datozen Verbnet-eko rolak jartzea, hau da, “agent” (EADBN ergatibo guztiak “esperimentatzaile” dira eta ez dago agenterik) eta “theme”. Honela geratuko litzateke:

Arg0: Agent, esperimentatzailea
Arg1: theme, gaia

- **Eremuren batean ez dakigunean zein balio jarri** edo berez ez badagokio inolako baliorik marratxoa (-) jarriko dugu.
- **Aditz arazleetan** sintagma ergatiboa **Arg1** da. Horrek esan nahi du kasu hauetan Arg0, Arg1, eta abar ez direla ohiko modura agertuko.

Adib.: Txominek Patxiri esanarazi dio.

Esaldi honetan, Arg1 “Txominek” litzateke eta Arg0 “Patxiri”. Eta biak “Agent” lirateke.

- **Inpertsoneletan** (adib. “beti hori esaten da”), Arg1 izango dugu (eta ez arg0). Pertsoneletan, hau da, iragankorretan gehienetan, arg0 (sintagma ergatiboa) eta arg1 (sintagma absolutiboa) izango ditugu.
- **Hautapen-murriztapenei** dagokienez, elementu baten hautapen-murriztapena argia bada, jarri egingo dugu.

Adib.: Idazkariak esan dio

Esaldi honetan, *idazkaria* [+giz] da.

Zalantzak ditugunean, edo ez jarri ezer (“-“) edo zer edo zer orokorra jarri ([-biz]).

Zenbait kasutan, “kokapena”z ari garenean, esaterako, interesgarria da kokapen hori biziduna, bizigabea, konkretua... edo zer den jakitea. Horrelakoetan, ahal dela, hautapen-murriztapena jarri egingo dugu.

- **Ohar berriak**: zenbait aditzen analisisan, jarraian beltzago nabarmendu eta zerrendatu ditugun oharretako bat edo batzuk adierazi ditugu. Oharra adierazteaz gain, oharraren ondotik zuriunea utzi, eta parentesi artean aztergai dugun aditza zein den zehaztu behar da.

Adib.: Egoera zein den ikusten joango gara.

Oharra: aditz-perifraasi aspektuala (joan)

Bi ohar batera jarri behar ditugunean, eslaxa erabiliko dugu bi aldetan zuriunea utziz eta ohar bakoitzaren ondotik aztergai dugun aditza adieraziz.

Adib.: Desarmaa aurrera badoa...

Oharra: d / aditz-esapidea (joan)

Hona hemen oharren zerrenda:

- **aditz-esapidea (aztergai dugun aditza):** *nesketara joan, usotara joan da*. Arg_info ez bete eta ez ezabatu, eta gorde OSOA karpetan. Lasai, hau oso subjektiboa da-eta.
- **aditz-perifraasi aspektuala (aztergai dugun aditza):** *egoera zein den [ikusten joango gara], [hasiz joan zen]*. Arg_info ez bete eta ez ezabatu, eta gorde OSOA karpetan.
- **argumenturik gabea (aztergai dugun aditza):** Zenbaitetan aditzak ez du argumentu-egiturarik aztertzen ari garen esaldian eta, beraz, ezin diogu argumentu edo rolik esleitu. Horrelakoetan, hutsik utziko dugu eta “osoa” karpetan gordeko. Horretaz gain, oharren eremuan “argumenturik gabea” oharra* jarriko dugu. Adibidez: *ikusten dugunez / joan den astean*.

*Ohar hau 2009-10-20tik aurrerako analisietan txertatu dugu; beraz, data honen aurreko aditzetako zenbait analisitan oharra falta daiteke. Data hori baino lehenago aztertutako aditzak dira: esan, joan, atera, jokatu, ahaztu, konturatu, eman (Ainararen fitxategiak), entzun, hil, aldatu, ikusi eta lortu.

- **balio berria:** ikusi 7.3. ataleko lehen puntuan zer diogun.
- **ordain berria:** ikusi 7.3. ataleko bigarren puntuan zer diogun.

Jarraian datozen ohar hauen ondotik ez da aztergai dugun aditza adieraziko; hau da, oharra baino ez da adieraziko:

- **d: dependentziak gaizki** daudela iruditzen bazaigu, aldatu egingo ditugu*. Hori adierazteko, oharretan “d” idatziko dugu. Hala ere, EPEC-DEPen emandako irizpideak jarraitzen ez dituztenak bakarrik aldatuko ditugu. Adibidez, *behar izan, nahi izan* eta horrelako aditzekin, erabakia izan zen horiekin batera agertzen zen aditza mendekotzat hartzea (“egin behar da”), edozein delarik ere ordena. Aldiz, “Liburuak ekarri nahi ditu” moduko esaldietan “ekarri”ren ezker aldean dauden elementuek duten lotura sintaktikoa (gunea) errespetatu egingo dugu. Horrelako kasu askotan aditzetako bat (“ekarri” kasu honetan) ezer gabe geratzen zaigu. Hori bere horretan utziko dugu.

*Gaizki etiketatuta dauden dependentziak aldatuko ditugu baldin eta lan handia eskatzen ez badute; alegia, dependentzia-zuhaitz osoa aldatu behar bada eta mementoz ez zuzentzea erabakitzen bada, esaldi horiek “zalantzazko karpetan” gordeko ditugu.

- **d_estilo zuzena:** ikusi zer diogun 7.1 ataleko hamaseigarren puntuan.
- **erabilera desegokia:** Adibideren batean arauetatik urruntzen den erabilerearen bat topatzen badugu, etiketatzen saiatuko gara, baina oharretan “erabilera desegokia” jarriko dugu. Adibidez, *erruletara jokatu* agertu zaigu eta badakigu erabilera desegokia edo okerra dela, *erruletan jokatu* beharko lukeelako, beraz, *erruletara* horri dagokion rola jarriko diogu, kasu honetan “jarduera” eta oharren eremuan “erabilera desegokia” iruzkina jarriko dugu.

- Zenbait esaldi, **estilo zuzena dela eta, ez dira sintaktikoki etiketatu**. Rolak etiketatzerakoan erabaki dugu horiek lantzea eta *arg_info* egiturako hirugarren eremuan puntuazio-ikurra jartzea (gunearekin lotura egiten duen elementu gisa). Kontuan izan behar da horrelako asko AbarHitzek aparteko esaldi gisa tratatzen dituela (esaldi-mugak horrelaxe definituta ditugulako); beraz, behin landuta gehienak “osorik” karpetan sartzekoak izango dira. Estilo zuzenean askotan komatxoak eta koma aldi berean ageri dira; horrelakoetan gune gisa komatxoak hartzea erabaki dugu.

Zenbaitetan estilo zuzena landuta dator dependentzietan (irizpideak aldatu egin zirelako momentu jakin batean), eta analisi okerra ematen da. Horrelakoetan, oharretan “d_estilo zuzena” jarriko dugu eta dependentzia-analisia ezabatu egingo dugu. Marka horren bidez horiek gerora “Estilo zuzena” karpetara pasa ahal izango dira.

7.2. Irizpide partikularrak eta adibideak

- **“Horretaz gain”** -> argM ADV
- **ArgM-etan noiz adierazi behar da ADV?** Hitza adverbio kategoriakoa demean ADV erabakiko da, eta gainerako kasuetan dagokion ezaugarriaren arabera: TMP, MNR...
- **“Oporretara joan”** eta horrelakoak argM: PRP (EADB rolean “-“).
- **“Ez da inoiz etorriko”**. “Ez”: ArgM:NEG, “inoiz” ArgM:ADV
- **“Nik neuk”**: dependentziei jarraitu (bi subjektu badaude, bi *arg_info* berdin; bestela, bakarra)
- **Adjuntuak** direnean, Aldezabalek bere tesi-lanean (2004) zehazten duen rolik ez dugu jarriko (marratxo erabiliko dugu)
- **Koordinazioan** zilegi da bi Arg0 edo Arg1 edo dena delakoak egotea.
- **Bokatiboak ArgM, DIS, -, - markatuko ditugu PropBank-eko gidalerroei jarraiki (ik. A Eranskinean 3.10 puntua). Adibidea: *Ama, eraman nazazu zurekin.***

Ama: ArgM, DIS, -, -

- **Perpau kontzesiboak, baldintzazko perpauak:** ArgM gisa etiketatuko ditugu eta ondorengo eremuak hutsik utziko ditugu (ArgM, -, -, -).

Adibidea: *ikusi nuen, nahiz eta ez nuen agurtu* (asmatua)

arg_info (see_01, ikusi, agurtu, ArgM, -, -, -)

Adibidea: Hala ez balitz, eskutitzok nirea bezain garai ero eta hutsalean agertuko balira, orduan zera eskatzen diot kutxa aurkitu duenari: bere lekuan utzi dezala berriz kutxatxo.

arg_info (ask_02, eskatzen, punt_koma, ArgM, -, -, -)

- **Adibideak**

Directionals (DIR)

arg_info(see_01, ikusten-[w59], derbirako-[w63], argM, DIR, -, -)

Temporal markers (TMP)

arg_info (see_01, ikusiko-[w297], Hortik aurrera-[pos7], argM, TMP, -, -)

ncpred-ak (*Jontxu alai ikustean...*)

arg_info (see_01, ikustean, alai, argM, MNR, -, -)

7.3. Propbank-ekiko ezadostasunak

- **Zer egin aditz batek PropBank-eko aditz bat baino gehiago dituenean lotuta?**

Lehenik eta behin, konprobatu behar dugu ingelesezko aditz horiei dagozkien rolset-zenbakia, rolak eta argumentuak berdinak ala desberdinak diren.

Berdinak badira, ez dago inolako arazorik: dagozkionak jarri eta kito (euskal aditzaren eta ingelesezkoaren arteko lotura egina dago zerrendan).

Gerta daiteke rolak eta argumentuak berdinak izanik, rolset-zenbakia desberdina izatea. Horrelakoetan, gorago azaldu dugunaren ildotik, bi aditzak zehaztuko dira lehenengo eremuan (adibide asmatua: take_04/bring_01). Rolak berdinak, baina argumentu-kopurua eta -zenbakiak bat ez datozenean, orduan desberdintzat hartuko ditugu, eta segidako paragrafoan adierazten duguna egingo dugu.

Desberdinak badira, aztertu beharko dugu zein den gure ustez egokiena. Horretarako oso kontuan izan beharko da ingelesezko aditzaren eta euskarazkoaren jokaera sintaktikoa, argumentu-kopuruari eta alternantziei dagokionez. Ezaugarri hauek antzekoen dituen aditza hautatu beharko da. Biak balizkoak dira eta beraz lehenengo eremuan zehaztu egin beharko da.

Zeregin honetan, *Ancora* corpora kontsultatzea ere lagungarri izan daiteke (<http://clic.ub.edu/ancora/>).

- **Zer egin aditz batek ez duenean loturarik inongo PropBank aditzekin edo ez gatozenean bat dituen loturekin?**

Horrelakoetan, kontsulta-prozesua gehiago luzatzen da; izan ere, lehenik eta behin, ingelesezko ordaina zein izan daitekeen pentsatu behar dugu (hori Wordreference-n eta Elhuyar-en on-line hiztegian begiratzuz atera daiteke: www.wordreference.com, <http://www.elhuyar.org/hizkuntza-zerbitzuak/EU/Hiztegi-kontsulta>), eta gero ordain hori/horiek PropBank-eko direktorioan begiratu. Hortik gure ustez egokiena izan daitekeen aditzaren rolset-zenbakia (01, 02...) aukeratuko dugu lehenengo eremurako, zenbakiarekin batera aditza zehaztuz (adib. have_03). Aditz bat baino gehiago jar dezakegu eslaxarekin bereizita (adib. have_03/own_01).

- PropBanken badira **rol batzuk ez direnak gure ustez argi bereizten**. Adibidez, **Location** eta **LOC**. Gure erabakia hauxe izan da: elementua argumentu izaerakoa bada (Arg1, Arg2...), “location” erabiliko dugu. Adjuntua bada (ArgM), LOC. Hots, hiru letrakoak adjuntuentzat utziko ditugu eta besteak argumentuentzako.

- **Ordain berria:** Ez dugunean PropBank balio egokirik topatzen, oharretan *ordain berria* jarriko dugu eta aditz horren adiera arruntenarekin lotuko dugu (normalean). Adibidez, “eskatu” aditzean ez dugu PropBank-ek ematen dituen ordainen artean egokirik aurkitu (*claim* eta *order*) eta iruditu zaigu kasu gehienetan ordain egokia *ask_02* dela; beraz, hau erabakita, oharretan *ordain berria* jarri dugu.

C eranskinean adierazi ditugu aditzaren ordain berriak.

- **Balio berria:** Ingelesezko aditzaren argumentu-egitura eta EADBarena erabat bat ez datozenean, oharretan *balio berria* jarriko dugu. Hots, Aldezabal-en tesi-lanean (2004) ageri den informazioa osatu edota aldatu beharrean aurkitzen garenean gertatuko zaigu hau.

Adibidez, aurreko atalean aipatu dugu “eskatu” aditzaren ordain berria den *Ask_02*-k dituen argumentuak eta rolak begiratu ditugu PropBank eta VerbNet-en eta 3 argumentu dituela ikusi dugu:

Arg0: Agent
Arg1: Theme (proposition)
Arg2: Patient

EADBN bi argumentu baino ez zaizkio esleitzen *eskatu* aditzari:

Arg0: esperimantatzailea
Arg1: gaia

Aipatzen da DAT ager daitekeela, baina ez da kasu azpikategorizatuen artean sartzen. *Ask_02* aditzak, ordea, sartu egiten du kasu azpikategorizatuen artean eta guk PropBank-en ereduari jarraitzea erabaki dugu; beraz, DAT agertzen zaigunean Arg2 gisa etiketatuko dugu:

Arg2: patient, abiapuntua.

Adib.: Nemesiok, joan baino lehen, Alejandro adiskideari eskatzen dio, zaindu dezala bere “x” zakurra.

arg_info (ask_02, eskatzen, Nemesiok, arg0, Agent, esperimantatzailea, +giz.)
arg_info (ask_02, eskatzen, lehen, argM, TMP, -, -)
arg_info (ask_02, eskatzen, adiskideari, arg2, patient, abiapuntua, +giz.)
arg_info (ask_02, eskatzen, zaindu, arg1, Theme, gaia, -biz.)

Hau dela eta, EADBN dagoen egiturari balio berri bat erantsi diogu, DAT elementua kasu azpikategorizatua gisa eskatzen duena, hain zuzen ere, eta beraz, **oharretan balio berria** jarriko dugu.

Baina ez dugu kasu guztietan berdin jokatu eta adibidez, *lortu* aditzaren kasuan, PropBank-ek Arg2 gisa ematen zuena, guretzat ArgM izango dela erabaki dugu. *Lortu* aditzak bi argumentu ditu EADBN:

Arg0: esperimentatzailea
Arg1: gaia

Bere ingelesezko ordaina “obtain” izango dela erabaki dugu, baina aditz honek hiru argumentu ditu PropBanken:

Arg0: Agent
Arg1: Theme
Arg2: source

Gure ustez, Arg2 hau ez da kasu azpikategorizatua euskaraz eta beraz, DAT elementua agertzen zaigunean, ArgM jartzea erabaki dugu. Kasu honetan ez dugu oharretan *balio berria* jarri behar ez garelako EADBEk ematen duen egituratik aldentzen.

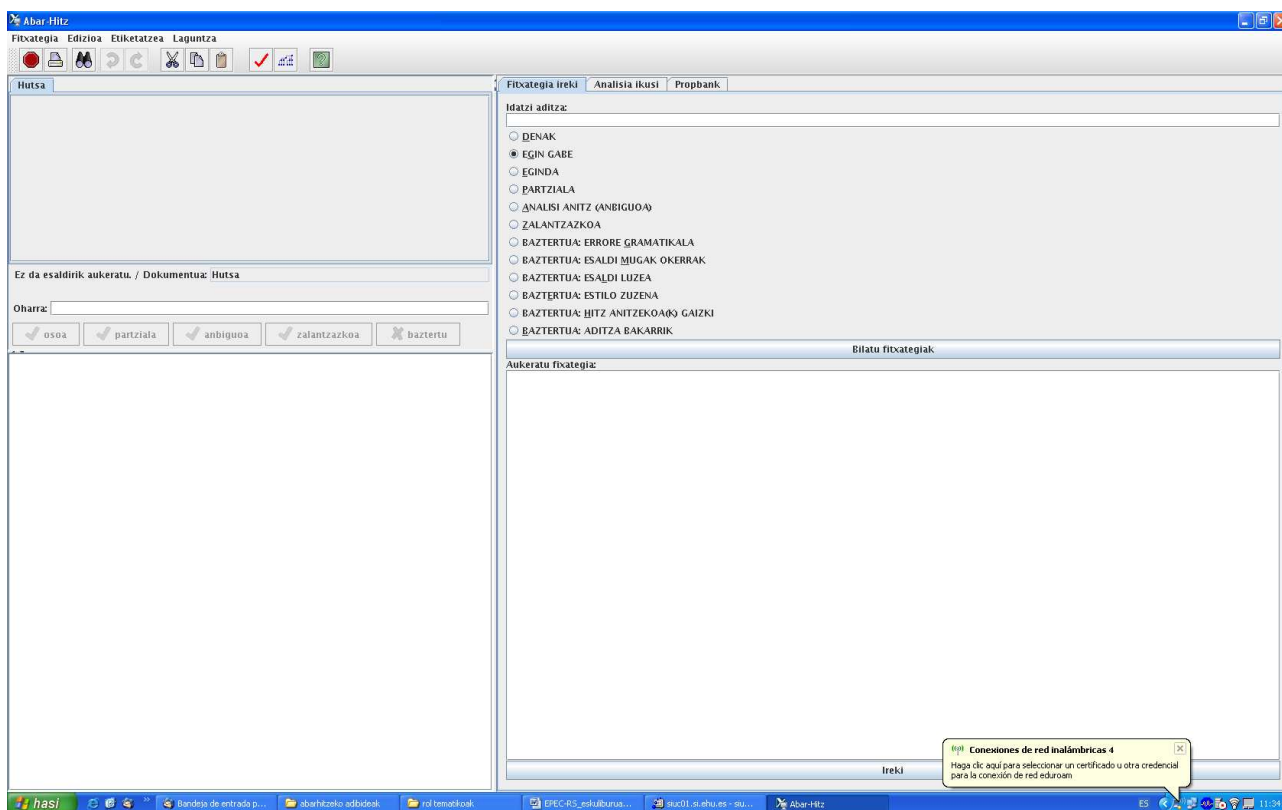
- *Joan* moduko aditzetan guk erabaki dugu subjektua **arg0** izatea. PropBanken arg1 gisa lantzen dute. Gure ustez gure hautua gehiago egokitzen zaio eurak proposatutako honako irizpideari: elementu bat biak izan badaiteke –arg0 eta arg1– aldi berean agentea eta ekintzaren jasalea delako, arg0 hautatuko da.
- Rol berri baten falta sumatu dugu: **path: bidea**

8. Interfazeari begira nahi genituzkeen hobekuntzak

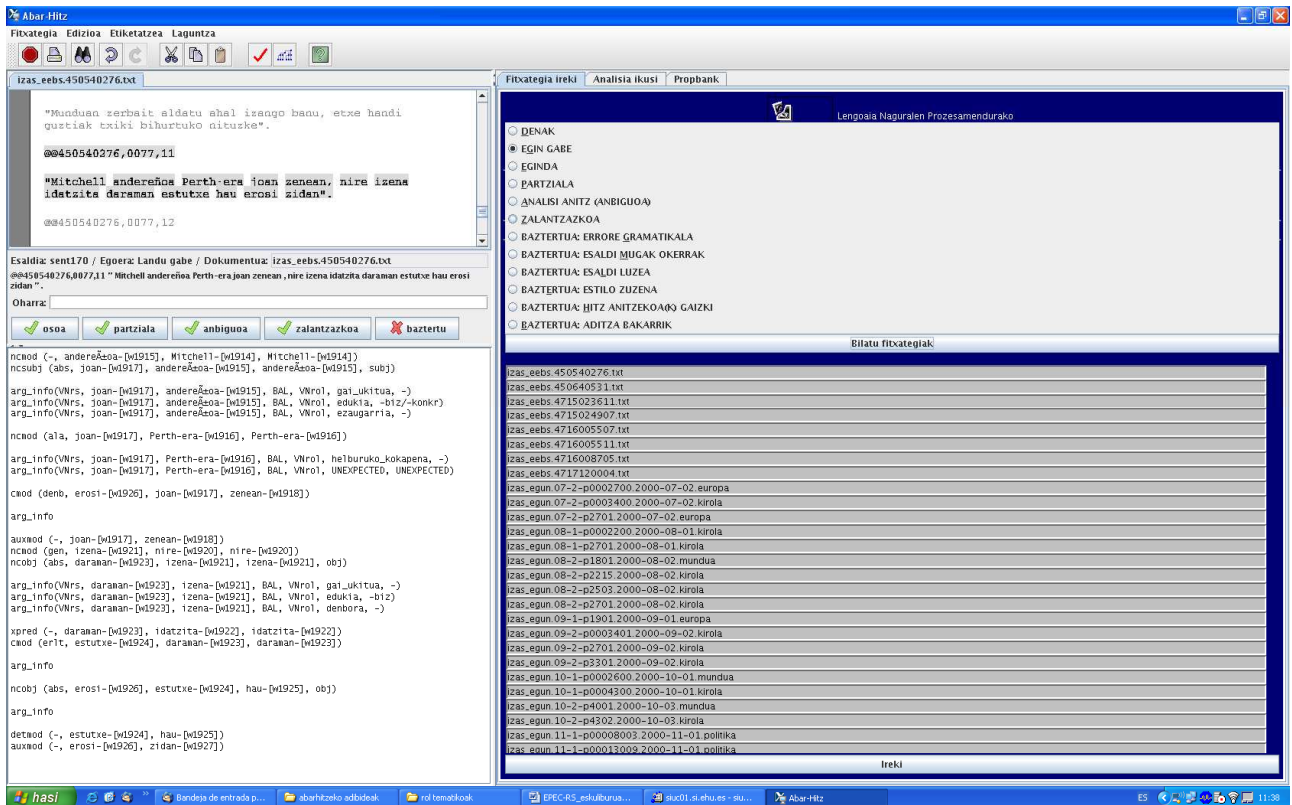
Rol sematikoak etiketatzeko erabili dugun tresna etiketatze sintaktikorako erabilitako Abarhitz bera da. Tresna honi buruzko informazio zehatza nahi izanez gero, jo beza irakurleak lan honetara: (Díaz de Ilarraza *et al.*, 2004). Hemen, batetik interfazeari dagozkion irudiak txertatuko ditugu, eta bestetik, interfazea hobetzeko proposamenak zerrendatuko ditugu.

8.1. Interfazeari dagozkion irudiak

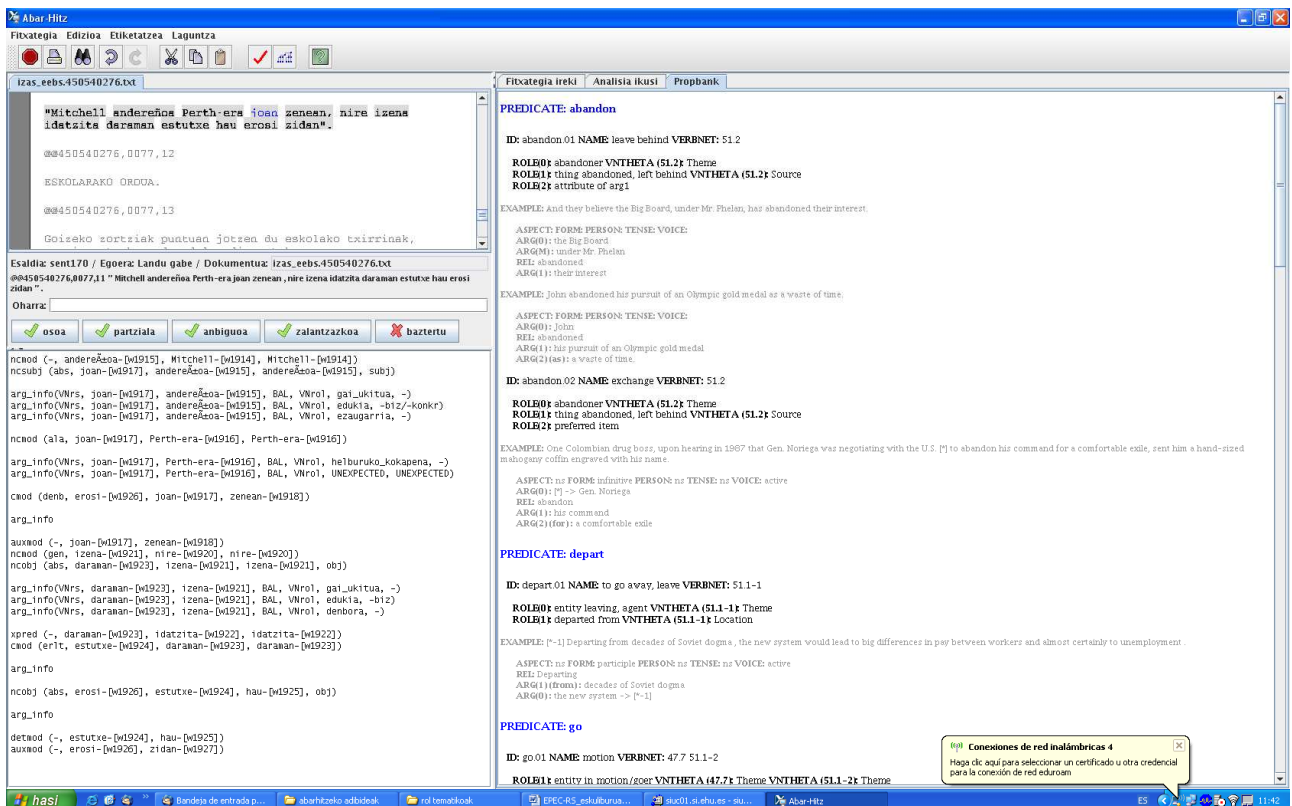
- lehen leihu honetan, etiketatu nahi dugun aditza adierazi behar da eskuineko “Idatzi aditza” atalean.



- “joan aditzari dagokion fitxategietako bat zabaldu ondoren, ezkerreko aldean aztergai den esaldiaren analisi automatikoa dago, eskuz gainbegiratua izan dadin.



c) Aurreko irudiarekin alderatuta, honetan eskuineko aldean ProBank/Verbnet datu-basetako informazioa da ageri dena.



8.2. Interfazea hobetzeko proposamenak

Hobekuntza hauek dira proposatu nahi direnak:

- Esaldien analisia zuhaitz edo diagrama modura ikusterakoan, rola eta gainontzeko informazio guztia ere ikus dadila.
- Lantzen ari garen aditza (burua aditz hori denean, alegia) nolabait markatzea, gu zuzenean joan ahal izateko landu behar den puntura.
- Behin betegabeko eremu bat osatutakoan, eremuaren deskribapena automatikoki desagertzea. Hau horrela da dagoeneko; gakoa da juxtu aurreko zeinua (koma edo parentesia) kentzea eta berriz idaztea.
- EADBko analisiak modu berean emanak etortzea: arg_info etiketa automatikoki osatuta datorrenean, osagai bat baino gehiagoko balioa azpimarrarekin dator (esaterako, *gai_ukitua*); etiketa hori geuk osatzen dugunean, aldiz, aukeretan azpimarrarik gabe dator balio hori (gai ukitua).
- Arg_info etiketatako informazioa baliatu behar izanez gero, kontuan izan etiketaren lehen eremuan aditz bat baino gehiago adierazi ditugunean ez dugula beti PropBank-eko hurrenkera jarraitu. Adibidez, amaitu aditzean, batzuetan (end_01/finish_01) idatzi dugu, eta beste batzuetan (finish_01/end_finish). 2009ko azaroaren 5a bitarteko aditzen analisisetan gerta daiteke hau; aurrerantzean PropBank-eko hurrenkera jarraitzea erabaki dugu. Data hori baino lehenago aztertutako aditzak dira: esan, joan, atera, jokatu, ahaztu, konturatu, eman (Ainararen fitxategiak), entzun, hil, aldatu, ikusi, lortu, eskatu, amaitu eta asmatu.

Erreferentziak

- Aduriz I., Aranzabe M., Arriola J., Atutxa A., Díaz de Ilarraza A., Ezeiza N., Gojenola K., Oronoz M., Soroa A., Urizar R. (2006). Methodology and steps towards the construction of EPEC, a corpus of written Basque tagged at morphological and syntactic levels for the automatic processing. *Corpus Linguistics Around the World. Book series: Language and Computers. Vol 56 (pag 1- 15). ISBN 90-420-1836-4 Ed. Andrew Wilson, Paul Rayson, and Dawn Archer. Rodopi. Netherlands.*
- Agirre E., Aldezabal I., Etxeberria J., Iruskieta M., Izagirre E., Mendizabal K., Pociello E. (2005). EUSEMCOR: euskarako corpora semantikoki etiketatze- eta editatze-, etiketatze- eta epaitze-lanak. *UPV/EHU/LSI/TR 23-2005*
- Aldezabal I., Aranzabe M., Atutxa A., Gojenola K., Sarasola K., Goenaga P. (2001). Extracción masiva de información sobre subcategorización verbal vasca a partir de corpus. *Actas del XVII Congreso de la SEPLN. ISSN 1135-5948. Nº. 27. 29-36 orriak. Universidad de Jaen, septiembre de 2001.*

- Aldezabal I. (2004). ADITZ-AZPIKATEGORIZAZIOAREN AZTERKETA SINTAXI PARTZIALETIK SINTAXI OSORAKO BIDEAN. 100 aditzen azterketa, Levin-en (1993) lana oinarri hartuta eta metodo automatikoak baliatuz. *Euskal Filologia Saila. Zientzia Fakultatea. Leioa. UPV/EHU.*
- Aldezabal I., Ceberio K., Esparza I., Estarrona A., Etxeberria J., Iruskieta M., Izagirre E., Uria L. (2007a). EPEC (Euskararen Prozesamendurako Erreferentzia Corpora) segmentazio-mailan etiketatzeko eskuliburua. *UPV/EHU / LSI / TR 11-2007*
- Aldezabal I., Aranzabe M., Arriola J., Díaz de Ilarraza A., Estarrona A., Fernandez K., Iruskieta M., Uria L. (2007b). EPEC (Euskararen Prozesamendurako Erreferentzia Corpora) dependentzietan etiketatzeko eskuliburua. *UPV/EHU / LSI / TR 12-2007*
- Aldezabal I. (2008). Aditz-azpikategoriazioa. *Artiagoitia*, X. & J. A. Lakarra (arg) *Gramatika Jaietan. Patxi Goenagaren omenez. ASJUren gehigarria*, 29-46, Donostia: Gipuzkoako Foru Aldundia-EHU.. ISBN:978-84-9860-085-8
- Ceberio K., Aduriz I., Díaz de Ilarraza A., García I. (2008). La anotación de la referencia sobre un corpus periodístico en euskara. *XXVI Congreso internacional de AESLA. Almería. ISBN: 978-84-612-2610-8*
- Díaz de Ilarraza A., Garmendia A., Oronoz M. (2004). Abar-Hitz: An Annotation Tool for the Basque Dependency Treebank. *Language Resources and Evaluation Conference (LREC 2004). Lisbon, Portugal. ISBN 2-9517408-1-6*
- Iruskieta M., Díaz de Ilarraza A., Lersundi M. (2008). Análisis de los marcadores del discurso para el euskera: denominación, clases, relaciones semánticas y tipos de ambigüedad. *XXVI Congreso internacional de AESLA. Almería.*
- Kingsbury P., Palmer M. (2002). From Treebank to PropBank. In *Proceedings of the 3rd International Conference on Language Resources and Evaluation (LREC-2002)*. Las Palmas, Spain.
- Levin B. (1993). *English Verb Classes and Alternations. A preliminary Investigation*. Chicago and London. The University of Chicago Press.
- Marti T., Marquez Ll., Taulé M., Bertran M. (2007). Anotación semiautomática con papeles temáticos de los corpus CESS-ECE Procesamiento del lenguaje natural, ISSN 1135-5948, N°. 38, pags. 67-76
- Palmer, M., Gildea, D., Kingsbury, P. (2005). "The Proposition Bank: A Corpus Annotated with Semantic Roles". In *Computational Linguistics Journal*. 31:1. Proceedings of Workshop on Linguistically Interpreted Corpora, EACL'99, Bergen. Sydney, July 2006. Association for Computational Linguistics.
- Vázquez G. & Fernández A. (2000). "Alternancias de diátesis: Estudio comparativo entre el comportamiento de los verbos en castellano e inglés. Hacia una propuesta de tipología de desajustes verbales" *X-TRACT Working Paper (Enero 2000), WP- 00/01.*
- Zapirain B., Agirre E., Márquez L. (2008). Robustness and Generalization of Role Sets: PropBank vs. VerbNet. *Proceedings of the 46th Annual Meeting of the Association of Computational Linguistics, ACL-08: HLT*, 550-558, Columbus, Ohio.

A ERANSKINA: PropBank-eko gidalerroak

PROPBANK ANNOTATION GUIDELINES

Olga Babko-Malaya

September 2005

1. PropBank Annotation Goals	2
2. Task 1: Argument Labeling	2
2.1 Frame Files.....	2
2.2 Choosing Arg0 versus Arg1.....	4
2.3 Annotation of null elements.....	6
2.3.1 Passive sentences	6
2.3.2. Fronted and dislocated arguments:	7
2.3.3. Questions and wh-phrases.....	9
2.3.4. Relative clauses.....	10
2.3.5. ICH traces (ICH: interpret constituent here).....	12
2.3.6. Right Node Raising (RNR) traces.....	13
2.3.7. *EXP* (`EXpletive").....	15
2.3.8. Other traces	17
2.3.9. Prepositional phrases with traces	19
2.4. Special cases	19
2.4.1. Verbs of Saying.....	19
2.4.2. Small clauses and sentential complements.	21
3. Task 2: Annotation of modifiers (ArgMs).	24
3.1. Directionals (DIR)	24
3.2. Locatives (LOC)	25
3.3. Manner Markers (MNR)	26
3.4. Temporal markers (TMP)	26
3.5. Extent Markers (EXT)	27
3.6. Reciprocals (REC)	28
3.7. Markers of secondary predication (PRD)	28
3.8. Purpose clauses (PNC: purpose, not cause).....	29
3.9. Cause clauses (CAU)	30
3.10. Discourse Markers (DIS)	30
3.11. Adverbials (ADV).....	32
3.12. Modals (MOD).....	32
3.13. Negation (NEG)	32
3.14. Stranded (STR)	33
3.15. Special cases	34
3.15.1. Modification in complex clauses	34
3.15.2. Multiple modifiers with the same tag	34
3.15.3. Multiple Propositions.....	36
4. Task 3: Coreference (for empty categories).....	37

1. PropBank Annotation Goals

PropBank is an annotation of syntactically parsed, or treebanked, structures with 'predicate-argument' structures. An important goal is to provide consistent argument labels across different syntactic realizations of the same verb, as in

[ARG0 John] broke [ARG1 the window]
[ARG1 The window] broke

As this example shows, the arguments of the verbs are labeled as numbered arguments: Arg0, Arg1, Arg2 and so on.

The second task of the PropBank annotation involves assigning functional tags to all modifiers of the verb, such as manner (MNR), locative (LOC), temporal (TMP) and others:

Mr. Bush met him privately, in the White House, on Thursday.

Rel: met
Arg0: Mr. Bush
Arg1: him
ArgM-MNR: privately
ArgM-LOC: in the White House
ArgM-TMP: on Thursday

And, finally, PropBank annotation involves finding antecedents for 'empty' arguments of the verbs, as illustrated below:

I made a decision [] to leave.*

The subject of the verb 'leave' in this example is represented as an empty category [*] in Treebank. In Propbank, all empty categories which could be co-referred with a NP within the same sentence are linked in 'co-reference' chains:

Rel: leave
Arg0: [*] -> I

These three tasks of Propbank annotation: argument labeling, annotation of modifiers, and creating co-reference chains for empty categories are discussed in detail below.

2. Task 1: Argument Labeling

2.1 Frame Files

The argument labels for each verb are specified in the frame files, which are available at <http://verbs.colorado.edu/framesets/>. Frame files provide verb-specific description of all possible semantic roles, as well as illustrate these roles by examples.

Frame File for the verb 'expect':

Roles:
Arg0: expecter
Arg1: thing expected

Example: Transitive, active:

Portfolio managers expect further declines in interest rates.

Arg0: Portfolio managers

REL: expect

Arg1: further declines in interest rates

For some verbs, it is impossible to provide one set of semantic roles for all senses of the verb. For example, the two senses of the verb 'leave' in the examples below take different arguments:

Mary left the room

Mary left her daughter-in-law her pearls in her will

In such cases, frame files distinguish two or more verb senses, which are called Framesets, and define argument labels specific to each Frameset:

Frameset leave.01 "move away from":

Arg0: entity leaving

Arg1: place left

Frameset leave.02 "give":

Arg0: giver

Arg1: thing given

Arg2: beneficiary

When annotating, annotators first select the frameset and then assign the argument labels as specified for this frameset. Please note that the annotation tool allows you to see the semantic roles and one example for the first frameset, but it is absolutely necessary to check the frame files to see if the verb has more than one frameset.

In some cases, frame files define not only several framesets for each verb, but also several predicates. If a verb has a particle (marked as PRT in TreeBank), then it is being considered as a different predicate, and has a different set of semantic roles. For example, the frame file for the verb 'keep' defines three predicates: predicate 'keep' (which has 3 framesets), and predicates 'keep_up' and 'keep_on'. The following example illustrates the definition of the predicate 'keep_up'. Note that the relation (REL) in PB annotation should include both the verb and the particle (which should be selected as one node, if possible, or as a concatenated constituent (i.e. [keep][up]), if one node is not available in Treebank).

Frameset: keep_up:

keep.05 "keep up: maintain position":

Arg0: maintainer of position

Arg1: relative to what

John can't keep up with Mary's rapid mood swings.

Arg0: John

Argm-MOD: ca

Argm-NEG: n't

REL: keep up

Arg1: with Mary's rapid mood swings

2.2 Choosing Arg0 versus Arg1

In most cases, choosing an argument label is straightforward, given the verb specific definition of this label in the frame files. However, in some cases, the decision needs to be made concerning choosing Arg0 or Arg1 labels.

The Arg0 label is assigned to arguments which are understood as agents, causers, or experiencers. The Arg1 label is usually assigned to the patient argument, i.e. the argument which undergoes the change of state or is being affected by the action.

Arg0 arguments (which correspond to external arguments in GB theory) are the subjects of transitive verbs and a class of intransitive verbs called unergatives.

John (Arg0) sang the song.
John (Arg0) sang.

Semantically, external arguments have what Dowty 1991 called Proto-Agent properties, such as

- volitional involvement in the event or state
- causing an event or change of state in another participant
- movement relative to the position of another participant

Internal arguments (labeled as Arg1) are the objects of transitive verbs and the subjects of intransitive verbs called unaccusatives:

John broke the window (Arg1)
The window (Arg1) broke

These arguments have Proto-Patient properties, which means that these arguments

- undergo change of state
- are causally affected by another participant
- are stationary relative to movement of another participant

Whereas for many verbs, the choice between Arg0 or Arg1 does not present any difficulties, there is a class of intransitive verbs (known as verbs of variable behavior), where the argument can be tagged as either Arg0 or Arg1.

A bullet-Arg1 landed at his feet
He-Arg0 landed

Arguments which are interpreted as agents should always be marked as Arg0, independent of whether they are also the ones which undergo the action.

In general, if an argument satisfies two roles, the highest ranked argument label should be selected, where Arg0 >> Arg1 >> Arg2>>... .

Given this rule, agents are ranked higher than patients. If an argument is both an agent and a patient, then Arg0 label should be selected.

Not all Arg0s are agentive, however. There are many inanimate as well as clausal arguments which are being marked as Arg0s. These arguments are usually the ones which cause an action or a change of state.

A notion which might be useful for selecting Arg0 arguments is the notion of ‘internally caused’ as opposed to ‘externally caused’ eventualities, as defined in Levin and Rapaport 1995. In internally-caused eventualities, “some property inherent to the argument of the verb is responsible for bringing about the eventuality... For agentive verbs such as *play*, *speak*, or *work*, the inherent property responsible for the eventuality is the will or volition of the agent who performs the activity. However, an internally caused eventuality need not be agentive. For example, the verbs *blush* and *tremble* are not agentive, but they, nevertheless, can be considered to denote internally caused eventualities, because these eventualities arise from internal properties of the arguments, typically an emotional reaction... In contrast to internally caused verbs, verbs which are externally caused inherently imply the existence of an external cause with an immediate control over bringing about the eventuality denoted by the verb: an agent, and instrument, a natural force, or a circumstance. Thus something breaks because of the existence of some external cause; something does not break because of its own properties” /Levin and Rapaport 1995/. The difference between internal and external causation is important for distinguishing Arg0s and Arg1s: the arguments which are responsible for bringing out the eventuality are Arg0s, whereas those which undergo an externally caused event are Arg1s.

To sum up, Arg0 arguments are the arguments which cause the action denoted by the verb, either agentively or not, as well as those which are traditionally classified as experiencers, i.e. the arguments of stative verbs such as love, hate, fear. Arg1 arguments, on the other hand, are those that change due to external causation, as well as other types of ‘patient’-like arguments.

2.3 Annotation of null elements

The inventory of null elements used in Penn Treebank is as follows (see guidelines for English Treebank in www.cis.upenn.edu/~treebank).

- [*T*] (trace of A-movement, including parasitic gaps)
- [(NP *)] (arbitrary PRO, controlled PRO, and trace of A-movement)
- [0] (null complementizer, including null wh-operator)
- [*U*] (unit)
- [*?*] (placeholder for ellipsed material)
- [*NOT*] (anti-placeholder in template gapping)
- [*RNR*] (pseudo-attach: right node raising)
- [*ICH*] (pseudo-attach: interpret constituent here)
- [*EXP*] (pseudo-attach: expletive)
- [*PPA*] (pseudo-attach: permanent predictable ambiguity)

This section presents some examples of most commonly used null elements and their Propbank annotation.

2.3.1 Passive sentences

Sentences can be either active (The executive committee approved the new policy) or passive (The new policy was approved by the executive committee). In active sentences, the subject is the agent or a do-er of the action, marked as Arg0 in Propbank. In passive sentences, the subject of the sentence is acted upon by some other agent or by something unnamed, and is being marked as Arg1 in Propbank.

Passive sentences are assumed to be derived from the corresponding active sentences by ‘movement’ of the object to the subject position. This movement leaves a trace,

represented as [*T*] in Treebank.

Active: Mary hit John
Passive: John was hit [*T*] by Mary.

Since the goal of the Propbank annotation is to mark the ‘maximal extent’ of the arguments, in the case of passive sentences what is being labeled as Arg1 is the chain [*T*] -> John.

Propbank annotation:
rel: hit
Arg1: [*T*] -> John
ArgM-MNR: hard
Arg0: by Mary

In most cases, the link between the trace and the NP is already provided by the Treebank annotation, which is indicated by the number ‘1’ in the trace (NP-3 *-1) and (NP-SBJ-1 he) below⁴.
Treebank annotation:

```
(S (NP-SBJ-1 he)
  (VP was
    (VP accused
      (NP-3 *-1)
      (PP-CLR of
        (S-NOM (NP-SBJ *-3)
          (VP (VP conducting
            (NP illegal business))
            and
            (VP possessing
              (NP illegal materials))))))))
```

Propbank annotation:
Arg1: [NP-3 *-1] -> he
Rel: accused
Arg2: of [*3*] conducting illegal business and possessing illegal materials

2.3.2. Fronted and dislocated arguments:

Other examples of moved constituents are fronted or otherwise dislocated arguments and adjuncts. As in the other cases of movement, fronted elements leave a trace, which should be coindexed with the moved constituent.

In the following example, the Arg2 argument of the verb ‘put’ is being fronted. In the Propbank annotation, this is indicated by the chain which links the trace [*T*-1] with the adverbial ‘There’:

Treebank annotation:
(S (ADVP-PUT-TPC-1 There)
,
(NP-SBJ I)
(VP put
(NP the book)
(ADVP-PUT *T*-1)))

Propbank annotation:

⁴ When Treebank provides a link, it can be reconstructed in rats by using the command ‘t’.

rel: put
 Arg0: I
 Arg1: the book
 Arg2: [*T*-1] -> There

Modifiers, or ArgMs, can be fronted as well, as the following example shows:

Treebank annotation:

(S (SBAR-PRP-TPC-9 Because
 (S (NP-SBJ I
 (VP 'm
 (NP-PRD such a bad boy))))
 (NP-SBJ I)
 (VP think
 (SBAR 0
 (S (NP-SBJ I
 (VP wo n't
 (VP get
 (NP a lollipop)
 (SBAR-PRP *T*-9))))))

Since the 'because' clause modifies the verb 'get' in this example, the trace originates as the modifier of 'get', and should be coindexed with the fronted phrase:

Propbank annotation:

rel: get
 Arg1: a lollipop
 Arg0: I
 ArgM-NEG: n't
 ArgM-MOD: wo
 ArgM-CAU: [*T*-9] -> Because I'm such a bad boy

In rare situations, movement does not leave a trace, but rather leaves a pronoun (called a resumptive pronoun). In such cases no chains are created, but rather the pronoun is being labeled as an argument, whereas a dislocated NP is being marked as ArgM-DIS:

Treebank annotation:

(S (NP-TPC John)
 '
 (NP-SBJ I)
 (VP like
 (NP him)
 (NP-ADV a lot)))

Propbank annotation:

Rel: like
 Arg0: I
 Arg1: him
 ArgM-MNR: a lot
 ArgM-DIS: John

2.3.3. Questions and wh-phrases

Another type of traces is a trace of a wh-phrase in questions.

What do you like?

As in the case of passive sentences, questions are assumed to be derived by movement. In the example below, the Arg1 argument of the verb 'like' is a wh-phrase 'what', which moves from the object position of the verb to the front of the sentence. This movement leaves a trace, as shown below:

*What do you like [**T**]?*

In Propbank, the argument Arg1 is a chain which links the trace and the moved NP:

Propbank annotation:

Rel: like

Arg0: you

Arg1: [**T**] -> What

In Treebank annotations, wh-phrases are marked as WHNP. As in the case of passive sentences, Treebank provides a link between the trace and the moved WHNP, which can be reconstructed in rats by using 't' command:

Treebank annotation:

```
(SBARQ (WHNP-1 what)
  (SQ do
    (NP-SBJ you)
    (VP like
      (NP *T*-1)))
```

Wh-phrases are not necessarily arguments. Questions can be formed with wh-phrases like when, where, how, in which case they should be tagged as ArgMs.

Treebank annotation:

```
SBARQ (WHNP-1 Which day)
  (SQ did
    (NP-SBJ you)
    (VP get
      (ADVP-DIR there)
      (NP-TMP *T*-1)))
```

Propbank annotation:

Arg1: you

rel: get

ArgM-TMP: [***] -> which day

Treebank annotation:

```
(SBARQ (WHADVP-42 How)
  (SQ did
    (NP-SBJ you)
    (VP fix
      (NP the car)
      (ADVP-MNR *T*-42)))
```

?)

Propbank annotation:

rel: fix

Arg0: you

Arg1: the car

ArgM-MNR: [*T*-42] -> [How]

Questions can also be embedded, as in the example below. Propbank annotation is not different from direct questions in this case:

John didn't know where his parents had met [].*

Arg0: his parents

rel: met

ArgM-LOC: [*] -> where

2.3.4. Relative clauses

Relative clauses are clauses which modify a N or a NP as in ‘answers that we’d like to have’. Relative clauses also include a trace, which is coindexed with the relative pronoun in Treebank (e.g. ‘that’/’which’/’who’). Propbank annotation differs from the Treebank annotation in that it also provides a link to the antecedent NP.

For example, in the following Treebank annotation, the Arg1 argument of the verb ‘have’ is the NP ‘answers’. The object position of the verb has a trace (NP *T*-6), which is being coindexed with the relative pronoun (WHNP-6 that/which). Propbank annotation reconstructs this chain, plus adds a link to the NP ‘answers’:

Treebank annotation:

(NP (NP answers)

(SBAR (WHNP-6 that/which)

(S (NP-SBJ-3 we)

(VP 'd

(VP like

(S (NP-SBJ *-3)

(VP to

(VP have

(NP *T*-6)))))))))

Propbank annotation:

Arg1: [NP *T*-6] -> which -> answers

rel: have

Arg0: [NP-SBJ *-3] -> we

Likewise, if a relative clause modifies a temporal or a locative, the chain should be constructed which links the trace, the relative pronoun ‘when’ or ‘where’ and the NP which specifies time or location:

*John found the place where his parents had met *.*

Arg0: his parents

rel: met

ArgM-LOC: [*] -> where -> the place

In some cases, a relative pronoun is missing in a relative clause (as in: ‘answers we would like to have’). Such clauses have a null complementizer ‘0’ inside SBAR in Treebank annotations:

Treebank annotation:
 (NP (NP **answers**)
 (SBAR (WHNP-3 0)
 (S (NP-SBJ-4 we)
 (VP 'd
 (VP like
 (S (NP-SBJ *-4)
 (VP to
 (VP have
 (NP *T*-3))))))))))

In such cases, the null complementizer should be included in the chain, instead of a relative pronoun:

Propbank annotation:
 Rel: have
 Arg1: [*T*-3] -> [WHNP-3 0] -> answers
 Arg0: we

A similar analysis applies to infinitival relatives:

Treebank annotation:
 (NP (NP **a movie**)
 (SBAR (WHNP-1 0)
 (S (NP-SBJ *)
 (VP to
 (VP see
 (NP *T*-1))))))

Propbank annotation:
 Rel: see
 Arg1: [*T*-1] -> [WHNP-1 0] -> a movies
 Arg0: [SBJ *]

2.3.5. ICH traces (ICH: interpret constituent here)

ICH traces are being used in Treebank to indicate a relationship of constituency between elements separated by intervening material. An example of such ‘split constituents’ are ‘heavy shift’ constructions, illustrated below:

Treebank annotation:
 (S (NP-SBJ (NP **a young woman**)
 (SBAR *ICH*-1)
 (VP entered
 (SBAR-1 (WHNP-2 **whom**)
 (S (NP-SBJ **she**)
 (PP-TMP **at**
 (ADVP **once**)))

(VP recognized
 (NP *T*-2)
 (PP-CLR as
 (NP Jemima Broadwood))))))

The subject NP in this case is being split into two constituents: the NP ‘a young woman’ and SBAR: ‘whom she at once recognized as Jemima Broadwood’. The ICH trace specifies a link to the SBAR node in this example.

All examples of this type should be annotated as concatenated constituents (using the command ‘A’ in rats), including the ICH trace:

Propbank annotation:

Arg0: [a young woman [*ICH*-1]][whom she recognized as Jemima Broadwood]

Rel: entered

Other typical examples of *ICH* traces are shown below:

*[Five *ICH*-1] ran, [out of the twenty-five that showed up].*

Arg0: [Five *ICH-1*][out of the twenty-five that showed up]

rel: ran

*Some people in Paris want to hear more [*ICH*] from me than those fellers over at the conference house do.*

Rel: hear

Arg1: [more [ICH]][than those fellers over at the conference house do]

Arg2: from me

2.3.6. Right Node Raising (RNR) traces

RNR traces are used when a constituent is interpreted simultaneously in more than one place. An example of a right node raising structure is given below:

Treebank annotation:

(NP (NP (ADJP so many) enchained demons)

(VP straining

(PP-MNR in

(NP anger))

(S (NP-SBJ *)

(VP to

(VP (VP tear

(NP *RNR*-1))

and

(VP gnaw

(PP-CLR on

(NP *RNR*-1)))

(NP-1 his bones))))))

In this example, the NP ‘his bones’ is interpreted as both the argument of the verb ‘tear’ and the verb ‘gnaw’. When annotating the verb ‘tear’, the chain should be constructed

which links the trace (NP *RNR*-1) and the NP:

Propbank annotation:

Rel: tear

Arg1: [*RNR*] -> his bones

Arg0: [*] -> so many enchained demons

Likewise, when annotating the verb ‘gnaw’, the chain should be constructed which links (PP-CLR on (NP *RNR*-1))) (see section 2.3.9 for annotation of prepositional phrases with traces) and the NP:

Propbank annotation:

Rel: gnaw

Arg1: on [*RNR*] -> his bones

Arg0: [*] -> so many enchained demons

A similar annotation applies when the [*RNR*] trace is a clausal argument:

*I want *RNR* and like *RNR* [* to eat ice-cream].*

Arg0: I

rel: want

Arg1: *RNR* -> [* to eat ice-cream]

If the RNR trace is part of the argument of the verb, then use concatenation instead (as always, including the trace):

His dreams had revolved around her so much and for so long that...

Treebank annotation:

(S (NP-SBJ His dreams)

(VP had

(VP revolved

(PP-CLR around

(NP her))

(UCP-ADV (ADVP (ADVP so much)

(SBAR *RNR*-1))

and

(PP-TMP for

(NP (NP so long)

(SBAR *RNR*-1)))

(SBAR-1 that...))))))

Propbank annotation:

Arg1: his dreams

rel: revolved

ArgM-LOC: around her

ArgM-MNR: [so much [*RNR]][that ...]

The following example illustrates annotation of RNR traces within a small clause (for annotation of small clauses see section 2.4.2 below). There are three concatenated constituents in this case:

But our outlook has been and continues to be defensive

```
(S But
  (NP-SBJ-2 our outlook)
  (VP (VP has
    (VP been
      (ADJP-PRD *RNR*-1)))
    ,
    and
    (VP continues
      (S (NP-SBJ *-2)
        (VP to
          (VP be
            (ADJP-PRD *RNR*-1))))))
    ,
    (ADJP-PRD-1 defensive)))
```

Propbank annotation:

rel: continue

Arg1: [our outlook][[*-2] to be *RNR-1][defensive]

2.3.7. *EXP* (‘EXpletive’)

Expletives like ‘it’ do not add any meaning to the sentence. In the following example, the syntactic subject of the sentence is an expletive, which includes a trace EXP-1. This trace refers to the logical subject of the sentence, marked as SBAR-1:

Treebank annotation:

```
(S (NP-SBJ (NP It)
  (SBAR *EXP*-1))
  (VP is
    (ADJP-PRD clear)
    (PP to
      (NP me))
    (SBAR-1 that
      (S (NP-SBJ this message)
        (VP is
          (ADJP-PRD unclear))))))
```

In Propbank annotations, expletives and EXP traces are NOT INCLUDED:

Propbank annotation:

Rel: clear

Arg1: that this message is unclear

Another example:

It required an energy he no longer possessed to be satirical about his father.

Propbank annotation:

Arg0: to be satirical about his father

Arg1: an energy he no longer possessed
Rel: required

In the examples below, the expletives are the objects, rather than the subjects. As in the case of expletive subjects, only the logical argument is being tagged, whereas an expletive and an EXP trace are not part of the Propbank annotation:

Mrs. Yeargin was fired [-1] and prosecuted [*-1] under an unusual South Carolina law that [*T*-79] makes it [*EXP*-2] a crime [*] to breach test security.*

Propbank annotation:

ARG0: [*T*-79] -> that -> an unusual South Carolina law
rel: makes
ARG2: a crime
ARG1: [*] to breach test security

*Any raider would find it [*EXP*-1] hard [*] to crack AG 's battlements.*

Trebank annotation:

(S
 (NP-SBJ (DT Any) (NN raider))
 (VP (MD would)
 (VP (VB find)
 (S
 (NP-SBJ
 (NP (PRP it))
 (S (-NONE- *EXP*-1)))
 (ADJP-PRD (JJ hard))
 (S-1
 (NP-SBJ (-NONE- *))
 (VP (TO to)
 (VP (VB crack)
 (NP
 (NP (NNP AG) (POS 's))
 (NNS battlements)))))))))
 (. .))

Propbank annotation:

ARG0: Any raider
ARGM-MOD: would
rel: find
ARG3: hard
Arg1: [*] to crack AG 's battlements

Common mistake: Please make sure to distinguish an expletive ‘it’ from the referring pronoun ‘it’, where ‘it’ refers to a previous NP, a clause, or an event. (hint: referring pronouns are not being followed by an EXP trace in Treebank). All referring pronouns, including ‘it’, should be marked as arguments in Propbank.

It sounds good.
Rel: sound
Arg1: it

ArgM-MNR: good

Italy 's Foreign Ministry said [0] it is investigating exports to the Soviet Union.

Rel: investigating

Arg0: it

Arg1: exports to the Soviet Union

2.3.8. Other traces

Other types of traces include null complementizer trace, ? trace (used in ellipsis constructions), and *PPA* trace in cases of predictable ambiguous attachments. Null complementizer trace should be included as part of the clausal argument.

Treebank annotation:

```
(S (NP-SBJ I)
  (VP believe
    (SBAR 0
      (S (NP-SBJ you)
        (VP are
          (ADJP-PRD smart))))))
```

Propbank annotation:

rel: believe

Arg0: I

Arg1: [[O] you are smart]

In cases of ellipsis, ? trace can be reconstructed if the sentence is a conjunction:

```
(S (S (NP-SBJ Robin)
  (VP likes
    (NP ice cream)))
  , and
  (S (NP-SBJ Kim)
    (VP does
      (VP *?*
        (ADVP too))))))
```

Multiple propositions are being used in this case (see section 3.14.3 for more discussion of multiple propositions). The ? trace is being reconstructed as a relation 'likes', introducing a second proposition:

Propbank annotation (2 propositions):

Rel: likes

Arg0: Robin

Arg1: ice cream

Rel: likes

Arg0: Kim

Arg1: ice cream

ArgM-DIS: too

However, if the antecedent of the ? trace is not in the conjoined clause, then it should not be annotated:

Treebank annotation:

```
(S (S (NP-SBJ She)
      (ADVP-TMP rarely)
      (VP sings))
  ,
  so
  (S (NP-SBJ I)
    (VP do n't
      (VP think
        (SBAR 0
          (S (NP-SBJ she)
            (VP will
              (VP *?*
                (NP-TMP tonight))))))))))
```

2.3.9. Prepositional phrases with traces

The following sentence illustrates a NP trace within a prepositional phrase (PP). The argument ARG1 is the PP in this sentence. The desired interpretation of this sentence should look like [for [[*-2] -> kid's car]], however the tool does not allow to create a chain within an argument.

Propbank annotation:

```
(S (NP-SBJ-2 (NP kid 's) cars)
  (VP are
    (ADVP-TMP often)
    (VP paid
      (PP-CLR for
        (NP *-2))
      (PP by
        (NP-LGS their parents))))))
```

The solution which has been adopted (mostly for practical reasons) is to link the PP with a trace and the extracted NP, as shown below:

```
Rel: paid
Arg0: by their parents
Arg1: [for *-2] -> kid's cars
ArgM-TMP: often
```

2.4. Special cases

2.4.1. Verbs of Saying

A verb of saying is any verb which has a speaker argument (Arg0) and an utterance (Arg1). If the utterance argument can be selected as one constituent, then Arg1 is a single constituent, or a chain, as shown below:

Treebank Annotation

```

(S ``
  (S-TPC-1 (NP-SBJ We)
    (VP will
      (VP win)))
  ;
  (NP-SBJ Mary)
  (VP said
    (S *T*-1))
  .))

```

Propbank Annotation

Rel: said

Arg1: [*T*-1] -> We will win

Arg0: Mary

Unfortunately, in many examples, the utterance does not correspond to one constituent in Treebank. In such cases, all pieces of the utterance should be concatenated, including the trace in the object position of the verb.

Among other things , they said [*T*-1] , Mr. Azoff would develop musical acts for a new record label .

Treebank annotation:

```

(S-1
  (PP (IN Among)
    (NP (JJ other) (NNS things) ))
  (PRN
    ( , )
    (S
      (NP-SBJ (PRP they) )
      (VP (VBD said)
        (SBAR (-NONE- 0)
          (S (-NONE- *T*-1) ))))
      ( , , )
      (NP-SBJ (NNP Mr.) (NNP Azoff) )
      (VP (MD would)
        (VP (VB develop)
          (NP
            (NP (JJ musical) (NNS acts) )
            (PP (IN for)
              (NP (DT a) (JJ new) (NN record) (NN label) )))))
      ( . . )
    )
  )

```

Propbank annotation:

ARG1: [Among other things] [Mr. Azoff] [would develop musical acts for a new record label] [[*T*-1]]

ARG0: they

rel: said

Punctuation generally should NOT be included, unless it is embedded in the constituent, so that it is impossible to select a constituent without punctuation.

`` By the end of the 1990s , " he said [*T*-4] , `` we want [*-2] to be producing roughly two vehicles [*ICH*-3] overseas for every vehicle that we export [*T*-1] from Japan . ''

ARG1: [By the end of the 1990s] [we] [want [*-2] to be producing roughly two vehicles [*ICH*-3] overseas for every vehicle that we export [*T*-1] from Japan] [*T* -4]
 ARG0: he
 rel: said

2.4.2. Small clauses and sentential complements.

This section is concerned with different types of clausal complements and modifiers. For example, verbs like ‘consider’ are given a small clause analysis in Treebank, which means that the syntactic complement of the verb ‘consider’ is a clause, marked as S. Propbank frame files do not always follow Treebank in this respect. In the frame files for the verb ‘consider’, for example, this verb has three semantic arguments, so that the clause S is decomposed into 2 arguments Arg1 and Arg2:

Treebank annotation:

(S (NP-SBJ I)
 (VP consider
 (S (NP-SBJ Kris)
 (NP-PRD a fool))))

Propbank annotation:

rel: consider
 Arg0: I
 Arg1: Kris
 Arg2: a fool

Propbank annotators should follow the frame files, and decompose S clauses into smaller constituents in such cases.

If a verb is passivized, the Arg1 argument is the chain linking the trace in the subject position of the small clause with the subject of the verb ‘consider’:

Treebank annotation:

(S (NP-SBJ-1 Kris)
 (VP is
 (VP considered
 (S (NP-SBJ *-1)
 (NP-PRD a fool))
 (PP by
 (NP-LGS most people))))))

Propbank annotation:

rel: consider
 Arg1: [*1]-> Kris
 Arg2: a fool
 Arg0: by most people

Other examples of verbs with small complements in Treebank, where the small clause is being decomposed in Propbank, are the verbs name, elect, tempt, lead, hold, force, call, deem, schedule, find, ask, make and some others.

(S (NP-SBJ-28 (NP Arthur A. Hatch)

,
(NP 59)
,)
(VP was
(VP named
(S (NP-SBJ *-28)
(NP-PRD (NP executive vice president)
(PP of
(NP the company))))))

Rel: named

Arg1: [*] -> Arthur A. Hatch, 59,

Arg2: executive vice president of the company

Small clauses (as well as other sentential complements) should only be decomposed if the frame files specify that the two decomposed constituents are different semantic arguments. In all other cases, the small clause analysis should be preserved. For example, in the following sentence, the clause S-CLR has a trace in the subject position of ‘asleep’, which is coindexed with the subject of the verb ‘fell’ ‘I’.

I fell asleep on the floor.

S (NP-SBJ-1 I)

(VP fell
(S-CLR (NP-SBJ *-1)
(ADJP-PRD asleep))
(PP-LOC on
(NP the lobby floor))))

When annotating the verb ‘fell’, the small clause (marked as S-CLR above) should not be decomposed, which means that the trace [NP*-1] is not part of the argument Arg1 of the verb ‘fell’. As the Propbank annotation below shows, the Arg1 argument of ‘fell’ is the NP ‘I’, rather than the chain [NP*-1] -> I, whereas the trace [NP*-1] is part of ArgM-PRD:

Propbank annotation:

Rel: fell

Arg1: I

ArgM-PRD: [NP-SBJ *-1] asleep

NOT:

Rel: fell

Arg1: [NP*-1] -> I

ArgM-PRD: asleep

Verbs like expect, allow and others are analyzed as having a clause as its argument. In this case, Propbank annotations follow Treebank analyses of these sentences, where the clausal complement is being marked as Arg1:

John expected Mary to come.

Propbank annotation:

Rel: expected

Arg0: John

Arg1: Mary to come

If such sentences are passivised, as shown below, then the Arg1 argument is a concatenation of the subject and the clausal complement.

Mary is expected [] to come*

Propbank annotation:

Rel: expected

Arg1: [Mary][* to come]

Similar analysis applies to verbs like ‘seem’ and ‘appear’, which are known as raising verbs. For example, the NP ‘everyone’ below is not the argument of the verb ‘seems’, but rather this sentence can be paraphrased as ‘It seems that everyone dislikes Drew Barrymore.’

Everyone seems to dislike Drew Barrymore

Treebank annotation:

(S (NP-SBJ-3 Everyone)

(VP seems

(S (NP-SBJ *-3)

(VP to

(VP dislike

(NP Drew Barrymore))))))

Since the logical argument of the verb *seems* is a clause, Propbank annotation of this sentence involves concatenation of the subject and the complement:

Rel: seems

Arg1: [Everyone][* to dislike Drew Barrymore]

And, finally, another class of verbs which follows this analysis is the class of aspectual verbs like *continue* and *start*, which take events as their arguments.

New loans continue [] to slow.*

Rel: continue

Arg1: [New loans][* to slow]

3. Task 2: Annotation of modifiers (ArgMs).

The following types of modifiers are being used in PropBank:

DIR: Directionals

LOC: Locatives

MNR: Manner
EXT: Extent
REC: Reciprocals
PRD: Secondary Predication
PNC: Purpose
CAU: cause
DIS: discourse
ADV: adverbials
MOD: modals
NEG: negation

3.1. Directionals (DIR)

Directional modifiers show motion along some path. Both "source" and "goal" are grouped under "direction." On the other hand, if there is no clear path being followed a "location" marker should be used instead. Thus, "walk along the road" is a directional, but "walk around the countryside" is a location. Directional modifiers are also used for some particles, as in back up.

Workers dumped large burlap sacks of the imported material into a huge bin , poured in cotton and acetate fibers and mechanically mixed the dry fibers in a process used [] [*] to make filters .*

ARG0: Workers

REL: dumped

ARG1: large burlap sacks of the imported material

ARGM-DIR: into a huge bin

*What sector is [*T*-46] stepping forward [*-2] to pick up the slack ? " he asked [*T*-1]*

ARG1: [*T*-46] → What sector

REL: stepping

ARGM-DIR: forward

ARGM-PNC: [*-2] to pick up the slack

*No one wants the U.S. to pick up its marbles and go home , " Mr. Hormats says [*T*-1] .*

ARG1: the U.S.

REL: go

ARGM-DIR: home

*That response annoyed Rep. Markey , House aides said [0] [*T*-1] , and the congressman snapped back that there had been enough studies of the issue and that it was time for action on the matter .*

ARG0: the congressman

REL: snapped

ARGM-DIR: back

ARG1: that there had been enough studies of the issue and that it was time for action on the matter

3.2. Locatives (LOC)

Locative modifiers indicate where some action takes place. The notion of a locative is not

restricted to physical locations, but abstract locations are being marked as LOC as well, as ‘[in his speech]-LOC he was talking about ...’.

The percentage of lung cancer deaths among the workers at the West Groton , Mass. , paper factory appears [-1] to be the highest for any asbestos workers studied [*] in Western industrialized countries , he said [0] [*T*-2] .*

ARG1: [*] → any asbestos workers

REL: studied

ARGM-LOC: in Western industrialized countries

*Areas of the factory [*ICH*-2] were particularly dusty where the crocidolite was used [*-8] [*T*-1] .*

ARGM-LOC: [*T*-1] → where

ARG1: [*-8] → the crocidolite

REL: used

*In his ruling , Judge Curry added an additional \$ 55 million [*U*] to the commission 's calculations .*

ARGM-LOC: In his ruling

ARG0: Judge Curry

REL: added

ARG1: an additional \$ 55 million [*U*]

ARG2-to: to the commission 's calculations

3.3. Manner Markers (MNR)

Manner adverbs specify how an action is performed. For example, "works well with others" is a manner. Manner tags should be used when an adverb be an answer to a question starting with ‘how?’.

*Among 33 men who [*T*-4] worked closely with the substance, 28 [*ICH*-1] have died -- more than three times the expected number.*

ARG0: [*T*-4] → who -> 33 men

REL: worked

ARGM-MNR: closely

ARG1-with: with the substance

Workers dumped large burlap sacks of the imported material into a huge bin, poured in cotton and acetate fibers and mechanically mixed the dry fibers in a process used [] [*] to make filters.*

ARG0: Workers

ARGM-MNR: mechanically

REL: mixed

ARG1: the dry fibers

ARGM-LOC: in a process used [*] [*] to make filters

The next morning, with a police escort, busloads of executives and their wives raced to the Indianapolis Motor Speedway , [-1] unimpeded by traffic or red lights .*

ARGM-TMP: The next morning

ARGM-MNR: with a police escort

ARG0: busloads of executives and their wives
REL: raced
ARG1-to: to the Indianapolis Motor Speedway
ARGM-ADV: [*-1] unimpeded by traffic or red lights

3.4. Temporal markers (TMP)

Temporal ArgMs show when an action took place, such as "in 1987", "last Wednesday", "soon" or "immediately". Also included in this category are adverbs of frequency (eg. often always, sometimes (with the exception of 'never', see NEG below), adverbs of duration (for a year/in an year), order (eg. first), and repetition (eg. again)..

A form of asbestos once used [] [*] to make Kent cigarette filters has caused a high percentage of cancer deaths among a group of workers exposed [*] to it more than 30 years ago , researchers reported [0] [*T*-1] .*

ARG1: [*] -> A form of asbestos
ARGM-TMP: once
REL: used
ARG2-PNC: [*] to make Kent cigarette filters

Four of the five surviving workers have asbestos-related diseases, including three with recently diagnosed cancer.

ARGM-TMP: recently
REL: diagnosed
ARG2: cancer

3.5. Extent Markers (EXT)

ArgM-EXT indicate the amount of change occurring from an action, and are used mostly for

- numerical adjuncts like "(raised prices) by 15%",
- quantifiers such as "a lot"
- and comparatives such as "(he raised prices) more than she did."

*PS of New Hampshire shares closed yesterday at \$ 3.75 [*U*], off 25 cents, in New York Stock Exchange composite trading .*

ARG1: PS of New Hampshire shares
REL: closed
ARGM-TMP: yesterday
ARGM-EXT: at \$ 3.75 [*U*], off 25 cents,
ARGM-LOC: in New York Stock Exchange composite trading

*``An active 55-year-old in Boca Raton may care more about Senior Olympic games, while a 75-year-old in Panama City may care more about a seminar on health, '' she says [*T*-1].*

ARG0: An active 55-year-old in Boca Raton
ARGM-MOD: may
REL: care
ARGM-EXT: more
ARG1-about: about Senior Olympic games
ARGM-ADV: while a 75-year-old in Panama City may care more about a seminar on health

Rep. Jerry Lewis , a conservative Californian , added a provision of his own, intended [] to assist Bolivia , and the Senate then broadened the list further by*

[-1] including all countries in the U.S. Caribbean Basin initiate as well as the Philippines - [*-1] backed [*] by the powerful Hawaii Democrat Sen. Daniel Inouye .*

ARG0: the Senate

ARGM-TMP: then

REL: broadened

ARG1: the list

ARGM-EXT: further

ARGM-MNR: by [*-1] including all countries in the U.S. Caribbean Basin initiate as well as the Philippines

ARGM-ADV: [*-1] backed [*] by the powerful Hawaii Democrat Sen.

Daniel Inouye

3.6. Reciprocals (REC)

These include reflexives and reciprocals such as *himself, itself, themselves, together, each other, jointly, both*, which refer back to one of the other arguments.

*But voters decided that if the stadium was such a good idea someone would build it himself, and rejected it 59% to 41% [*U*].*

ARGM-ADV: if the stadium was such a good idea

ARG0: someone

ARGM-MOD: would

REL: build

ARG1: it

ARGM-REC: himself

*But while history can suggest what [*T*-1] is reasonable [0] [*] to expect [*T*-2] there 's no guarantee that the past will repeat itself.*

ARG1: the past

ARGM-MOD: will

REL: repeat

ARGM-REC: itself

3.7. Markers of secondary predication (PRD)

These are used to show that an adjunct of a predicate is in itself capable of carrying some predicate structure.

Typical examples include

- Resultatives: as in ‘The boys pinched them dead’ or

‘She kicked the locker lid [*] shut

- Depictives

- ‘as’-phrases, e.g. ‘supplied as security in the transaction’ and other cases of secondary predication

Pierre Vinken , 61 years old , will join the board as a nonexecutive director Nov. 29 .

ARG0: Pierre Vinken , 61 years old ,

ARGM-MOD: will
REL: join
ARG1: the board
ARGM-PRD: as a nonexecutive director
ARGM-TMP: Nov. 29

Prior to his term , a teacher bled to death in the halls , [-1] stabbed [*-2] by a student .*ARGM-TMP: *Prior to his term*

ARG1: a teacher
REL: bled
ARGM-PRD: to death
ARGM-LOC: in the halls
ARGM-ADV: [*-1] stabbed [*-2] by a student

*This wage inflation is bleeding the NFL dry, the owners contend [*T*-1] .*

ARG0: This wage inflation
REL: bleeding
ARG1: the NFL
ARGM-PRD: dry

[-2] Glamorous and pure-voiced as ever , Ms. Collins sang Joni Mitchell 's `` For Free " -- about an encounter with a street-corner clarinetist , to which Mr. Stoltzman contributed a clarinet obligatto [*T*-1] -- and Mr. Douglas 's lush setting of a Gaelic blessing , `` Deep Peace . "*

ARGM-PRD: [*-2] Glamorous and pure-voiced as ever
ARG0: Ms. Collins
REL: sang
ARG1: Joni Mitchell 's `` For Free " -- about an encounter with a street-corner clarinetist , to which Mr. Stoltzman contributed a clarinet obligatto [*T*-1] -- and Mr. Douglas 's lush setting of a Gaelic blessing , `` Deep Peace

3.8. Purpose clauses (PNC: purpose, not cause)

Purpose clauses are used to show the motivation for some action. Clauses beginning with "in order to" are canonical purpose clauses.

More than a few CEOs say [0] the red-carpet treatment tempts them to return to a heartland city for future meetings .

ARG1: them
REL: return
ARG4-to: to a heartland city
ARGM-PNC: for future meetings

*In a disputed 1985 ruling , the Commerce Commission said [0] Commonwealth Edison could raise its electricity rates by \$ 49 million [*U*] [*-1] to pay for the plant .*

ARG0: Commonwealth Edison
ARGM-MOD: could
REL: raise
ARG1: its electricity rates
ARG2-by: by \$ 49 million [*U*]

ARGM-PNC: [*-1] to pay for the plant

3.9. Cause clauses (CAU)

Similar to "Purpose clauses", these indicate the reason for an action. Clauses beginning with "because" or "as a result of" are canonical cause clauses. Also questions starting with 'why':

Pro-forma balance sheets clearly show why Cray Research favored the spinoff
[*T*-1].

ARGM-CAU: [*T*-1] → why

ARG0: Cray Research

REL: favored

ARG1: the spinoff

However, five other countries -- China, Thailand, India, Brazil and Mexico -- will remain on that so-called priority watch list as a result of an interim review, U.S. Trade Representative Carla Hills announced [0] [*T*-1].

ARGM-DIS: However

ARG1: five other countries -- China, Thailand, India, Brazil and Mexico

--

ARGM-MOD: will

REL: remain

ARG3-on: on that so-called priority watch list

ARGM-CAU: as a result of an interim review

3.10. Discourse Markers (DIS)

These are markers which connect a sentence to a preceding sentence.

Examples of discourse markers are: *also, however, too, as well, but, and, as we've seen before, instead, on the other hand, for instance, etc.*

Note that conjunctions such as *but* or *and* are only marked in the beginning of the sentence.

But for now, they're looking forward to their winter meeting -- Boca in February.

ARGM-DIS: But

ARGM-TMP: for now

ARG0: they

REL: [looking] [forward]

ARG1-to: to their winter meeting -- Boca in February

Do not mark *and, or, but*, when they connect two clauses in the same sentence.

Another type of discourse markers includes vocatives, which are marked as VOC in Treebank:

Treebank annotation:

(S (NP-VOC Kris),

(NP-SBJ *)

(VP go

(ADVP-DIR home)))

Propbank annotation:

ArgM-DIS: Kris
Rel: go
Arg0: [*]
ArgM-DIR: home

Vocative NPs in imperative sentences as shown above should not be tagged as chains, i.e. Arg0: [*] -> Kris. in order to make annotation consistent with other examples of Vocative NPs, which do not include traces:

I ai n't kidding you, Vince
ArgM-DIS: Vince
Rel: kidding
Arg0: I
Arg1: you
ArgM-NEG: n't

And, finally, the class of Discourse markers includes interjections such as 'oh my god' 'ah', and 'damn'

I might point out that your inability to report to my office this morning has not ah limited my knowledge of your activities as you may have hoped.
ArgM-DIS: ah
Rel: limited
Arg1: my knowledge of your activities
Arg0: your inability to report to my office this morning
ArgM-ADV: as you may have hoped

3.11. Adverbials (ADV)

These are used for syntactic elements which clearly modify the event structure of the verb in question, but which do not fall under any of the headings above.

1. Temporally related (modifiers of events)
Treasures are just lying around, waiting to be picked up
2. Intensional (modifiers of propositions)
Probably, possibly
3. Focus-sensitive
Only, even
4. Sentential (evaluative, attitudinal, viewpoint, performatives)
Fortunately, really, legally, frankly speaking,
clauses beginning with 'given that', 'despite', except for, 'if'

As opposed to ArgM-MNR, which modify the verb, ARGM-ADVs usually modify the entire sentence.

In some cases, modifiers like 'happily' can be ambiguous between MNR and ADV interpretations, as shown below:

She sang happily.

ArgM-MNR: happily

Happily, she sang. (paraphrasable as 'I am happy that she sang')

ArgM-ADV: happily

3.12. Modals (MOD)

Modals are: *will, may, can, must, shall, might, should, could, would*. "Phrasal modals" such as "going (to)", "have (to)" and "used (to)" are also included, although unlike the regular modals, these are also annotated as verbs in their own right, where they take their own Negation and Adverbial markers, but not any numbered arguments. Thus, in the sentence "John does not have to run", "have" is a modal adjunct of "run", but "not" is a negation adjunct of "have", and not of "run".

3.13. Negation (NEG)

This tag is used for elements such as "not", "n't", "never", "no longer" and other markers of negative sentences. Negation is an important notion for Propbank annotation; therefore, all markers which indicate negation should be marked as NEG. For example, when annotating adverbials like 'never', which could be marked as either TMP or NEG, the NEG tag should be used.

3.14. Stranded (STR)

Stranded tag is not a modifier tag, meaning that it does not mark a certain type of modifier and does not appear in the final, or released version, of Propbank annotations. The use of the tag is motivated by technical reasons. Specifically, STR tags are used when the tool doesn't allow you to combine everything that should be one argument. There are typically two kinds of stranded situations. First, an argument is a chain that points to something that cannot be selected as one node, e.g.:

Arg1: *1 -> *2 -> that -> [X] [Y] [Z]

Second, the argument involves concatenating a chain with something else, e.g.:

Arg1: [* -> 0 -> X] [Y] [Z]

In both cases, the chain should be constructed, which refers to the main element from the concatenated constituent. This chain is labeled as a numbered argument ArgN. The remaining constituents are being tagged as ArgN-STR:

*[Two members] [of the group] wanted *1 to continue *2 running.*

Ideally:

Arg0: *2 -> *1 -> [Two members][of the group]
rel: running

In practice:

Arg0: *2 -> *1 -> [Two members]
Arg0-STR: [of the group]
rel: running

Likewise,

*John wanted *1 to appear *2 to be funny.*

Ideally:

Arg1: [* -> John][*2 to be funny]
rel: appear

In practice:

Arg1: * -> John
Arg1-STR: [*2 to be funny]
rel: appear

Another example:

Italy 's Foreign Ministry said [0] it is investigating exports to the Soviet Union by an

Ing . C. Olivetti & Co. subsidiary called [] OCN-PPL that [*T*-1] makes numerically controlled machine tools.*

ARG0-STR: called [*] OCN-PPL
ARG0: [*T*-1] -> that -> an Ing . C. Olivetti & Co. subsidiary
rel: makes
ARG1: numerically controlled machine tools

3.15. Special cases

3.15.1. Modification in complex clauses

When annotating ArgMs, please make sure that they modify the verb being annotated, and not another verb in the sentence. For example, in the following sentence, ArgM-TMP modifies the verb *come*, but not *ask*:

[About 5 years ago]-TMP, Handley came to ask me if he could see the tattered register.

In the next example, ArgM-TMP modifies the verb *interested*, but not *awakened*.
It awakened [RNR] and, [for a moment]-TMP, interested [RNR] him

3.15.2. Multiple modifiers with the same tag

As a general rule, try to follow Treebank constituency, whenever it is possible. This means that ArgMs should not be decomposed or concatenated.

For example, if there is no constituent in Treebank which contains *yesterday* and *at 5 pm*, then they should be analyzed as separate ArgMs:

He was in the library yesterday at 5pm

ArgM-TMP: yesterday
ArgM-TMP: at 5 pm

NOT:
ArgM-TMP: [yesterday][at 5 pm]

If there is a constituent which contains both modifiers, then such constituent should be selected.

The following examples illustrate two ArgM-Adv and ArgM-DIS in the same sentence:

Of all places in Poland Andrei hated Lublin the most.

ArgM-Adv: Of all places in Poland
ArgM-Adv: the most
NOT:
ArgM-ADV: [Of all places in Poland][the most]

Oh, well, you can't really blame Lolotte.

DIS: Oh
DIS: well
NOT:
ArgM-DIS: [Oh]{well}

Another consequence of the rule 'follow the syntax' is that in the following sentence, 'all three' should be marked as ArgM-ADV, rather than concatenated with the argument 'they'.

They were [all three]-ADV bent over a shabby riding boot.

Rel: bent
Arg1: they
ArgM-ADV: all three

NOT:
Rel: bent
Arg1: [they][all three]

Likewise, annotators should not decompose ArgMs which are analyzed as constituents in Treebank, even if they can be semantically decomposed:

[] To kayo him and maybe or maybe not kill*
ArgM-ADV: maybe or maybe not
Rel: kill

A more complicated situation is when there are two arguments with the same semantic role. This usually happens when the argument is a location or a direction. For example, the Arg2 argument of the verbs 'locate' and 'put' is a location. In the following examples, there are two constituents which could be viewed as locations (e.g. 'at the curb' and 'in front of an apartment house' for the first sentence).

The rule of thumb which was adopted for such cases is that the constituent closest to the verb is marked as a numbered argument, and the other one as ArgM:

I located the car parked [] at the curb in front of an apartment house.*

Arg2: at the curb
ArgM-LOC: in front of an apartment house

She put the slipper neatly by its mate at the foot of the bed

Arg2-LOC: by its mate
ArgM-LOC: at the foot of the bed

3. 15. 3. Multiple Propositions

Multiple propositions are mostly being used in the case of gapping or ellipsis:

John not only went to the store but also__ to the bank.

In the second clause, the verb is missing, but the gap is understood as referring to the verb in the first clause. Sentences of this type are annotated as having two propositions:

Propbank annotation (2 propositions):

Arg1: John
ArgM-Dis: not only
Rel: went
Arg4: to the store

Arg1: John
ArgM-Dis: but also
Rel: went
Arg4: to the bank

Some instances of conjunction can also be annotated as having multiple annotations, but only if it is not possible to annotate it as one proposition.

John went to the store and then to the beach.

In this case, the ArgM-DIS ‘then’ modifies the second conjunct only, which can only be captured by having 2 propositions:

Propbank annotation (2 propositions):

Rel: went
Arg0: John
ArgM-DIR: to the store

Rel: went
Arg0: John
ArgM-DIR: to the beach
ArgM-DIS: then

Do not decompose conjunctions into separate propositions, if a conjoined phrase denotes an argument or a modifier:

John went to the store and to the beach.

Rel: went
Arg0: John
ArgM-DIR: to the store and to the beach

An existentialist is a man who perceives himself only as “esse”, as existence without substance.

Arg0: who -> a man

Rel: perceives

Arg1: himself

ArgM-PRD: only as “esse”, as existence without substance

4. Task 3: Coreference (for empty categories)

One of the most difficult issues that you will confront when annotating coreference is deciding whether a given NP is an argument of the verb, which should therefore be coindexed in a chain, or an implicit or arbitrary argument.

The following examples illustrate typical cases of annotator errors with respect to coreference annotation:

Long-debated proposals [] to simplify the more than 150 civil penalties and make them fairer and easier [0] [*] to administer [*T*-1] are in the House tax bill.*

1. ARG0: [*] -> proposals
2. rel: make
3. ARG1: them
4. ARG2-PRD: fairer and easier [0] [*] to administer [*T*-1]

Indexing for the most part has involved [-2] simply buying [*RNR*-1] and then holding [*RNR*-1] stocks in the correct mix [*] to mirror a stock market barometer, such as Standard & Poor 's 500-stock index, and match its performance.*

1. ARG0: [*-2] -> Indexing
2. ARGM-ADV: simply
3. ARGM-DIS: then
4. rel: holding
5. ARG1: stocks in the correct mix -> [*RNR*-1]

According to these annotations, ‘the proposals are making penalties fairer’, and ‘indexing is holding stocks in the correct mix’, which is certainly not the correct analysis of these sentences.

The correct analysis is shown, where the argument Arg0 is an arbitrary subject (the empty category in understood as arbitrary PRO), which could be substituted by ‘someone’ in these examples.

1. ARG0: [*]
2. rel: make
3. ARG1: them
4. ARG2-PRD: fairer and easier [0] [*] to administer [*T*-1]

1. ARG0: [*-2]
2. ARGM-ADV: simply
3. ARGM-TMP: then
4. rel: holding

5. ARG1: stocks in the correct mix -> [*RNR*-1]

More difficult examples involve ambiguous interpretations, as shown below:

The patent for Interleukin-3 covers materials and methods used [] [*] to make the human blood cell growth factor via recombinant DNA technology.*

Annotator 1:

ARG2: [*] -> [*] -> materials and methods

rel: make

ARG1: the human blood cell growth factor

ARGM-MNR: via recombinant DNA technology

Annotator 2:

Arg0: [*] -> [*]

rel: make

ARG1: the human blood cell growth factor

ARGM-MNR: via recombinant DNA technology

Both annotations are correct interpretations of this sentence, since the trace can either refer to the Arg0 argument (i.e. the agent of making, which is arbitrary or implicit in this case), or the Arg2 argument (which corresponds to the instrument).

The rule adopted for such cases is similar to the rule used for resolution of ambiguous argument labels, and is based on the hierarchy of arguments, as discussed in section 2.2. above. According to this rule, the highest ranked argument should be selected, where Arg0 >> Arg1 >> Arg2 >>... . In this example, the trace should be interpreted as suggested by the annotator 2.

B ERASKINA: EADBko aditzak beren balio sintaktiko/semantikoekin.

ADITZA bere <i>bssekin</i>	abs	erg	dat	ela	ine	ins	soz	abl	ala	KASU SEMANTIKOAK MURRIZTAPENAK	AZPIKATEGORIZATUEN ETA	OSAGAI HAUTAPEN-	BSSAREN BALIO SEMANTIKOA
abestu-DU-1	√	√	+	-	+	+	+	+	-	gaia [-biz]_abs; esperimentatzailea [+biz]_erg			Entitate baten jarduera
abestu-DU-2	-	√	+	-	+	+	+	+	-	esperimentatzailea [+biz]_erg			Entitate baten jarduera
adierazi-DU-1	√	√	+	-	+	+	+	+	-	gaia [-biz; -konkr]_abs; esperimentatzailea_erg			Entitate baten jarduera, adierazpenezkoa
adierazi-DU-2	-	√	+	√	+	+	+	+	-	esperimentatzailea_erg; gai a_konp-ela			Entitate baten jarduera, adierazpenezkoa
afaldu-DU-1	√	√	-	-	+	+	+	-	-	gaia [-biz; +konkr]_abs; esperimentatzailea [+biz]_erg			Entitate baten jarduera
afaldu-DU-2	-	√	-	-	+	+	+	-	-	esperimentatzailea [+biz]_erg			Entitate baten jarduera
agertu-DA-1	√	-	+	-	√	+	+	-	-	gaia_abs; kokapena_ine			Entitate baten kokapena
agertu-DA-2	√	-	+	√	+	+	+	-	-	gaia_konp-ela; kokapena_ine			Entitate baten kokapena
agertu-DA-3	√	-	+	-	+	+	+	√	√	gai ukitua_abs; abiapuntuko kokapena / bidea_abl; helburuko kokapena_ala			Entitate baten kokapen-aldaketa
agertu-DA-4	√√	-	+	-	+	-	+	-	-	gaia [+giz]_abs; egoera_abs			Entitate baten egoera
agertu-DU-5	√	√	+	-	+	+	+	-	-	gaia [-biz]_abs ; esperimentatzailea_erg			Entitate baten jarduera
agertu-DU-6	-	√	+	√	+	+	+	-	-	esperimentatzailea_erg; gaia_konp-ela			Entitate baten jarduera

ADITZA bere <i>bs</i> sekin	abs	erg	dat	ela	ine	ins	soz	abl	ala	KASU SEMANTIKOAK MURRIZTAPENAK	AZPIKATEGORIZATUEN ETA	OSAGAI HAUTAPEN-	BSSAREN BALIO SEMANTIKOA
ahaztu-DA-1	√	-	-	-	+	√	+	-	-	esperimentatzailea [+giz]_abs; gaia [+giz?]_ins			Entitate baten jarduera, erreakziozkoa
ahaztu-ZAIO-2	√	-	√	-	+	+	+	-	-	gaia [-biz?]_abs; esperimentatzailea [+giz]_dat			Entitate baten jarduera, erreakziozkoa
ahaztu-DU-3	√	√	-	-	+	+	+	-	-	gaia_abs; esperimentatzailea [+giz]_erg			Entitate baten jarduera, erreakziozkoa
ahaztu-DU-4	-	√	-	√	+	+	+	-	-	esperimentatzailea [+giz]_erg; gaia_konp-ela			Entitate baten jarduera, erreakziozkoa
aldatu-DA-1	√	-	+	-	+	+	+	√	√	gai ukitua_abs; abiapuntua_abl; helburua_ala			Entitate baten kokapen- edo egoera aldaketa
aldatu-DA-2	√	-	+	-	+	√	+	-	-	gai ukitua_abs; kokapena/egoera_ins			Entitate baten kokapen- edo egoera aldaketa
aldatu-DU-3	√	√	+	-	+	+	+	√	√	gai ukitua_abs; kausa_erg; abiapuntua_abl; helburua_ala			Entitate baten kokapen- edo egoera aldaketa
aldatu-DU-4	√	√	+	-	+	√	+	-	-	gai ukitua_abs; kausa_erg; kokapena/egoera_ins			Entitate baten kokapen- edo egoera aldaketa
aldatu-DU-5	√	√	+	-	+	√	+	-	-	gaia_abs; esperimentatzailea [+giz]_erg; gaia_ins			Entitate baten jarduera: entitate bat ordezkatzea
amaitu-DA-1	√	-	+	-	+	+	+	-	-	gai ukitua? [-biz]_abs			Entitate baten egoera-aldaketa: iraupena agortzea
amaitu-DU-2	√	√	+	-	+	+	+	+	+	gai ukitua? [-biz]_abs; kausa_erg			Entitate baten egoera-aldaketa: iraupena agortzea
amaitu-DU-3	-	√	-	-	+	+	√	-	-	kausa_erg; gai ukitu?_soz			Entitate baten egoera-aldaketa: iraupena agortzea
amaitu-DU-4	-	√	-	-	√	+	+	-	-	gai ukitua_erg; helburuko kokapena_ine			Entitate baten kokapen-aldaketa
argitu-DU-1	√	√	+	-	+	+	+	+	-	gai ukitua_abs; kausa_erg			Entitate baten egoera-aldaketa
argitu-DU-2	-	√	+	√	+	+	+	+	-	kausa_erg; gai ukitua_konp-ela			Entitate baten egoera-aldaketa
argitu-DU-3	√	-	-	-	+	+	+	-	-	gai ukitua: <i>egun</i> _abs; [kausa: <i>eguzki</i> _erg]			Entitate baten egoera-aldaketa

ADITZA bere <i>bs</i> sekin	abs	erg	dat	ela	ine	ins	soz	abl	ala	KASU SEMANTIKOAK MURRIZTAPENAK	AZPIKATEGORIZATUEN ETA	OSAGAI HAUTAPEN-	BSSAREN BALIO SEMANTIKOA
argitu-DU-4	-	-	-	-	+	+	+	-	-	[gai ukitua: <i>egun_abs</i>]; [kausa: <i>eguzki_erg</i>]			Entitate baten egoera-aldaketa
aritu-DA-1	√	-	+	-	√	+	+	+	-	esperimentatzailea [+biz]_abs; jarduera_ine			Entitate baten jarduera
aritu-DA-2	√	-	+	-	+	+	+	+	-	esperimentatzailea [+biz]_abs			Entitate baten jarduera
asmatu-DU-1	√	√	+	-	+	+	+	+	-	gaia [-konkr]_abs; esperimentatzailea [+giz]_erg			Entitate baten jarduera
asmatu-DU-2	-	√	+	-	+	+	√	+	-	esperimentatzailea [+giz]_erg			Entitate baten jarduera
asmatu-DU-3	-	√	-	√	+	+	+	-	-	esperimentatzailea [+giz]_erg; gaia_konp-ela			Entitate baten jarduera
atera-DA-1	√	-	+	-	+	+	+	√	√	gai ukitua_abs; abiapuntuko kokapena / bidea_abl; helburuko kokapena_ala			Entitate baten kokapen-aldaketa
atera-DA-2	√	-	-	-	+	+	+	√	-	gai sortua [-konkr]_abs; iturria [-biz]_abl			Entitate baten sortzea
atera-ZAIO-3	√	-	√	-	+	+	+	+	√	gai ukitua_abs; abiapuntuko kokapena_dat; helburuko kokapena_ala			Entitate baten kokapen-aldaketa
atera-ZAIO-4	√	-	√	-	+	+	+	+	-	gai sortua [-konkr]_abs; iturria [+giz]_dat			Entitate baten sortzea
atera-DU-5	√	√	+	-	+	+	+	√	√	kausa_erg; gai ukitua_abs; abiapuntuko kokapena / bidea_abl; helburuko kokapena_ala			Entitate baten kokapen-aldaketa
atera-DU-6	√	√	+	-	+	+	+	+	-	gai sortua [-biz]_abs; iturria [+biz]_erg			Entitate baten sortzea
atera-DU-7	√	-	-	-	+	+	+	-	-	gai sortua: <i>haize_abs</i> ; [iturria: ?_erg]			Entitate baten sortzea
atera-DIO-8	√	√	√	-	+	+	+	+	√	kausa_erg; gai ukitua_abs; abiapuntuko kokapena [+biz]_dat; helburuko kokapena_ala			Entitate baten kokapen-aldaketa
aurkitu-DA-1	√	-	-	-	√	+	+	-	-	gaia_abs; kokapena_ine			Entitate baten kokapena
aurkitu-DU-2	√	√	+	-	√	+	+	-	-	gaia_abs; esperimentatzailea [+giz]_erg; kokapena_ine			Entitate baten kokapena, eta horren

ADITZA bere <i>bs</i> sekin	abs	erg	dat	ela	ine	ins	soz	abl	ala	KASU SEMANTIKOAK MURRIZTAPENAK	AZPIKATEGORIZATUEN ETA	OSAGAI HAUTAPEN-	BSSAREN BALIO SEMANTIKOA
aurkitu-DU-3	√√	√	-	-	+	-	-	-	-	gaia_abs; egoera_abs; esperimentatzailea [+giz]_erg			esperimentazioa Entitate baten egoera, eta horren esperimentazioa
baieztatu-DU-1	√	√	+	-	+	+	+	+	-	gaia [-biz; -konkr]_abs; esperimentatzailea_erg			Entitate baten jarduera, adierazpeneko
baieztatu-DU-2	-	√	+	√	+	+	+	+	-	gaia_konp-ela; esperimentatzailea_erg			Entitate baten jarduera, adierazpeneko
banandu-DA-1	√	-	+	-	+	+	+	√	-	gai ukitua_abs; gai ukitua_abl			Entitate batzuen egoera-aldaketa
banandu-DA-2	√	-	+	-	√	+	+	-	-	gai ukitua_abs; helburuko egoera_ine			Entitate batzuen helburuko egoera
banandu-DU-3	√	√	+	-	+	+	+	√	-	gai ukitua_abs; kausa_erg; gai ukitua_abl			Entitate batzuen egoera-aldaketa
banandu-DU-4	√	√	+	-	√	+	+	-	-	gai ukitua_abs; kausa_erg; helburuko egoera_ine			Entitate batzuen helburuko egoera
barkatu-DU-1	√	√	-	-	+	+	+	-	-	abiapuntua [+giz]_erg; helburua [+giz]_abs			Entitate baten elkarraldatzea
barkatu-DIO-2	-	√	√	-	+	+	+	-	-	abiapuntua [+giz]_erg; helburua [+giz]_dat			Entitate baten elkarraldatzea
barkatu-DIO-3	√	√	√	-	+	+	+	-	-	abiapuntua [+giz]_erg; gaia [-biz]_abs; helburua [+giz]_dat			Entitate baten elkarraldatzea
bazkaldu-DU-1	√	√	-	-	+	+	+	-	-	gaia [-biz; +konkr]_abs; esperimentatzailea [+biz]_erg			Entitate baten jarduera
bazkaldu-DU-2	-	√	-	-	+	+	+	-	-	esperimentatzailea [+biz]_erg			Entitate baten jarduera
besarkatu-DA-1	√	-	-	-	+	+	+	-	-	esperimentatzailea+esperimentatzailea [+giz]_abs			Entitate batzuen jarduera (erreziprokoa)
besarkatu-DU-2	√	√	-	-	+	+	+	-	-	esperimentatzailea [+giz]_erg; esperimentatzailea [+biz]_abs			Entitate batzuen jarduera (erreziprokoa)
bilakatu-DA-1	√√	-	+	-	+	+	+	-	-	gai ukitua_abs; helburuko egoera_abs			Entitate baten egoera-aldaketa
bilakatu-DU-2	√√	√	+	-	+	+	+	-	-	gai ukitua_abs; helburuko egoera_abs; kausa_erg			Entitate baten egoera-aldaketa

ADITZA bere <i>bs</i> sekin	abs	erg	dat	ela	ine	ins	soz	abl	ala	KASU SEMANTIKOAK MURRIZTAPENAK	AZPIKATEGORIZATUEN ETA	OSAGAI HAUTAPEN-	BSSAREN BALIO SEMANTIKOA
bete-DA-1	√	-	+	-	√	+	+	-	-	helburuko kokapena_abs; gai ukitua_ins			Entitate baten kokapen-aldaketa.
bete-DA-2	√	-	+	-	+	+	+	-	-	gaia/gai ukitua [-konkr]_abs			Entitate baten jarduera
bete-DU-3	√	√	+	-	√	+	+	-	-	kausa_erg; helburuko kokapena_abs; gai ukitua_ins			Entitate baten kokapen-aldaketa
bete-DU-4	√	√	-	-	+	+	+	-	-	kausa/esperimentatzailea [+giz]_erg; gai ukitua/gaia [-konkr]_abs			Entitate baten jarduera
bete-DU-5	√	√	-	-	+	+	+	-	-	kausa+ gai ukitua_erg; helburuko kokapena_abs			Entitate baten kokapen-aldaketa [+holist]
bete-DU-6	-	√	-	-	+	+	+	-	-	kausa/esperimentatzailea [+giz]_erg; [gai ukitua/gaia: behar_abs]			Entitate baten jarduera
bisitatu-DU-0	√	√	-	-	+	+	+	-	-	helburuko kokapena [+giz/+lek]_abs; gai ukitua [+giz]_erg			Entitate baten kokapen-aldaketa
dedikatu-DIO-0	√	√	√	-	+	+	+	+	-	gaia [-biz]_abs; abiapuntua [+giz]_erg; helburua [+biz]_dat			Entitate baten elkarraldatzea
deitu-DU-1	√	√	-	-	+	+	+	+	√	helburua [+giz]_abs; esperimentatzailea [+giz]_erg; jarduera_ala			Entitate baten jarduera (norbaiti zuzendua)
deitu-DU-2	√	√	-	-	+	+	+	-	-	jarduera_abs; esperimentatzailea [+giz]_erg			Entitate baten jarduera (norbaiti zuzendua)
deitu-DU-3	-	√	-	-	+	+	+	-	√	esperimentatzailea [+giz]_erg; jarduera_ala			Entitate baten jarduera (norbaiti zuzendua)
deitu-DU-4	√	√	-	-	+	+	+	-	-	abiapuntua [+giz]_erg; helburua_abs; ezaugarria_abs			Entitate bati ezaugarria esleitzea
deitu-DIO-5	-	√	-	-	+	+	+	√	√	abiapuntua [+giz]_erg; helburua [+giz]_dat; kokapena_abl; kokapena_ala			Entitate baten jarduera (norbaiti zuzendua), bi entitateen kokapenak azpimarratuta
deitu-DIO-6	√	√	√	-	+	+	+	-	-	abiapuntua [+giz]_erg; helburua_dat; ezaugarria_abs			Entitate bati ezaugarria esleitzea

ADITZA bere <i>bs</i> sekin	abs	erg	dat	ela	ine	ins	soz	abl	ala	KASU SEMANTIKOAK MURRIZTAPENAK	AZPIKATEGORIZATUEN ETA	OSAGAI HAUTAPEN-	BSSAREN BALIO SEMANTIKOA
edan-DU-1	√	√	+	-	+	+	+	+	-	gaia [-biz]_abs; esperimentatzailea [+biz]_erg			Entitate baten jarduera
edan-DU-2	-	√	-	-	+	+	+	+	-	esperimentatzailea [+biz]_erg			Entitate baten jarduera
egin-DU-0	√	√	+	-	+	+	+	+	-	gai sortua [-biz]_abs; iturria [+biz]_erg			Entitate baten sortzea
egokitu-DA-1	√	-	+	-	+	+	+	+	√	gai ukitua_abs; helburuko egoera_ala			Entitate baten egoera-aldaketa
egokitu-DA-2	√	-	+	-	+	+	+	-	-	gaia_abs			Ezusteko gertakizuna
egokitu-ZAIO-3	√	-	√	-				-	-	ezaugarria_abs; helburua_dat			Entitate bati ezaugarria esleitzea
egokitu-DU-4	√	√	+	-	+	+	+	+	√	kausa_erg; gai ukitua_abs; helburuko egoera_ala			Entitate baten egoera-aldaketa
egokitu-DIO-5	√	√	√	-	+	+	+	-	-	abiapuntua_erg; ezaugarria_abs; helburua_dat			Entitate bati ezaugarria esleitzea
egon-DA-1	√	-	-	-	√	+	+	-	-	gaia_abs; kokapena_ine			Entitate baten kokapena
egon-DA-2	√√	-	-	-	-	+	+	-	-	gaia_abs; egoera_abs			Entitate baten egoera
ehizatu-DU-1	√	√	-	-	+	+	+	-	-	gaia [+biz]_abs; esperimentatzailea [+biz]_erg			Entitate baten jarduera
ehizatu-DU-2	-	√	-	-	+	+	+	-	-	esperimentatzailea [+biz]_erg			Entitate baten jarduera
ekarri-DU-1	√	√	+	-	+	+	+	√	√	gai ukitua_abs; kausa_erg; abiapuntuko bidea_abl; helburuko kokapena	kokapena	/	Entitate baten kokapen-aldaketa (hona=deiktikoa)
ekarri-DU-2	√	√	+	-	+	+	+	-	-	gai sortua [-konkr]_abs; iturria_erg			Entitate baten sortzea
ekarri-DU-3	√	√	-	-	√	+	+	-	-	edukia_abs; edukitzailea_erg; kokapena_ine			Entitate baten edukitzea
ekarri-DU-4	√	√	-	-	+	+	+	-	√	gai ukitua_abs; kausa [+giz]_erg; helburuko egoera_ala			Entitate baten egoera -aldaketa
elkartu-DA-1	√	-	-	-	+	+	√	+	-	gai ukitua_abs; gai ukitua_so			Entitate batzuen egoera-aldaketa

ADITZA bere <i>bs</i> sekin	abs	erg	dat	ela	ine	ins	soz	abl	ala	KASU SEMANTIKOAK MURRIZTAPENAK	AZPIKATEGORIZATUEN ETA	OSAGAI HAUTAPEN-	BSSAREN BALIO SEMANTIKOA
elkartu-DA-2	√	-	+	-	√	+	+	-	-	gai ukitua_abs; helburuko egoera_ine			Entitate batzuen helburuko egoera
elkartu-ZAIO-3	√	-	√	-	+	+	+	-	-	gai ukitua_abs; gai ukitua_dat			Entitate batzuen egoera-aldaketa
elkartu-DU-4	√	√	+	-	+	+	√	+	-	gai ukitua_abs; kausa_erg; gai ukitua_soz			Entitate batzuen egoera-aldaketa
elkartu-DU-5	√	√	+	-	√	+	+	-	-	gai ukitua_abs; kausa_erg; helburuko egoera_ine			Entitate batzuen helburuko egoera
eman-DU-1	√	√	√	-	+	+	+	+	-	gaia_abs; abiapuntua_erg; helburua_dat			Entitate baten elkarraldatzea
eman-DU-2	√	√	-	-	√	+	+	-	-	denbora_abs; esperimentatzailea [+giz]_erg; jarduera_ine			Entitate baten jarduera; denbora-epe mugatu baten barruan
eman-DU-3	-	-	-	√	-	-	-	-	-	gaia_konp-ela			Entitate baten jarduera; gertakizun baten inguruko ustea adierazten duena.
entzun-DU-1	√	√	+	-	+	+	+	+	-	esperimentatzailea [+giz]_erg; gaia_abs			Entitate baten jarduera
entzun-DU-2	-	√	+	-	+	+	+	-	-	esperimentatzailea [+giz]_erg			Entitate baten jarduera
entzun-DU-3	-	√	+	√	+	+	+	+	-	esperimentatzailea [+giz]_erg; gaia_ela-konp			Entitate baten jarduera
erabili-DU-0	√	√	+	-	+	+	+	-	-	gaia_abs; esperimentatzailea [+giz]_erg			Entitate baten jarduera
eragin-DU-1	√	√	+	-	√	+	+	-	-	gai sortua [-biz]_abs; iturria_erg, kokapena_ine			Entitate baten sortzea
eragin-DIO-2	-	√	√	-	+	+	+	-	-	iturri bitartekoa?_erg; [gai sortua: sentipen bat?_abs]; iturria [+biz]?_dat			Entitate baten sortzea
eraman-DU-1	√	√	+	-	+	+	+	√	√	gai ukitua_abs; kausa_erg; abiapuntuko kokapena / bidea_abl; helburuko kokapena_ala			Entitate baten kokapen-aldaketa
eraman-DU-2	√	√	-	-	√	+	+	-	-	edukia [-biz]_abs; edukitzailea_erg; kokapena_ine			Entitate bat edukitzea
eraman-DU-3	√	√	-	-	√	+	+	-	-	denbora_abs; esperimentatzailea_erg; jarduera_ine			Entitate baten jarduera; denbora-epe baten

ADITZA bere <i>bs</i> sekin	abs	erg	dat	ela	ine	ins	soz	abl	ala	KASU SEMANTIKOAK MURRIZTAPENAK	AZPIKATEGORIZATUEN ETA	OSAGAI HAUTAPEN-	BSSAREN BALIO SEMANTIKOA
													barruan
erantzun-DU-1	-	√	+	√	+	+	+	+	-	esperimentatzailea_erg ; gaia_konp-ela;			Entitate baten jarduera
erantzun-DU-2	-	√	+	-	+	+	+	+	-	esperimentatzailea_erg; helburua_dat			Entitate baten jarduera
Erre-DA-1	√	-	+	-	+	+	+	+	-	gai ukitua_abs			Entitate baten egoera-aldaketa
erre-DU-2	√	√	+	-	+	+	+	+	-	gai ukitua_abs; kausa_erg			Entitate baten egoera-aldaketa
erre-DU-3	-	√	-	-	+	+	+	-	-	esperimentatzailea [+giz]_erg; [tabako_abs]			Entitate baten jarduera
erreparatu-DU-1	√	√	-	-	+	+	+	-	-	esperimentatzailea [+giz]_erg; gaia [-biz]_abs;			Entitate baten jarduera; esfortzuaz
erreparatu-DU-2	-	√	-	√	+	+	+	-	-	esperimentatzailea [+giz]_erg; gaia_konp-ela			Entitate baten jarduera; esfortzuaz
erreparatu-DU-3	-	√	-	-	√	+	+	-	-	esperimentatzailea [+giz]_erg; gaia [-biz]_ine			Entitate baten jarduera; esfortzuaz?
erreparatu-DIO-4	-	√	√	-	+	+	+	-	-	esperimentatzailea [+giz]_erg; gaia_dat			Entitate baten jarduera; esfortzuaz?
esan-DU-1	√	√	+	-	+	+	+	+	-	esperimentatzailea [+giz]_erg; gaia [-konkr]_abs			Entitate baten jarduera, adierazpenezkoa
esan-DU-2	-	√	+	√	+	+	+	+	-	esperimentatzailea [+giz]_erg; gaia_ela-konp			Entitate baten jarduera, adierazpenezkoa
esan-DIO-3	√	√	√	-	+	+	+	-	-	abiapuntua [+giz]_erg; helburua_dat; ezaugarria_abs			Entitate bati ezaugarria esleitzea
eskaini-DU-0	√	√	√	-	+	+	+	+	-	gaia_abs; abiapuntua_erg; helburua_dat			Entitate baten elkarraldatzea
eskatu-DU-1	√	√	+	-	+	+	+	+	-	gaia_abs; esperimentatzailea_erg			Entitate baten jarduera
eskatu-DU-2	-	√	+	√	+	+	+	+	-	esperimentatzailea_erg; gaia_konp-ela			Entitate baten jarduera
etorri-DA-1	√	-	+	-	+	+	+	√	√	gai ukitua_abs; abiapuntuko kokapena / bidea_abl; helburuko kokapena_ala			Entitate baten kokapen-aldaketa (hona=deiktikoa)

ADITZA bere <i>bs</i> sekin	abs	erg	dat	ela	ine	ins	soz	abl	ala	KASU SEMANTIKOAK MURRIZTAPENAK	AZPIKATEGORIZATUEN ETA	OSAGAI HAUTAPEN-	BSSAREN BALIO SEMANTIKOA
etorri-DA-2	√	-	-	-	+	+	+	√	-	gai sortua [-konkr]_abs; iturria [-biz]_abl			Entitate baten sortzea
etorri-DA-3	√	-	-	-	√	+	+	-	-	edukia [-biz]_abs; edukitzailea [-biz]_ine			Entitate bat edukitzea
etorri-DA-4	√√	-	-	-	+	-	-	-	-	gaia_abs; ezaugarria_abs			Entitate baten deskribapena
etorri-ZAIO-5	√	-	√	-	+	+	+	+	-	gai sortua [-biz]_abs; iturria [+biz]_dat			Entitate baten sortzea
eutsi-DU-1	√	√	-	-	+	+	+	+	-	gaia [-biz]_abs; esperimentatzailea_erg			Entitate baten jarduera; esfortzuzkoa
eutsi-DIO-2	-	√	√	-	+	+	+	+	-	esperimentatzailea_erg; gaia_dat			Entitate baten jarduera; esfortzuzkoa
existitu-DA-1	√	-	-	-	+	-	-	-	-	gaia_abs			Entitate baten izatea/existitzea
existitu-DU-2	-	√	-	-	+	-	-	-	-	gaia_erg			Entitate baten izatea/existitzea
ezkondu-DA-1	√	-	+	-	+	+	√	-	+	gai ukitua [+giz]_abs; gai ukitua [+giz]_soz			Bi entitateren egoera-aldaketa
ezkondu-DU-2	√	√	+	-	+	+	√	-	-	gai ukitua [+giz]_abs; kausa [+giz]_erg; gai ukitua [+giz]_soz			Bi entitateren egoera-aldaketa
flotatu-DU-0	-	√	-	-	√	+	+	-	-	gaia_erg; kokapena_ine			Entitate baten kokapena
gaintitu-DU-0	√	√	-	-	+	+	+	+	-	gaia_abs; esperimentatzailea_erg			Entitate baten jarduera
gertatu-DA-1	√	-	+	-	√	+	+	-	-	gaia_abs; kokapena_ine			Entitate baten kokapena
gertatu-DA-2	√√	-	+	-	+	+	+	-	-	gaia [-konkr]_abs; egoera_abs			Entitate baten egoera
gosaldu-DU-0	-	√	-	-	+	+	+	-	-	esperimentatzailea [+biz]_erg			Entitate baten jarduera
grabatu-DU-1	√	√	+	-	+	+	+	+	+	gaia [-biz]_abs; esperimentatzailea [+biz]_erg			Entitate baten jarduera
grabatu-DU-2	-	√	+	-	+	+	+	+	+	esperimentatzailea [+biz]_erg			Entitate baten jarduera

ADITZA bere bssekin	abs	erg	dat	ela	ine	ins	soz	abl	ala	KASU AZPIKATEGORIZATUEN SEMANTIKOAK ETA MURRIZTAPENAK	OSAGAI HAUTAPEN-	BSSAREN BALIO SEMANTIKOA
hartu-DU-1	√	√	+	-	+	+	+	+	-	gai_ukitua_abs; kausa [+giz]_erg; abiapuntuko kokapena_abl		Entitate baten kokapen-aldaketa
hartu-DU-2	√	√	-	-	√	+	+	-	-	edukia_abs; edukitzailea_erg; kokapena_ine		Entitate bat edukitzea
haserretu-DA-1	√	-	-	-	+	+	+	-	-	gai ukitua [+biz]_abs; kausa_soz		Entitate baten egoera (psikologikoa)ren aldaketa
haserretu-DU-2	√	√	-	-	+	+	+	-	-	gai ukitua [+biz]_abs; kausa_erg		Entitate baten egoera (psikologikoa)ren aldaketa
hasi-DA-1	√	-	+	-	+	+	+	+	-	gai ukitua [-biz, -konkr]_abs		Entitate baten egoera-aldaketa; iraupenaren hasiera
hasi-DA-2	√	-	-	-	√	+	+	-	-	esperimentatzailea [+giz]_abs; jarduera_ine		Entitate baten jardueraren hasiera
hasi-DA-3	√	-	-	-	+	+	√	-	-	esperimentatzailea [+giz]_abs; jarduera_soz		Entitate baten jardueraren hasiera
hasi-DU-4	√	√	-	-	+	+	+	+	-	gai ukitua [-biz, -konkr]_abs; kausa [+giz]_erg		Entitate baten egoera-aldaketa; iraupenaren hasiera
hautatu-DU-0	√	√	-	+	+	+	+	+	-	gaia_abs; esperimentatzailea [+giz]_erg		Entitate baten jarduera
hautsi-DA-1	√	-	+	-	+	+	+	+	-	gai ukitua [-biz]_abs		Entitate baten egoera-aldaketa
hautsi-DU-2	√	√	+	-	+	+	+	+	-	gai ukitua [-biz]_abs; kausa_erg		Entitate baten egoera-aldaketa
hazi-DA-1	√	-	+	-	+	+	+	-	-	gai ukitua_abs		Entitate baten egoera-aldaketa
hazi-DU-2	√	√	+	-	+	+	+	-	-	gai ukitua_abs; kausa_erg		Entitate baten egoera-aldaketa
hil-DA-1	√	-	+	-	+	+	+	-	-	gai ukitua [+biz]_abs		Entitate baten egoera-aldaketa
hil-DU-2	√	√	+	-	+	+	+	-	-	kausa_erg; gai ukitua [+biz]_abs		Entitate baten egoera-aldaketa
hornitu-DA-1	√	-	+	-	+	√	+	+	-	abiapuntuko kokapena_abs; gai ukitua_ins		Entitate baten kokapen-aldaketa

ADITZA bere <i>bs</i> sekin	abs	erg	dat	ela	ine	ins	soz	abl	ala	KASU SEMANTIKOAK MURRIZTAPENAK	AZPIKATEGORIZATUEN ETA	OSAGAI HAUTAPEN-	BSSAREN BALIO SEMANTIKOA
hornitu-DU-2	√	√	+	-	+	√	+	-	-	abiapuntuko kokapena _abs; kausa [+giz]_erg; gai ukitua_ins			Entitate baten kokapen-aldaketa
hornitu-DU-3	√	√	-	-	+	+	+	-	-	abiapuntuko kokapena_abs; kausa+gai ukitua_erg			Entitate baten kokapen-aldaketa
hustu-DA-1	√	-	+	-	+	+	+	-	-	helburuko kokapena_abs			Entitate baten kokapen-aldaketa
hustu-DU-2	√	√	+	-	+	+	+	-	-	helburuko kokapena _abs; kausa_erg			Entitate baten kokapen-aldaketa
hustu-DU-3	√	√	+	-	+	+	+	+	+	gai ukitua [-biz; +konkr]_abs; kausa [+giz]_erg			Entitate baten kokapen-aldaketa
igo-DA-1	√	-	+	-	+	+	+	√	√	gai ukitua_abs; abiapuntua / bidea_abl; helburua_ala			Entitate baten egoera- edo kokapen-aldaketa (norabide batean)
igo-DA-2	√√	-	+	-	+	+	+	-	-	gai ukitua_abs; neurria_abs			Entitate baten egoera- edo kokapen-aldaketa (norabide batean)
igo-DU-3	√	√	+	-	+	+	+	√	√	gai ukitua_abs; kausa_erg; abiapuntua / bidea_abl; helburua_ala			Entitate baten egoera- edo kokapen-aldaketa (norabide batean)
igo-DU-4	√√	√	+	-	+	+	+	-	-	gai ukitua_abs; kausa_erg; neurria_abs			Entitate baten egoera- edo kokapen-aldaketa (norabide batean)
igo-DU-5	√	√	-	-	+	+	+	+	-	bidea_abs; gai ukitua [+biz]_erg			Entitate baten egoera- edo kokapen-aldaketa (norabide batean)
ikasi-DU-1	√	√	+	-	+	+	+	+	-	jarduera_abs; experimentatzailea [+biz]_erg			Entitate baten jarduera
ikasi-DU-2	-	√	+	-	+	+	+	-	-	[jarduera:batxillergo/ikasgai_abs]; experimentatzailea [+biz]_erg			Entitate baten jarduera
ikasi-DU-3	-	√	-	√	+	+	+	+	-	experimentatzailea [+biz]_erg; gaia_konp-ela			Entitate baten jarduera
ikasi-DU-4	-	√	-	-	√	+	+	-	-	experimentatzailea [+biz]_erg; jarduera_ine			Entitate baten jarduera
ikusi-DU-1	√	-	+	-	+	+	+	+	-	experimentatzailea [+giz]_erg; gaia_abs			Entitate baten jarduera

ADITZA bere <i>bs</i> sekin	abs	erg	dat	ela	ine	ins	soz	abl	ala	KASU SEMANTIKOAK MURRIZTAPENAK	AZPIKATEGORIZATUEN ETA	OSAGAI HAUTAPEN-	BSSAREN BALIO SEMANTIKOA
ikusi-DU-2	-	√	-	-	+	+	+	+	-	esperimentatzailea [+giz]_erg			Entitate baten jarduera
ikusi-DU-3	-	√	-	√	+	+	+	+	-	esperimentatzailea [+giz]_erg; gaia_ela-konp			Entitate baten jarduera
irakin-DU-0	-	√	-	-	+	+	+	-	-	gai ukitua [-biz, +konkr]_erg			Entitate baten egoera-aldaketa
irakurri-DU-1	√	√	+	-	+	+	+	+	-	gaia [-biz]_abs; esperimentatzailea [+giz]_erg			Entitate baten jarduera
irakurri-DU-2	-	√	+	-	+	+	+	+	-	esperimentatzailea [+giz]_erg			Entitate baten jarduera
irakurri-DU-3	-	√	+	√	+	+	+	+	-	esperimentatzailea [+giz]_erg; gaia_konp-ela			Entitate baten jarduera
iraun-DU-1	√	√	+	-	+	+	+	-	-	denbora_abs; esperimentatzailea_erg			Entitate baten jarduera; epe batean
iraun-DU-2	-	√	+	-	√	+	+	-	-	esperimentatzailea_erg; denbora_ine			Entitate baten jarduera; epe batean
iraun-DU-3	-	√	+	-	+	√	+	-	-	esperimentatzailea_erg; denbora_ins			Entitate baten jarduera; epe batean
iritzi-DA-1	√	-	+	-	+	+	+	√	√	gai ukitua_abs; abiapuntuko kokapena / bidea_abl; helburuko kokapena_ala			Entitate baten kokapen-aldaketa
iritzi-DU-2	√	√	+	-	+	+	+	+	-	gai ukitua_erg; helburua_abs			Entitate baten kokapen-aldaketa
isildu-DA-1	√	-	+	-	+	+	+	-	-	gai ukitua_abs			Entitate baten egoera-aldaketa
isildu-DU-2	√	√	+	-	+	+	+	-	-	gai ukitua_abs; kausa_erg			Entitate baten egoera-aldaketa
isuri-DA-1	√	-	-	-	+	+	+	√	-	gai sortua [-biz]_abs; iturria [-biz]_abl			Entitate baten sortzea
isuri-DA-2	√	-	+	-	+	+	+	√	√	gai ukitua [-biz]_abs; abiapuntuko kokapena / bidea_abl; helburuko kokapena_ala			Entitate baten kokapen-aldaketa
isuri-DU-3	√	√	-	-	+	+	+	-	-	gai sortua [-biz]_abs; iturria_erg			Entitate baten sortzea

ADITZA bere <i>bs</i> sekin	abs	erg	dat	ela	ine	ins	soz	abl	ala	KASU AZPIKATEGORIZATUEN SEMANTIKOAK ETA OSAGAI MURRIZTAPENAK HAUTAPEN-	BSSAREN BALIO SEMANTIKOA
isuri-DU-4	√	√	+	-	+	+	+	+	√	gai ukitua [-biz]_abs; kausa [+biz]_erg; abiapuntuko kokapena / bidea_abl; helburuko kokapena_ala	Entitate baten kokapen-aldaketa
Izan-DA-1	√	-	-	-	√	+	+	-	-	gaia_abs; kokapena_ine	Entitate baten kokapena
izan-DA-2	√√	-	-	-	+	+	+	-	-	gaia_abs; ezaugarria_abs	Entitate baten deskribapena
izan-DA-3	-	-	-	√	+	+	+	-	-	gaia_ela-konp; ezaugarria_abs	Entitate baten deskribapena
izan-DU-4	√	√	-	-	+	+	+	-	-	edukitzailea_erg; edukia_abs	Entitate bat edukitzea
jaitsi-DA-1	√	-	+	-	+	+	+	√	√	gai ukitua_abs; abiapuntua/bidea_abl; helburua_ala	Entitate baten egoera- edo kokapen-aldaketa (norabide batean)
jaitsi-DA-2	√√	-	+	-	+	+	+	-	-	gai ukitua_abs; neurria_abs	Entitate baten egoera- edo kokapen-aldaketa (norabide batean)
jaitsi-DU-3	√	√	+	-	+	+	+	√	√	gai ukitua_abs; kausa_erg; abiapuntua/bidea_abl; helburua_ala	Entitate baten egoera- edo kokapen-aldaketa (norabide batean)
jaitsi-DU-4	√√	√	+	-	+	+	+	-	-	gai ukitua_abs; kausa_erg; neurria_abs	Entitate baten egoera- edo kokapen-aldaketa (norabide batean)
jaitsi-DU-5	√	√	-	-	+	+	+	+	-	bidea_abs; gai ukitua_erg	Entitate baten egoera- edo kokapen-aldaketa (norabide batean)
jan-DU-1	√	√	+	-	+	+	+	+	-	gaia [-biz]_abs; esperimentatzailea [+biz]_erg	Entitate baten jarduera
jan-DU-2	-	√	-	-	+	+	+	+	-	esperimentatzailea [+biz]_erg	Entitate baten jarduera
Jarri-DA-1	√	-	+	-	√	+	+	-	-	gai ukitua_abs; helburuko kokapena_ine	Entitate baten kokapen-aldaketa
jarri-DA-2	√√	-	+	-	+	+	+	-	-	gai ukitua_abs; helburuko egoera_abs	Entitate baten egoera-aldaketa
jarri-ZAIO-3	√	-	√	-	+	+	+	-	-	gai sortua [-konkr]_abs; iturria [+giz]_dat	Entitate baten sortzea

ADITZA bere <i>bs</i> sekin	abs	erg	dat	ela	ine	ins	soz	abl	ala	KASU SEMANTIKOAK MURRIZTAPENAK	AZPIKATEGORIZATUEN ETA	OSAGAI HAUTAPEN-	BSSAREN BALIO SEMANTIKOA
jarri-DU-4	√	√	+	-	√	+	+	-	-	gai ukitua_abs; kausa_erg; helburuko kokapena_ine			Entitate baten kokapen-aldaketa
jarri-DU-5	√√	√	+	-	+	+	+	-	-	gai ukitua_abs; helburuko egoera_abs; kausa_erg			Entitate baten egoera-aldaketa
jarri-DIO-6	√	√	√	-	+	+	+	-	-	ezaugarria_abs; abiapuntua [+giz]_erg; helburua_dat			Entitate bati ezaugarria esleitzea
jasan-DU-0	√	√	+	-	+	+	+	-	-	gaia_abs; experimentatzailea_erg			Entitate baten jarduera
Jaso-DU-1	√	√	+	+	+	+	+	√	√	kaus_erg; gai ukitua_abs; helburua_ala; abiapuntua/bidea_abl			Entitate baten egoera- edo kokapen-aldaketa (norabide batean)
jaso-DU-2	√	√	+	+	+	+	+	√	-	helburua [+giz]_erg; gaia [-biz]_abs; abiapuntua_abl			Entitate baten elkarraldatzea
jaso-DU-3	√	√	-	+	√	+	+	-	-	edukitzailea_erg; edukia_abs; kokapena_ine			Entitate bat edukitzea
jo-DU-1	√	√	+	-	+	+	+	+	-	gaia [-biz]_abs; experimentatzailea_erg			Entitate baten jarduera: doinu bat ekoiztea
jo-DU-2	-	√	-	-	+	+	+	-	-	experimentatzailea [-giz]_erg			Entitate baten jarduera: doinu bat ekoiztea
jo-DU-3	-	√	+	-	+	+	+	-	-	[gaia: <i>musika_abs</i>]; experimentatzailea_erg			Entitate baten jarduera: doinu bat ekoiztea
jo-DU-4	-	√	-	-	+	+	+	-	-	gaia [-konkr]_abs; [experimentatzailea: <i>kanpai_erg</i>]			Entitate baten jarduera: doinu bat ekoiztea
jo-DU-5	√	√	-	-	+	+	+	-	-	gaia_abs; experimentatzailea_erg			Entitate baten jarduera: kolpekatze-jarduera
jo-DU-6	-	√	-	-	+	+	+	-	-	experimentatzailea_erg			Entitate baten jarduera: kolpekatze-jarduera
jo-DU-7	-	√	+	-	√	+	+	+	-	experimentatzailea [+giz]_erg; gaia [+konkr]_ine			Entitate baten jarduera: kolpekatze-jarduera
jo-DU-8	-	√	-	-	+	+	+	√	√	gai ukitua_abs, abiapuntuko kokapena / bidea_abl; helburuko kokapena_ala			Entitate baten kokapen-aldaketa
jo-DU-9	√√	-	-	-	+	+	+	-	-	abiapuntua [+giz]; ezaugarria_abs; helburua_dat			Entitate bati ezaugarria esleitzea
										gaia [-biz: kolpea]_abs; experimentatzailea_erg;			Entitate baten jarduera: kolpekatze-jarduera

ADITZA bere <i>bs</i> sekin	abs	erg	dat	ela	ine	ins	soz	abl	ala	KASU SEMANTIKOAK MURRIZTAPENAK	AZPIKATEGORIZATUEN ETA	OSAGAI HAUTAPEN-	BSSAREN BALIO SEMANTIKOA
jo-DIO-10	√	√	√	-	+	+	+	+	-	gaia/helburua?_dat			
joan-DA-1	√	-	+	-	+	+	+	√	√	gai ukitua_abs; abiapuntuko kokapena / bidea_abl; helburuko kokapena_ala			Entitate baten kokapen-aldaketa
joan-ZAIO-2	√	-	√	-	+	-	-	+	√	gai ukitua_abs; abiapuntuko kokapena [+biz]_dat; helburuko kokapena_ala			Entitate baten kokapen-aldaketa
joan-ZAIO-3	√	-	√	-	+	+	+	+	-	edukia [-biz, -konkr]_abs; edukitzailea_dat			Entitate baten ezaugarri-agortzea, desagertzea, ez-edukitzea
joan-ZAIO-4	√√	-	√	-	+	+	+	-	-	ezaugarria_abs; gaia/helburua?_dat			Entitate bati ezaugarria esleitzea
jokatu-DA-1	√	-	-	-	+	+	+	-	-	esperimentatzailea [+giz] + esperimentatzailea [+giz]_abs			Entitate batzuen jarduera
jokatu-DU-2	√	√	+	-	+	+	+	-	-	esperimentatzailea [+giz]_erg; jarduera_abs			Entitate baten jarduera
jokatu-DU-3	-	√	-	-	+	+	+	-	-	esperimentatzailea [+giz]_erg; [jarduera: partida_abs]			Entitate baten jarduera
jokatu-DU-4	-	√	-	-	√	+	+	-	-	esperimentatzailea [+giz]_erg; jarduera_ine			Entitate baten jarduera
jokatu-DU-5	√	√	-	-	+	+	+	-	-	esperimentatzailea [+giz]_erg; modua_abs			Entitate baten jarrera
jokatu-DU-6	-	√	-	-	+	√	+	-	-	esperimentatzailea [+giz]_erg; modua_ins			Entitate baten jarrera
jolastu-DA-1	√	-	-	-	√	+	√	-	-	esperimentatzailea [+biz]_abs; esperimentatzailea [+biz]_soz	jarduera_ine;		Entitate batzuen jarduera
jolastu-DU-2	-	√	-	-	√	+	√	-	-	esperimentatzailea [+biz]_erg; esperimentatzailea [+biz]_soz	jarduera_ine;		Entitate batzuen jarduera
kezkatu-DA-1	√	-	-	-	+	+	+	-	-	gai ukitua [+giz]_abs; kausa_ins			Entitate baten egoera (psikologikoa)ren aldaketa
kezkatu-DU-2	√	√	-	-	+	+	+	-	-	gai ukitua [+giz]_abs; kausa_erg			Entitate baten egoera (psikologikoa)ren aldaketa

ADITZA bere <i>bs</i> sekin	abs	erg	dat	ela	ine	ins	soz	abl	ala	KASU SEMANTIKOAK MURRIZTAPENAK	AZPIKATEGORIZATUEN ETA	OSAGAI HAUTAPEN-	BSSAREN BALIO SEMANTIKOA
kokatu-DA-1	√	-	+	-	√	+	+	-	-	gai ukitua_abs; helburuko kokapena_ine			Entitate baten kokapen-aldaketa
kokatu-DU-2	√	√	+	-	√	+	+	-	-	gai ukitua_abs; kausa_erg; helburuko kokapena_ine			Entitate baten kokapen-aldaketa
konparatu-DU-0	√	√	-	-	+	+	√	-	-	gaia_abs; esperimentatzailea [+giz]_erg; gaia_soz			Entitate baten jarduera beste entitate batzuekiko
konturatu-DA-1	√	-	-	-	+	√	+	-	-	esperimentatzailea [+giz]_abs; gaia [-konkr]_ins			Entitate baten jarduera; erreakziozkoa
konturatu-DA-2	√	-	-	√	+	+	+	-	-	esperimentatzailea [+giz]_abs; gaia_konp-ela			Entitate baten jarduera; erreakziozkoa
landatu-DU-1	√	√	+	-	√	+	+	-	-	gai ukitua_abs; kausa [+giz]_erg; helburuko kokapena_ine			Entitate baten kokapen-aldaketa
landatu-DU-2	√	√	+	-	+	√	+	-	-	helburuko kokapena_abs; kausa [+giz]_erg; gai ukitua_ins			Entitate baten kokapen-aldaketa.
landu-DU-0	√	√	+	-	+	+	+	-	-	gaia [-biz]_abs; esperimentatzailea [+giz]_erg			Entitate baten jarduera
laztandu-DU-0	√	√	+	-	+	+	+	+	-	gaia_abs; esperimentatzailea [+giz]_erg			Entitate baten jarduera
loratu-DA-0	√	-	+	-	+	+	+	-	-	gai sortua [-konkr]_abs			Entitate baten sortzea
lortu-DU-0	√	√	+	-	+	+	+	+	-	gaia_abs; esperimentatzailea_erg			Entitate baten jarduera
mintzatu-DA-1	√	-	-	-	+	√	√	+	-	abiapuntua [+giz]_abs; gaia_ins; helburua [+giz]_soz			Entitate baten elkarreldatzea; adierazpenezkoa
mintzatu-ZAIO-2	√	-	√	-	+	√	+	+	-	abiapuntua [+giz]_abs; gaia_ins; helburua [+giz]_dat			Entitate baten elkarreldatzea; adierazpenezkoa
moztu-DU-0	√	√	+	-	+	+	+	+	-	gaia [-biz]_abs; esperimentatzailea_erg			Entitate baten jarduera
mugitu-DA-1	√	-	+	-	+	+	+	√	√	gai ukitua_abs; abiapuntuko kokapena / bidea_abl; helburuko kokapena_ala			Entitate baten kokapen-aldaketa

ADITZA bere <i>bs</i> sekin	abs	erg	dat	ela	ine	ins	soz	abl	ala	KASU SEMANTIKOAK MURRIZTAPENAK	AZPIKATEGORIZATUEN ETA	OSAGAI HAUTAPEN-	BSSAREN BALIO SEMANTIKOA
mugitu-DU-2	√	√	+	-	+	+	+	√	√	gai ukitua_abs; kausa_erg; abiapuntuko bidea_abl; helburuko kokapena_ala	kokapena /		Entitate baten kokapen-aldaketa
nahastu-DA-1	√	-	+	-	+	+	√	-	-	gai ukitua_abs; gai ukitua_soz			Entitate batzuen egoera-aldaketa
nahastu-DA-2	√	-	+	-	+	+	+	-	-	gai ukitua_abs			Entitate baten egoera-aldaketa (bere baitan)
nahastu-DU-3	√	√	+	-	+	+	√	-	-	gai ukitua_abs; kausa_erg; gai ukitua_soz			Entitate batzuen egoera-aldaketa
nahastu-DU-4	√	√	+	-	+	+	+	-	-	gai ukitua_abs; kausa_erg			Entitate baten egoera-aldaketa (bere baitan)
onartu-DU-1	√	√	+	-	+	+	+	+	-	gaia_abs; esperimentatzailea [+giz]_erg			Entitate baten jarduera
onartu-DU-2	-	√	+	√	+	+	+	+	-	esperimentatzailea [+giz]_erg; gaia_konp-ela			Entitate baten jarduera
oroitu-DA-1	√	-	-	-	+	√	+	-	-	esperimentatzailea [+giz]_abs; gaia_ins			Entitate baten jarduera
oroitu-DA-2	√	-	-	√	+	+	+	-	-	esperimentatzailea [+giz]_abs; gaia_konp-ela			Entitate baten jarduera
oroitu-DU-3	√	√	+	-	+	+	+	-	-	gaia_abs; esperimentatzailea [+giz]_erg			Entitate baten jarduera
oroitu-DU-4	-	√	+	√	+	+	+	-	-	esperimentatzailea [+giz]_erg; gaia_konp-ela			Entitate baten jarduera
Otu-ZAIO-0	-	-	√	√	+	+	+	-	-	esperimentatzailea [+giz]_dat; gaia_konp-ela			Entitate baten jarduera
pasatu-DA-1	√	-	+	-	+	+	+	√	√	gai ukitua_abs; abiapuntua/bidea_abl; helburua_ala			Entitate baten kokapen-aldaketa nahiz egoera-
pasatu-DA-2	√	-	+	-	+	+	+	-	-	gaia [-konkr]_abs			Entitate baten jarduera, denbora epe baten barruan
pasatu-DU-3	√	√	+	-	+	+	+	√	√	kausa_erg; gai ukitua_abs; abiapuntua/bidea_abl; helburua_ala gai ukitua_erg; bidea_abs			Entitate baten kokapen-aldaketa nahiz egoera-

ADITZA bere <i>bs</i> sekin	abs	erg	dat	ela	ine	ins	soz	abl	ala	KASU AZPIKATEGORIZATUEN SEMANTIKOAK ETA MURRIZTAPENAK	OSAGAI HAUTAPEN-	BSSAREN BALIO SEMANTIKOA
pasatu-DU-4	√	√	-	-	+	+	+	-	-	esperimentatzailea [+giz]_erg; gaia [-konkr]_abs		Entitate baten kokapen-aldaketa
pasatu-DU-5	√	√	-	-	+	+	+	-	-	abiapuntua [+giz]_erg; gaia [-biz]_abs; helburua [+giz]_dat		Entitate baten jarduera, denbora epe baten barruan
pasatu-DIO-6	√	√	√	-	+	+	+	-	-			Entitate baten elkarraldatzea
sartu-DA-1	√	-	+	-	√	+	+	-	-	gai ukitua_abs; helburuko kokapena_ine		Entitate baten kokapen-aldaketa
sartu-DA-2	√	-	+	-	+	+	+	√	√	gai ukitua_abs; abiapuntuko kokapena / bidea; helburuko kokapena_ala		Entitate baten kokapen-aldaketa
sartu-DU-3	√	√	+	-	√	+	+	-	-	gai ukitua_abs; kausa_erg; helburuko kokapena_ine		Entitate baten kokapen-aldaketa
sartu-DU-4	√	√	+	-	+	+	+	√	√	gai ukitua_abs; kausa_erg; abiapuntuko kokapena / bidea; helburuko kokapena_ala		Entitate baten kokapen-aldaketa
topatu-DA-1	√	-	-	-	+	+	√	+	-	esperimentatzailea [+giz]_abs; gaia_soz		Entitate baten jarduera
topatu-DU-2	√	√	+	-	+	+	+	+	-	esperimentatzailea [+giz]_erg; gaia_abs;		Entitate baten jarduera
ukitu-DU-0	√	√	+	-	+	+	+	+	-	gaia_abs; esperimentatzailea_erg		Entitate baten jarduera
ulertu-DU-1	√	√	+	-	+	+	+	+	-	gaia_abs; esperimentatzailea [+giz]_erg		Entitate baten jarduera
ulertu-DU-2	-	√	+	-	+	+	+	+	-	esperimentatzailea [+giz]_erg		Entitate baten jarduera
ulertu-DU-3	-	√	+	√	+	+	+	+	-	esperimentatzailea [+giz]_erg; gaia_konp-ela		Entitate baten jarduera
zeharkatu-DU-0	√	√	+	-	+	+	+	+	-	bidea_abs; gai ukitua_erg		Entitate baten kokapen-aldaketa
zintzilikatu-DA-1	√	-	+	-	+	+	+	√	-	gai ukitua_abs; helburuko kokapena_abl		Entitate baten kokapen-aldaketa
zintzilikatu-DU-2	√	√	+	-	+	+	+	√	-	gai ukitua_abs; kausa_erg; helburuko kokapena_abl		Entitate baten kokapen-aldaketa

C ERANSKINA: Landutako aditzentzat aukeratutako “ordain berriak”.

ATERA

- take_out_01
- come_out_01
- publish_01
- take_01
- turn_out_01
- make_01
- get_out_01
- get_01
- go_out_01
- work_out_01
- raise_01

JOKATU

- bet_01

AHAZTU

- forget_01

KONTURATU

- realize_01

EMAN

- spend_02

ALDATU

- replace_01

ESKATU

- ask_02
- demand_01
- require_01

ASMATU

- succeed_01

ERRE

- burn_01
- smoke_01

GERTATU

- be_01

ESKAINI

- dedicate_01

GOSALDU

- breakfast_01

ERAMAN

- spend_02
- win_01