



Universidad del País Vasco Euskal Herriko Unibertsitatea

Euskarazko espazio-egituren etiketatzea: lehen urratsak

Egilea: Jasone Urruzola Balerdi

Tutoreak: Izaskun Aldezabal eta Ainara Estarrona

hap/lap

Hizkuntzaren Azterketa eta Prozesamendua / Language Analysis and Processing Masterreko titulua
lortzeko bukaerako proiektua

2016ko ekaina

Sailak: Lengoia eta Sistema Informatikoak, Konputagailuen Arkitektura eta Teknologia, Konputazio Zientziak eta Adimen Artifiziala, Euskal hizkuntza eta Komunikazioa, Komunikazioen Ingeniaritza.

Laburpena

Proiektu honen helburua euskarazko espazio-egiturak etiketatzeko lehen eredia sortzea da, etorkizunean, IXA taldea ikertzen ari den euskarazko etiketatze semantikoa aberastu asmoz. Horretarako, ingeleserako sortutako ISO-Space eredia izan dugu oinarri eta proposatutako eskema hori euskarara egokitzea izan da egitasmoaren funtsa.

Hala, ISO-Space ereduak proposatzen dituen *oinarrizko espazio-elementuen* izaera eta euren arteko *loturen* nolakotasuna aztertu eta zehaztu ditugu. Eredua bera ondo ulertzeko, euskarazko testu-zati bat etiketatu dugu, MAE tresnaren bidez.

Hizkuntza bakoitzaren gramatika-ezaugarriek kasu batzuetan egokitzapena arreta handiagoz egitea badakarte ere, zailtasunak zailtasun, eredia euskaraz erabiltzeko modukoa dela ikusi dugu.

Dena den, espazio-egituren esparruan sakondu ahala, arloa bera, oro har, hastapenetan dagoela nabaritu ahal izan dugu. Izan ere, ereduko hainbat kontzepturen esanahia azalpenetan topatzea zail suertatu zaigu eta, halaber, alderdi teorikoa eta praktikoa ez direla modu edo aldi berean garatzen ari ikusi dugu.

► Hitz gakoak: euskarazko etiketatze semantikoa, ISO-Space, oinarrizko espazio-elementuak, loturak, MAE

Abstract

The aim of this project is to create the first model for tagging spatial structures in Basque, in order to expand the semantic tagging that is being developed for Basque in the IXA group. For this purpose, we have followed the ISO-Space model created for English. In this project we have focused on adapting it to the Basque language.

We have analyzed the nature of the *basic spatial elements* and *spatial relationships* proposed in ISO-Space. To better understand the model, we have tagged a sample of a Basque text, using the MAE annotation tool.

Although the grammar of each language asks for more attention in some cases, even with some difficulties, we could conclude that the model is also useful for Basque.

However, as we went further on the scope of spatial structures, we noticed that the area *per se* is still in an early stage. In fact, in some cases, it has been difficult to find explanations about some concepts; moreover, it seems that the theoretical aspect and the practical one are not developing at the same time.

► Key words: semantic tagging for Basque, ISO-Space, basic spatial elements, spatial relationships, MAE

Aurkibidea

Aurkibidea	3
1 Sarrera	4
2 ISO-Space eredu euskarara ekarrita.....	5
2.1 Oinarrizko elementuak	6
2.1.1 Kokapen-etiketak	6
Place	6
Path	6
2.1.2 Kokapenezkoak ez direnak	7
Spatial Entity	7
Non-consuming tag	8
Non-motion event	9
Motion	15
Motion_signal	16
Spatial_signal	17
Measure	17
2.2 Lotura-etiketak.....	18
QSLINK	18
OLINK	19
MOVELINK	20
MLINK	21
METALINK	21
3 Ereduaren zailtasunak	22
4 Etiketatzeko tresna: MAE	24
4.1 Tresnaren funtzionamendua	25
4.2 Zailtasunak	28
4.3 Etiketatzeko prozesua: metodologia, kopurua eta denbora	29
5 Ondorioak	30
Bibliografia	31
Eranskina: etiketatutako lagina	32

1 Sarrera

Proiektu honetan, euskarazko espazio-egiturak etiketatzeko lehen urratsak eman ditugu, IXA taldea¹ ikertzen ari den euskarazko etiketatze semantikoa aberastea helburu.

Euskararako gida edo eskema sortzeko, alor honek ingelesean izan duen garapena izan dugu aztergai eta eredu. Lan horretan, SpatialML, Space Role Labeling eta ISO-Space ereduak egin dugu topo; guztiak ere, SemEval lehiaketaren baitan, SpaceEval atazari erantzuteko sortuak. Jarraian, eredu bakoitzaren zertzelada batzuk emango ditugu.

SpatialML

Aipaturiko beste ereduaren oinarrian dagoena da, lehen bertsioa, aitzindaria, nolabait (Mani et al., 2010). Oinarritzko etiketak darabiltza, nagusiki kokapen edo leku geografikoei dagozkienak, leku-izen edo izen-adierazpideak (ahal den heinean, koordenada geografiko jakinak dituztenak) eta eskualde gisa hartzen ditugun lekuen arteko harreman topologikoak. Eredu honetan, leku-izenak hiztegi geografikoei (*gazetteer*) esker desanbiguatzen dira.

Horrez gain, harreman topologiko motak zehaztu zituen lehenengo ereduak dugu SpatialML (RCC8 eskema). Aztergai dugun ISO-Space ereduak ere harreman topologiko mota horiek hartzen dituzenez aintzat, dagokion lotura-etiketaren baitan azalduko dugu sakonago (1.2 Lotura-etiketak, *QSLINK*).

Space Role Labeling

SpaceEval lehiaketa-esparrua Spatial Role Labeling (SpRL) atazaren gainean eraiki zen. 2012.urteko SemEval lehiaketan aurkeztu zen lehen aldiz (Kordjamshidi et al., 2012) eta, gerora, 2013.urtekoan ere erabili zen (Kolomiyets et al., 2013). Ataza horietarako oinarritzko etiketatze-eskema 2010ean aurkeztu zen eta 2011n egin ziren lehen praktika enpirikoak, ingeleserako.

2012ko SemEval-eko SpRL ereduak oinarritzko rolak eta euren arteko espazio-harremanak atzematea zuen xede, batik bat, harreman estatiko edo egonkorak. Hala, *trajectory* (espazioarekiko harremana duen entitatea), *landmark* (zein espaziorikiko dagoen erlazioa) eta *spatial indicator* (harremana eragiten duen elementua) izaerako elementuak eta euren arteko espazio-harremanak identifikatzen zituen; hasiera batean, maila orokorrean, norabidezko harremana (*directional*), topologikoa (*topological*) eta distala (*distal*) bereiziz (Pustejovsky et al., 2015). 2013rako, berriz, eredu hori mugimendu-adierazle eta bide-izaera duten elementuetara ere hedatu zen. Mugimenduzko harremanak edo harreman dinamikoak ere atzematea helburu, *motion indicator* (mugimendu-adierazlea) eta *path* (bidea) elementuak gehitu zituzten. Denborarekin, Space Eval-ean, ataza hedatuz joan zen, analisi maila aberastuz (espazio-egitura dinamiko zein estatikoetan) eta corpusen kopurua zein nolakotasuna (domeinua) zabalduz.

Etiketatzeko-prozesuak alderatuz, bada beste ezberdintasun bat SpRL eta ISO-Space ereduaren artean. Lehengoak *espazio-harremanaren* kontzeptu orokorra definitzen du lehenik, eta harremanaren adiera edo semantikari dagokion informazioa atributuen bidez zehazten du gero.

¹ <http://ixa.si.ehu.es/Ixa>

Bigarren ereduari, berriz, adiera mota bakoitzak harreman mota jakin bat “aktibatzen” du. Bestela esanda, lehena harreman edo loturetatik abiatzen da eta gerora zehazten du loturaren esanahia zein harremaneko partaideen zer-nolakoa. Bigarrenaren prozesua, berriz, kontrakoa da; elementu bakoitzaren semantikatik abiatzen da, horren arabera, harreman bat edo bestea identifikatzeko.

ISO-Space

Aipaturiko SpatialML eta Spatial Role Labeling ereduak, helburuari dagokionez, ISO-Space ereduaren antzekoak badira ere, SpaceEval-ek azken hau – espazio-informazioa atzemateko estandar berria dena – hartzen du aintzakotzat gaur egun (Pustejovsky et al., 2011a; Moszkowicz and Pustejovsky, 2010; ISO/TC 37/SC 4/WG 2, 2014). Espazio-informazioa zein espazio-denborazkoa (*Spatial and Spatiotemporal information*) atzemateko diseinatu zen eta 2015eko SpaceEval-az geroztik gailendu den eredu da, beraz, dokumentu honetan, nagusiki, hizpide izango duguna, horren harira, SpaceEval lehiaketara aurkeztutako gidalerroekin batera.

Lehenik eta behin, 2. atalean, ereduak aurkeztearekin batera, euskararako egokitzapenak azalduko ditugu, hizkuntza batetik besterako aldeak – kasu batzuetan, batik bat, euskara, ingelesa ez bezala, hizkuntza eranskaria izateari zor zaizkionak – azpimarratuz. Ereduak aurkezten dituen etiketak adibideez hornituko ditugu (asko ingelesetik itzuli eta moldatuak). 3. atalean, berriz, ereduak egokitzeak eta erabiltzeak berak izan dituen edo izan ditzakeen zailtasunak azalduko ditugu, kontraesanak edo zalantzak sortzen dituzten adibideak bilduz. 4. atalean, etiketatze-lanetarako erabili dugun MAE tresna aurkeztuko dugu; tresnaren funtzionamendua bera, gure etiketatze-prozesuaren nondik norakoak eta bertan izan ditugun zailtasunak azalduz. 5. atalean, egokitzapen-prozesutik eta sortutako eskema lehen aldiz erabiltzetik (hau da, corpusaren lagin bat ereduari jarraiki etiketatzetik) ateratako hainbat ondorio azaleratuko ditugu. Azkenik, ereduari jarraiki etiketatutako lagina aurkeztuko dugu, eranskin gisa.

2 ISO-Space ereduak euskarara ekarrita

Euskarazko espazio-egiturak etiketatzeko eskema sortzeko, 2015eko SpaceEval-ean aurkeztutako ingeleserako gidalerroak² izan ditugu eredu. Atal honetan, ereduak eskaintzen dituen etiketak, euren atributuak eta balioak azalduko ditugu, bai ingelesetik euskarara bere horretan ekar daitezkeenak, bai egokitzapenak behar izan dituztenak (behar denetan, azalpenak eskainiz).

Gure lagina etiketatzeko, ingeleserako egin bezala, MAE tresna erabiliko dugu. Hori dela eta, proiektu honetan aurkeztuko ditugun etiketatzen, ingelesezko izenei eutsiko diegu eta dagozkien euskarazko itzulpen edo azalpenak gehitu.

Etiketak bi multzotan bereiz daitezke: *oinarrizko elementuak* izendatzen dituztenak batetik, eta euren arteko espaziotan *lotura* edo *harremanak* adierazten dituztenak, bestetik. Lehenengoen artean, halaber, kokapen-etiketak eta kokapenezkoak ez direnak bereizten dira. Jarraian, hurrenez hurren, multzo bakoitzeko etiketak aurkeztu eta azalduko ditugu kasu gehienetan, etiketari buruzko informazioa edo zehaztapenak – bakoitzari dagozkion balioak, atributuak, hedadura (hots, unitate aztergaiaren ohiko kategoria linguistikoa) eta adibide adierazgarri bat(zuk)– taula batean bilduz.

² <http://jamespusto.com/blog/category/news/>

2.1. Oinarrizko elementuak

2.1.1 Kokapen-etiketak

Place

Leku-izaera berezkoa duten elementuak dira; eskualde edo entitate geografikoak (mendiak, lakuak...) zein administratiboak (herriak, hiriak...). Izen-sintagmak (IS) izan ohi dira eta kasu horietan, ISren burua da etiketatzen dena (izen berezia nahiz arrunta), 1. taulan bildu ditugun atributu eta balioen bidez:

Place

Atributua ³	Balioa	Hedadura
id (identifikatzailea)	pl1, pl2, pl3...	ISren burua (izena)
form (forma)	NOM, NAM (Izen arrunta, Izen berezia)	
dimensionality (dimentsioa)	POINT, LINE, AREA, VOLUME (puntua, lerroa, eremua, bolumena)	
elevation (goragunea)	Measure ID (neurriaren IDa)	
Mod (modifikatzaileak)		
dcl (Document Creation Location)	TRUE/FALSE (egiazkoa/faltsua)	
countable (zenbakarria)	TRUE/FALSE (egiazkoa/faltsua)	
Adib.: “Matia <u>k</u> alean gaude” BAINA “ <u>M</u> atian gaude” NOM NAM		

1. taula: *Place* etiketaren atributu eta balioak

Path

“Ibilbide” zentzua edo espazioen arteko muga funtzioa duten kokapenak dira. Izen arruntak (bidea, ibaia...) zein bereziak (Matia [kalea], A8 [autobidea]...) izan daitezke. Ez da beti erraza izaten, ordea, elementu batek *leku* eta *bide* izaera ote duen erabakitzea, esaldiaren baitan hartzen duen funtzioak zalantzak sor baititzake. Esaterako, “ibaia jarraitu” eta “ibaia zeharkatu” adibideak alderatuz, zilegi da inork bigarren adibidean “ibaia” *place* modura etiketatzea (berezko “ibilbide” esanahia galdu egin baitu adibide horretan).

Bereizketarako, ereduak proposatzen duen heuristikoetako bat *egon/hartu* (*be-at/take*) bereizketa da. Izan ere, espazio batean *egon* egin badaiteke, *place* zein *path* etiketa esleitu diezaiokegu, baina espazio hori *hartu* edo *jarraitu* egin badaiteke, hautatu beharrekoa *path* izango da, ezinbestean.

³ Azalpen hauetan ez dugu atributu eta balio guztietan sakonduko, kasu batzuetan, itzulpena bera nahiko argia delakoan (azalpena beharrezko deritzogunean soilik sakonduko dugu).

⁴ Kako artekoak elidituta dauden kasuetan, izen berezia ISren burua den kasuetan, alegia.

Etiketatzailen arteko adostasun edo koherentziaren mesederako, ereduak “ibilbide” izaera berezko duten elementuak – errepideak (*roads*), ibaiak (*rivers*), eskilarak (*stairways*), mendikateak (*mountain ranges*) – *path* gisa etiketatzea hobesten du (testuinguru jakinetan *leku* zentzua izan badezakete ere).

2. taulan ditugu etiketa honi dagozkion atributu eta balioak:

Path		
Atributuak	Balioak	Hedadura
id (identifikatzailea)	p1, p2, p3...	- ISren burua (izena) - Bide “estatikoak” (ez egoerazkoak)
type (mota)	BODYOFWATER, MTS ⁵ , ROAD (ur-masa, ?, errepidea)	
beginID (hasierako ID)	edozein leku, bidearekiko hasierako muga dena	
endID (bukaerako ID)	edozein leku, bidearekiko bukaerako muga dena	
midIDs (tarteko IDak)	edozein leku, bidearekiko tarteko zatia dena	
dimensionality (dimentsioa)	POINT, LINE, AREA, VOLUME (puntua, lerroa, eremua, bolumena)	
form (forma)	NOM, NAM (Izen arrunta, izen berezia)	
elevation (goragunea)	neurriaren IDa	
mod (modifikatzaileak)		
countable (zenbakarria)	TRUE/FALSE (egiazkoa / faltsua)	
Adib.: “Urumea <u>ibajan</u> zehar dabil piraguan” BAINA “ <u>Urumean</u> zehar dabil piraguan” NOM NAM		

2. taula: *path* etiketaren atributu eta balioak

2.1.2 Kokapenezkoak ez direnak

Spatial Entity

Oro har, honelako entitatetzat jotzen dira, berezko espazio-izaerarik ez duten elementuak eta, beraz, “leku” eta “bide” kategoriekin ezin lot daitezkeenak. Esan nahi baita, espazio-entitate batek espazio errealean kokatuta egoteaz gain, lotura- edo harreman-etiketa bateko partaide izan behar du.

Praktikan, mugitzen diren edo mugikortasuna duten objektuak edo egoera-aldaketaren bat jasaten dutenak dira, batik bat, etiketa hau jasotzen dutenak, 3. taulako atributu eta balioekin:

5 Itzulpenik gabe – hots, itzulpenetan '?' ikurrarekin – utzi ditugun balioen gaineko azalpen argirik ez dago ereduari.

Spatial Entity

Atributuak	Balioak	Hedadura
id (identifikatzailea)	se1, se2, se3...	ISren burua (izena)
type (mota)	FAC, VEHICLE, PERSON, DYNAMIC_EVENT... (?, ibilgailua, pertsona, gertaera dinamikoa...)	
dimensionality (dimentsioa)	POINT, LINE, AREA, VOLUME (puntua, lerroa, eremua, bolumena)	
form (forma)	NOM, NAM (Izen arrunta, izen berezia)	
mod (modifikatzaileak)		
countable (zenbakarria)	TRUE/FALSE (egiazkoa / faltsua)	
Adib.: - “ <u>Jon</u> ek Boston bisitatu zuen” - “ <u>Dorrea</u> erori egin da tximistak jo duenean”		

3. taula: *spatial_entity* etiketaren atributu eta balioak**Non-consuming tag**

Zeharka erreferentzia egiten zaien eta espazio-izaera duten elementuak dira, esplizituki testuan ageri ez direnak.

Normalean, elementu hau ez da ezinbestekoa objektu eta harreman espazial esanguratsuak atzemateko eta, horregatik, ereduak etiketa hau tentuz erabiltzea gomendatzen du. Honako egoeretan erabiltzea jotzen du beharrezkotzat:

1. Neurri (*measure*) bat agertzen denean

Goragune batek (*elevation*) zeharka kokapen esanguratsu bat erreferentziatzen duenean. *Ibaitik 3 milatara kanpatu genuen* esaldian, esaterako, *kanpatu* den lekua (*place* etiketa jasoko lukeen elementu eliptikoa) erreferentziatzen da zeharka.

2. Gurutzatu, zeharkatu, -n zehar eta antzeko elementuak daudenean

Zeharkatzen den eremu edo bidearen nondik norakoa testuan esplizituki ageri ez bada. Mugimenduzko gertaeren artean, *cross* motan (*motion_class*) sailkatzen direnekin gertatu ohi da (ik. *Motions*). *Jon herrian zehar ibili zen* esaldian, adibidez, *Jon*-ek egindako bidearen hasierako kokapena edo abiapuntua (*source*) eta bukaerakoa edo helmuga (*goal*) dira elipsian dauden lekuak.

3. Multzo bateko partaideak aipatzen direnean

Testuan elementu multzoak erreferentziatzen direnean, batzuetan beharrezkoa da elementu eliptiko hauek azalratzea, gerora, lotura edo harremanetan parte har dezaten. Esaterako, aurrekari banatuaren kasuan, elipsian dagoen aurrekaria azalratzea beharrezkoa izango da, ondoren, harremana azpikorreferentzia (*subcorreference*) edo korreferentzia banatua (*splitcorreference*) dela adierazteko (ik. 1.2 Lotura-etiketak, *METALINK*).

*Jon eta Miren (Ø) dendan zeuden. Haiek erosketak egitera joan ziren.*⁶

Adibide horretan, ereduaren arabera, beharrezkoa da *Jon* eta *Miren* multzo berean bilduko dituen elementua sortu edo azaleratzea (Ø), gerora, *Haiek* elementuarekiko duen azpi-korreferentzia edo korreferentzia banatuko harremana gauzatzeko.

Euskarazko eredurako egiten dugun proposamena, ordea, aipatu multzoetako hirugarren kasua baino ez etiketatzea da, *elipsi testualtzat* hartuko genituzkeen elementu eliptikoak baino ez azaleratzea, alegia (gure proposamenaren nondik norakoak sakonago aztertuko ditugu 3. atalean).

Non-Motion Event

Mugimendurik gabeko gertaerei ematen zaien izena da. Era berean, *gertaerak* une jakin batekoak edo denbora-tarte bat hartzen dutenak izan daitezke. Etiketa honen helburua, mugimendurik adierazi ez arren, espazio-elementuren batekin harreman zuzena duten gertaerak atzematea da. Hala, horrelako gertaerek *place*, *path*, *spatial_entity* edo beste *nonmotion_event* zein *motion* etiketadun elementuekin izango dute lotura edo harremana. Are gehiago, esaterako, *place* etiketan azaldu ditugun *elevation*, *mod* eta *countable* atributuen bidez zehazten da.

Bestalde, egoera emozional edo esperimentalei etiketa hau ez esleitzea proposatzen du ereduak, baldin eta gertaera horiek “espazio errealean” existitzen ez badira, emozio eta esperientziak “ez-kokapenezkotzat” har baitaitezke, esperimentatzaile batekin (espazio-entitate bat litzatekeena) lotura badute ere. Esaterako, *Jon asko poztu da* esaldian, gertaera *Jon*-en egoera-aldaketan datza eta, beste testuinguru batean *Jon* espazio-entitatetzat jo bagenezake ere, kasu honetan gertaera ez da espazio errealean gauzatu, hots, ez du espazio jakin batekiko harremanik. Ikus dezagun argiago antzekoak diruditen bi adibide konparatuz:

- a) *Festak 3 ordu iraun zituen*
- b) *Teilatuan egin zen festak 3 ordu iraun zituen.*

a) eta b) esaldietan dugun *fiesta* hitzak mugimendurik gabeko gertaera bat adierazten du. Alabaina, b) adibidean gertaera hori kokapen edo espazio jakin batekin dago lotuta (*teilatua*), a) esaldikoa ez bezala. Hori horrela, ereduak b) kasuaren parekoak etiketatzea proposatzen du.

Hedapenari dagokionez, etiketa honek token soilak (hitz bakarrekoak) markatzen ditu. Dena den, mota askotako hitzak izan daitezke gertaera adierazten dutenak eta, kasu honetan, hizkuntza bakoitzak ditu bereak.

Aditzak

Ereduak egoera bat adierazten duten aditzak ere gertaera gisa etiketatzea proposatzen du, aditz edo aditz-sintagma guztiek, betiere, gertaera bat adierazten dutelako ideiatik abiatuz.

⁶ Egia da, halaber, honelako adibideetan ingelesa eta euskara ez datozela guztiz bat. Izan ere, euskarak bigarren esaldiko *Haiek* elementua ere elipsian ematea ahalbidetzen digu (are gehiago, ohikoena litzateke), ingelesak ez bezala. Horrelakoetan, bi elementu eliptikoren (*non-consuming tag*) artean sortu beharko dugu harremana.

Aditz-sintagma sinplea zein konposatua izan, etiketak sintagmaren burua baino ez du biltzen (euskararen kasuan, aditz nagusia); berdin aditz-lokuzio edo esamoldeen kasuan ere.

Saudi Arabiak soilik du petrolio-erreserba gehiago esaldian, esaterako, mugimendua baino, egoera bat da adierazten dena, baina badu erlaziorik kokapen jakin batekin (*Saudi Arabian* dago petrolio-erreserba gehiago); beraz, etiketa hau *du* aditza markatzeko erabiliko dugu.

Izenak

Izenen kasuan, TimeML-ko informazioa (Pustejovsky et al., 2006) zuzenean erabiltzen da ISO-Spacen (bi azterketak lotuta daude, beraz). Bertan ez dira izen guztiak gertaeratzat hartzen. Honako testuinguruetan agertzeko aukera izan behar dute horretarako:

- 1- IZENA + *X segundu/minutu/egun/urte... irau*n zuen.
- 2- IZENA + *azkarra/berehalakoa/bat-batekoa... izan* zen.
- 3- IZENA + denborazko adierazpena + *gertatu* zen.
- 4- IZENA + denborazko adierazpena + *hasi/jarraitu/amaitu* zen/zuen.

Behin testuinguruaren baldintza beteta, honelako izenen artean lau mota bereizten dira eta mota bakoitzerako irizpide batzuk proposatzen dira zer etiketatu eta zer ez zehazteko.

1. Izenaren modifikatzaile gisa jokatzeko duten gertaera-izenak (*prepositional modifiers*). Hauek ez dira gertaera gisa etiketatzen:

Bileran, euren hauteskunde-porrotaren gaineko ikerketa hasi zuten.

2. Egoera ez-egonkor (*sortal states*) motakoak. Honelakoak izan ohi dira:

- Jarduera edo ekintza jakinen eragileei erreferentzia egiten dieten izenak (*agentive nominals*), kasu gehienetan, adizetik eratorriak direnak (*irabazle, galtzaile, garbitzaile...*).
- Lanbideak, rolak edo karguak adierazten dituzten izenak (*medikua, iturgina, zuzendaria...*)
- Erreferentziatzaile zurrinak (*rigid designators*), hots, munduko edozein toki izanda ere, entitate bera erreferentziatzen duten termino edo egiturak (*Ameriketako hirugarren presidentea*).

Elementu hauek osagarri predikatiboaren buru gisa jokatzeko dutenean baino ez dira hartuko gertaeratzat. Osagarri horrek predikazio bat adierazten du eta, besteak beste, jarraian zerrendatzen diren aditz motaren baten osagarri izango da:

- Predikazio kopulatiboak (*copulative predicates*), eskuarki, *izan* eta *iruditu* aditzekin .

Gaur egun bera da [herriko elkarte gastronomikoko presidentea].

- Predikazio inkoatiboak (*inchoative predicates*), *bihurtu* aditzarekin. Egoera-aldaketa bat adierazten dute.

1821ean Simon Bolivar [Boliviako errepublikako lehen presidente] bihurtu zen.

- Aspektu-predikazioak (*aspectual predicates*), *hasi, jarraitu, bukatu* eta antzeko aditzekin.

[Espazio-egiturei buruzko ikerketako zuzendari gisa] jarraitu zuen.

- Egoera-aldaketa adierazten duten predikazioak (*change of states predicates*), *hots, iristear dagoen edo bukatu den egoera baten berri ematen dutenak. Erretiratu, izendatu, hautatu, uko egin* eta antzeko aditzekin.

? 1998an erretiratu zen [irakasle gisa].

Hori baino ohikoagoa da euskaraz 1998an erretiratu zen irakaskuntzatik edo antzeko egituraren bat erabiltzea, beraz, kasu honetan, euskaraz ez dugula egitura predikatiborik esan genezake.

- Ebaluazio- eta deskribapen-predikazioak (*predicates of evaluation and description*). *Deskribatu, ebaluatu, -tzat hartu/jo* eta antzeko aditz edo egiturekin.

[*Linguae Vasconum Primitiae* lehen euskal liburutzat hartzen da].

3. Gertaerak adierazten dituzten beste izen guztiak. Hauek ere beti etiketatzen dira gertaera gisa.

Gizonak maskaradun mutil batek egindako erasoa izan zela azaldu zuen.

4. Hitz anitzeko gertaera-izenak. Kasu hauetan, berriz, determinatzaileak, espezifikatzaileak, osagarriak, modifikatzaileak... eta halakoak kanpo utzi eta sintagmaren burua baino ez da etiketatzen.

[Aita Santuaren bisitak] milaka lagun bildu ditu.

Adjektiboak

Adjektiboek entitate bati ezaugarri edo atributu bat erantsi ohi diote eta, beraz, izaera egonkor edo estatikoko gertaerak adierazten dituzte. Kokapen atributibo edo predikatiboan ager daitezke eta ereduak gertaera gisa azken mota horretakoak baino ez etiketatzea proposatzen du.

1. Adjektibo atributiboak: izenaren modifikatzaile gisa jokatzen dutenak dira (*lorategi polita, min jasanezina...*)

2. Adjektibo predikatiboak: adierazten duten ezaugarria iraunkorra ez bada markatuko dira. Ezaugarri asko izan daitezke mota honetakoak. Esaterako, pertsonen euren nazionalitatea alda dezakete eta autoak, jatorrizkoa ez den beste kolore batez margo daitezke. Hori horrela, *Itsasontziko jende gehiena senegaldarra zen* esaldiko *senegaldarra* adjektiboa gertaera gisa markatuko da; eta baita *Akusatuaren autoa gorria zen* esaldiko *gorria* adjektiboa ere.

Bestela esanda, markatzekotan, gutxienez honako baldintzetako bat bete beharko dute:

- Ezaugarria argi eta garbi aldakorra denean (ez-iraunkorra). *Aurpegia gorri jartzen zaio haserretzen denean.*
- Ezaugarria behin-behinekoa bada, garai edo une jakin batekiko (aipatutako *senegaldarra* eta *gorria* adjektiboak dituzten esaldiak, kasu).
- Ezaugarriak iritzia, ezagutza edo sinesmena adierazten duenean edo eztabaidagai denean. *Konpontzea zail ikusten dut.*

Mota honetako adjektiboek aditz mota hauetako baten (besteak beste) osagarri predikatibo gisa jokatzen dute:

- Predikazio kopulatiboak. *Ikasleek nekatuak dirudite hiru astez klaseak izan ondoren.*
- Predikazio inkoatiboak. *Bere bititza jasanezin bihurtu zelako egin zuen ihes.*
- Aspektu-predikazioak. *Familiak itxaropentsu jarraitu zuten senideen zain.*
- Predikazio kausatiboak. *Taldeak zoriontsu egin zituen ikusleak.*
- Egoera-aldaketa adierazten duten predikazioak, oro har.⁷
- Pertzepzio-predikazioak. *Jokalariek gogotsu ziruditen.*
- Ebaluazio- eta deskribapen-predikazioak. *Alfertzat jo dute askotan.*

Aurreko lerroetan azaldutako kasu asko zailak izan daitezke ebaluatzen. Ereduaren aholkua zalantza-kasuetan adjektiboak gertaera gisa ez etiketatzea da.

Preposizio-sintagmak

Euskararen kasuan postposizio-sintagmak ditugu, beraz, preposizioez baino, postposizioez arituko gara puntu honetan. Izenek, adjektiboek edo aditzek berez gertaerarik adierazten ez dutenean, postposizioek hartzen dute betebehar hori. Honelakoak ondorengo aditz motaren baten (besteak beste) osagarri predikatibo funtzioa dutenean soilik etiketatuko ditugu.

- Predikazio kopulatiboak. *Emakume bat ez da misio bat[en arduradun] izan orain arte.*
- Predikazio inkoatiboak. Euskararen kasuan ez dugu halako adibiderik, *bihurtu* aditzarekin ez baitugu inolako postposiziorik erabiltzen (mugagabea edo mugatua, modifikatzailea konplexuagoa den heinean, erabiltzen da).

⁷ Ereduak ez du mota honetako adibiderik eskaintzen.

- Aspektu-predikazioak. *Politikariak hautagaitzatik kanpo jarraitzen du.*
- Predikazio kausatiboak. *Gertaera honek erabakia epailearen esku jartzen duela esan du.*

Aipaturiko kasu horietan, postposizioa soilik etiketatuko dugu. Era berean, postposizio-sintagmaren baitan dagoen aditz, izen edo adjektiboak berez gertaerarik adierazten ez dutenean baino ez ditugu etiketatuko postposizio horiek.

Gertaera adierazten duten beste elementu batzuk

Gertaerak beste elementu batzuen bidez ere adieraz daitezke, eskuarki, leku-aditzondoekin (*hemen* edo *han*, esaterako).

Denbora luzez geratzeko asmoz gaude hemen.

Gertaera adierazten duten egitura konplexuak

Gerta liteke gertaera bat baino gehiago egitura beraren bidez adieraztea. Kasu gehienetan, aditzen bidez izango da, baina baliteke izen-, adjektibo- edo postposizio-sintagmekin ere topo egitea. Jarraian egitura ohikoenak azalduko ditugu.

- Aspektu-egiturak. Aspektu-aditz (*hasi, amaitu, gelditu...*) edo izen batek (*hastea, amaitzea...*) eta osagarri batek osatua. Azken hori, aditz- zein -izen-sintagma bidez adieraz daiteke. Honelako egituretan, ereduak bai aspektu-predikazioa, bai bere osagarria gertaera beregain gisa etiketatzea proposatzen du. Ondorengo adibideetan, aspektu-predikazioa beltzez adierazi dugu eta osagarria, berriz, azpimarratuta:

Estatu Batuek ez diote beste herrialdeetako politiketan nahasteari uzten.
(Aspektu-predikazioa aditz bidez)

Bush eta bere aholkulariak krisiaren ondorio militarri begira zeuden.
(Aspektu-predikazioa izen bidez)

- Egitura inkoatiboak. *Bihurtu* eta antzeko aditzek egoera-aldaketa bat adierazten dute eta euren osagarriek prozesua edo azken egoera. Kasu honetan ere, ereduak bi elementuak gertaera gisa markatzearen aldekoa da.

Printzea igel bihurtu zen.

- Egitura kausatiboak. Hauen artean bi motatakoak aurki ditzakegu:
 - GERTAERA_{e1} *eragin/ekarri*_{e2} GERTAERA_{e3}: kausazko adierazpena (e2), bere subjektu logikoa (e1) eta osagarria (e3); guztiak gertaera beregain gisa markatzen dira.

[Euriak_{e1}] [uholdeak_{e3}] [eragin_{e2}] zituen.

- ENTITATEA *eragin/ekarri*_{e1} GERTAERA_{e2} : bai kausazko adierazpena (e1), bai bere osagarria (e2), gertaera beregain gisa markatzen dira.

Jonek [sua_{e1}] [eragin_{e2}] zuen

- Aditz arinen egitura. Aditz arin batek eta bere osagarri gisa jokatzeko duen gertaera-izen batek osatzen dute egitura hau. Aditz arinak eduki semantiko ahuleko aditzak dira, euskaraz hitz anitzeko unitate lexikaletan (HAUL) agertu ohi direnak (*egin, eman, izan...*). Kasu honetan ere, ereduak bai aditza eta bai osagarria (izena) gertaera beregain hartzea proposatzen du.

Ez dute nire iritzia kontuan hartu.

- Egitura kopulatiboak. *Izan* edo *iruditu* aditza buru duten aditz-sintagmak dira, osagarri gisa izen-, adjektibo- edo postposizio-sintagma dutenak. Honelako kasuetan, aditz-predikazioa zein bere osagarri predikatiboa markatzen dira gertaera gisa. Hurrengo adibideetan, predikazio kopulatiboa beltzen markatu dugu eta osagarriaren burua, berriz, azpimarratuta.

Emakume hori elkarteko buru berria da.

- Izen funtzional bidezko egiturak. *Temperatura, neurria, zabalera, populazioa* edo *biztanleria, intentsitatea* eta antzeko izenak. Honelakoek banako bat hartzen dute argumentutzat eta balio jakin bat itzultzen dute, zenbakizkoa izan daitekeena, edo ez (*altua, baxua, handia, txikia...*). Mota honetako izenak hiru egitura ezberdinetan ager daitezke. Guztietan, izen funtzionala (adibideetan azpimarratuta) gertaera modura markatzen da, bere perpaus nagusiko aditz nagusiarekin (adibideetan beltzez) batera⁸:

- IZENA_{funtz.} X *izan*

*Gaurko temperaturarik altuena 30 gradukoa **izan** da.*

- BANAKOA X-ko IZENA_{funtz.} *eduki/ukan.*

Kataluniak 7 milioi pertsona inguruko biztanleria du

- IZENA_{funtz.} *aldatu/altxatu/askatu/hazi/handitu...* (X-tik) (Y-ra)

*9 hilabetetan zorra % 4 **hazi** da.*

⁸ Ingelesak ez bezala, euskaraz ordena askea duenez, egitura hauek aukera gehiago izan ditzakete.

Motion

Mugimendu-izaera duten gertaerek jasotzen duten etiketa da. Horietako elementu bakoitza mugimendu-gertaerako partaide den beste edozein elementurekin egongo da harremanetan. Hau da, etiketa hau erabiltzen dugun bakoitzean, kontuan izan beharko dugu *MOVELINK* deritzon harreman-etiketa bat sortu beharko dugula (ik. 1.2. Lotura-etiketak).

Mota honetako gertaeretan hiru aukera daude:

- Moduzko mugimendua: *Jon oinez joan zen.* (ingelesezko *walk*)
- Bide-mugimendua: *Jonek etxea utzi zuen.*
- Mugimendu konposatua (aurreko bien batura): *Jonek etxea utzi zuen korrika edo Jon oinez joan zen etxera.*

Aukera hauek, orokorki onartua den Talmy-ren (Talmy, 1985) tipologian dute jatorria, eta hain zuzen lan horretan azaltzen dena oinarri hartuta, puntu honetan ingelesak eta euskarak ez dute bat egiten; izan ere, ingelesa satelite-hizkuntza litzateke (*satellite-framed language*), mugimenduaren norabidea azaltzen duen elementua aditzetik kanpo duen hizkuntza baita; eta euskara, berriz, aditz-hizkuntza (*verb-framed language*), informazio hori aditzaren baitan ematen duen heinean. Ingelesaren kasuan, esaterako, *go* aditzak berez adierazten du mugimendua baina preposizio baten beharra du mugimendu horren norabidea edo bidea zehazteko. Hartara, preposizio bat edo bestea erabili, mugimenduaren norabidea ezberdina izango da: *go out, go in, go up, go down...* euskaraz, berriz, aipatu adibideen baliokideak aditz hutsarekin adierazten dira, hau da, aditzak berak barnebiltzen du bai mugimendu-izaera eta bai mugimenduaren norabidea zehazten dituen informazioa: *atera* (kanpoaldera), *sartu* (barrualdera), *igo* (gorantz), *jaitsi* (beherantz)...

Horrekin lotuta, dirudienez, satelite-hizkuntzek joera handiagoa dute "modua" aditzean lexikalizatzeko, eta horregatik du ingelesak "moduzko mugimendua" deitu dutena. Aditz-hizkuntzek, aldiz, joera handiagoa dute modu hori aditzetik kanpoko elementu batez adierazteko, eta ohikoagoa da "bide-mugimendua" edo "konposatua" agertzea. Horretaz gain, euskararen kasuan, esan behar dugu, mugimenduaren helburua edo helburuko kokapena aditzean txertatuta ager daitekeela (*laster etxeratu ginen bezalakoetan*), eta hori horrela izanik, beste aukera bat gehitu beharko genuke agian: helburu-mugimendua. Laburtuz eta beste modu batera esanda, hizkuntza bakoitzak bere modua du "bidea", "norabidea", "helburua", "modua"... lexikalizatzeko.

Etiketa hau aditzei esleitzen zaie, baldin eta "espazio errealean" leku-aldaketa adierazten dutenak badira. Esaterako, *Jonek bidea jarraitu zuen* esaldian, *jarraitu* aditzak benetako mugimendua adierazten du; *Jonek mapa jarraitu du* esaldian, ordea, aditza berbera izanda ere, ez dago espazio errealean gertatzen den mugimendurik.

Motion etiketak 4. taulako atributu eta balioak jasotzen ditu:

Motion

Atributuak	Balioak	Hedadura
id (identifikatzailea)	m1, m2, m3...	Aditza
motion_type (mugimendu mota)	MANNER, PATH, COMPOUND (modua, bidea, konposatua)	
motion_class (mugimendu klasea)	MOVE, MOVE_EXTERNAL, MOVE_INTERNAL, LEAVE, REACH, DETACH, HIT, FOLLOW, DEVIATE, CROSS ⁹	
Motion_sense (mugimenduaren zentzua)	LITERAL, FICTIVE, INTRINSIC_CHANGE (literala, fikziozkoa, berezkoa)	
mod (modifikatzaileak)		
countable (zenbakarria)	TRUE/FALSE (egiazkoa / faltsua)	
Adib.: “Jon <u>iritsi</u> zen”		

4. taula: *motion* etiketaren atributu eta balioak

Motion_signal

Mugimendu-seinaleak, mugimendu-gertaeraren baitan, biderei edo moduari buruzko informazioa ematen duten partikula, preposizio, aditz edo aditzondoak dira. Euskaraz, jakina, preposizioak baino, postposizioak ditugu eta bereizketa hori etiketa batean baino gehiagotan egitea suertatuko da aurrerantzean. Dena den, hori ez da eragozpen euskararen moduko hizkuntzetan – hots, hizkuntza eranskarietan– ere eredu hau baliatzeko, koreerarako egindako egokitzapen-lanek erakusten duten bezala (Lee et al., 2011).

Kasu honetan, identifikatzaileaz gain, ematen duten informazio mota izango da atributu eta balioen bidez zehaztu beharreko bakarra (5. taula):

Motion Signal

Atributuak	Balioak	Hedadura
id (identifikatzailea)	a1, a2, a3...	- Bide-mugimendua denetan, postposizioa, partikula edo elementu soltea soilik. (1) - Moduzko mugimendua edo mugimendu konposatua denean, postposizio- edo aditzondo-egitura (egitura adberbial) osoa. (2)
motion_signal_type (mota)	MANNER, PATH (modua, bidea)	
Adib.: - “Dend <u>ara</u> sartu zen” (1) - “ <u>Autoz</u> etorri ginen” (2)		

5. taula: *motion_signal* etiketaren atributu eta balioak

⁹ Euskarazko sailkapena aztertzeke dago eta momentuz, ingelesezkoa erabil badaiteke ere (guk hala egin dugu), etorkizunerako zeregin interesgarria litzateke, hizkuntzen artean aldeak egon baitaitezke. Hori dela eta utziko ditugu, momentuz, balio horiek itzuli gabe.

Spatial_signal

Espazio-seinaleak lekuzko bi elementuren arteko harreman jakina adierazten duten preposizioak edo preposizio-egiturak dira (euskararen kasuan, postposizioak). Aurreko puntuan ikusi ditugun mugimendu-seinaleen aldean, hauek ematen duten informazioa bestelakoa da. Kasu honetan, elementuen arteko harremanari esleitzen zaion informazioa topologikoa edo kualitatiboa da, hots, elementuek espazio jakin batean zein harreman mota erakusten duten (espazioa partekatzen duten ala ez, zein modutan partekatzen duten, bata bestearikiko duen kokapena...).

Seinale hauen etiketatzea 6. taulako atributu eta balioen bidez adierazten da:

Spatial Signal

Atributuak	Balioak	Hedadura
id (identifikatzailea)	s1, s2, s3...	Postposizioa
cluster (multzoa)	Postposizioaren zentzua	
semantic_type (mota semantikoa)	DIRECTIONAL (1), TOPOLOGICAL (2), DIR_TOP (3) (norabidezkoa, topologikoa, norabidezkoa + topologikoa)	
Adib.: - “Boston New York <u>iparraldean</u> dago” (1) - “Donostia Gipuzko <u>an</u> dago” (2) - “Edalontzia mahai <u>(ren)</u> <u>gainean</u> dago” (3)		

6. taula: *spatial_signal* etiketaren atributu eta balioak

Mota semantikoaren arabera (hots, gaineratzen duten informazioaren nolakotasunaren arabera), harreman mota jakin bat aktibatuko da (ik. 1.2. Lotura-etiketak):

- Norabidezkoa (DIRECTIONAL): *OLINK* lotura-etiketa.
- Topologikoa (TOPOLOGICAL): *QSLINK* lotura-etiketa.
- Norabidezkoa + topologikoa (DIR_TOP): *OLINK* eta *QSLINK* lotura-etiketak.

Measure

Neurria espazioen arteko distantzia edo dimentsioa adierazten duen elementua da, zenbakizko balio batez eta neurri-unitate (m, km...) batez edo *-tik/-dik gertu* eta *-tik/-dik urrun* eta antzeko distantzia-neurri zehaztu gabez osatua. Honelako egiturek, *MLINK* deritzon lotura-etiketari ematen diote bide (ik. 1.2. Lotura-etiketak). 7. taulako atributu eta balioak erabiltzen dira etiketa honen ezaugarriak zehazteko:

Measure

Atributuak	Balioak	Hedadura
id (identifikatzailea)	me1, me2, me3...	Zenbakia eta neurri-unitatea edo distantzia-neurri zehaztu gabea adierazten duten postposizio-egiturak
Value (balioa)	Zenbakizko osagaia edo <i>-tik/-dik gertu</i> , <i>-tik/-dik urrun</i> ...	
unit (unitatea)	Neurri-unitatea adierazten duen egitura	
Adib.: - “Jonek <u>5 kilometro</u> egin zituen oinez” - “Jon <u>parketik gertu</u> ibili zen”		

7. taula: *measure* etiketaren atributu eta balioak

1.2 Lotura-etiketak

Lotura-etiketak oinarrizko elementu etiketatuen arteko loturak ezartzen dituztenak dira. Harreman motaren arabera, *QSLINK*, *OLINK*, *MOVELINK*, *MLINK* eta *METALINK* etiketak aurkituko ditugu.

QSLINK (Qualitative Spatial Link)

Erlazio honek bi espazio-objektuen arteko harreman topologikoa adierazten du. Hori dela eta, *topological* edo *dir-top* motako espazio-seinaleen beharra izango du, sortuko bada.

Harreman-etiketa honek 8. taulako atributu eta balioak hartzen ditu:

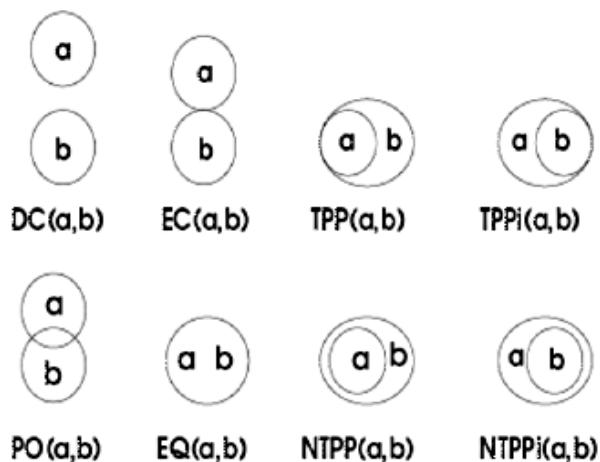
QSLINK

Atributuak	Balioak	Lotura aktibatzen duen elementua
id (identifikatzailea)	qsl1, qsl2, qsl3...	TOP edo DIR_TOP motako espazio-seinalea
reltype	RCC8+ eskemako harremana	
trajector	Erlazioa duena (eskuarki objektu mugikorra)	
landmark	Zerekiko erlazioa duen (eskuarki elementu estatikoa)	
trigger	Lotura eragiten duen espazio-seinalea	

Adib.: “[Liburua_{se1}] [mahai_{se2}] [gainean_{s1}] dago”
trajector=se1; landmark=se2; trigger=s1; reltype= EC

8. taula: *QSLINK* etiketaren atributu eta balioak

Objektu espazialak elkarren artean nola lotzen diren adierazteko, SpaceEval-ek RCC (Region Connection Calculus) eskema darabil. RCC8+ deritzona, eskema horren ildotik sortua da, honako oinarrizko 8 harremanen gainean eraikia:



- DC (*disconnected*): euren artean inolako ukipenik ez duten eremuak direnean.
- EC (*external connexion*): harremanetan edo kontaktuan dagoena eremuen kanpoaldea denean.
- TPP (*tangential proper part*): A eremua B eremuaren parte izan eta ukipen-harremana dagoenean.
- TPPi (*TPP inverse*): B eremua A eremuaren parte izan eta ukipen-harremana dagoenean.
- PO (*partial overlap*): eremuen arteko harreman edo kontaktua (gainjartzea, kasu honetan) partziala denean.
- EQ (*equal*): izendapen ezberdinekin eremu bera erreferentziatzen denean.
- NTPP (*non tangential proper part*): A eremua B eremuaren parte izan eta ukipen-harremanik ez dagoenean.
- NTPPi (*NTPP inverse*): B eremua A eremuaren parte izan eta ukipen-harremanik ez dagoenean.

OLINK (Orientation Link)

Berezko harreman topologikorik ez duten bi kokapenen arteko harremanak atzematen dituen lotura da.

QSLINK bidez atzeman ezin daitezkeen hiru informazio mota zehazten ditu:

- Absolutua (*absolute*): ikuspegi objektiboa (“*bird’s eye*” view).
- Berezkoa (*intrinsic*): espazio-objektuak orientazio edo norabide jakina (berezkoa) duen elementuren bat duenean (*telebistaren aurrean* diogunean, esaterako).
- Erlatiboa (*relative*): Norabidea ikuspuntu jakin baten arabera denean (*ezkerraldean dagoen mahaia* diogunean, kasu).

Horrez gain, espazio-seinaleen puntuan aipatu bezala, norabidezkoak (*DIRECTIONAL*) edo norabidezkoak zein topologikoak (*DIR_TOP*) diren seinaleek aktibatzen dute harreman mota hau.

Harreman-etiketa honek 9. taulako atributu eta balioak hartzen ditu:

OLINK

Atributuak	Balioak	Lotura aktibatzen duen elementua
id (identifikatzailea)	ol1, ol2, ol3...	DIR edo DIR_TOP motako espazio-seinalea
reltype	ABOVE, BEHIND, NEXT TO, NORTH OF ... (gainean, azpian, ondoan, iparraldean...)	
trajector	Erlazioa duena	
landmark	Zerekiko erlazioa duen	
trigger	Lotura eragiten duen espazio-seinalea	
frame_type	ABSOLUTE, INTRINSIC, RELATIVE (absolutua, berezkoa, erlatiboa)	
referencePt	Norabide kardinala, espazio-entitatea, ikuslea	
projective	TRUE/FALSE (egiazkoa: beti dagoen leku horretan/ faltsua: une jakin batean bakarrik dagoen leku horretan)	
<p>Adib.¹⁰: “[Boston_{p11}] [New York_{p12}] [iparraldean_{s1}] dago” trajector=p11; landmark=p12; trigger=s1; reltype=”NORTH”; frame_type=ABSOLUTE; referencePt=NORTH; projective=TRUE</p>		

9.taula: *OLINK* etiketaren atributu eta balioak

¹⁰ Lotura-etiketen kasuan, MAE tresnarekin etiketatutako ingelesezko laginetik zuzenean hartuak edo moldatuak.

MOVELINK (Movement Link)

Mugimendua adierazten duten elementuak mugitzen diren elementuekin lotzen dituen harremana da. Hori horrela, *motion* etiketadun elementu baten beharra du, sortuko bada. Halaber, harreman honen etiketatzea mugimendu motaren (*motion_type*) araberakoa izango da.

Harreman-etiketa hau 10. taulako atributu eta balioen bidez zehazten da:

MOVELINK

Atributuak	Balioak	Lotura aktibatzen duen elementua
id (identifikatzailea)	mv1, mv2, mv3...	Mugimendua (<i>motion</i>) adierazten duen aditza
trigger	Mugimendua (<i>motion</i>) adierazten duen aditza, lotura eragiten duena.	
source	Abiapuntuko kokapena	
goal	Helburuko kokapena	
midPoint	Tarteko kokapena(k)	
mover	Leku-aldaketa jasaten duena	
ground	Zerekiko erlazioa duen mugitzen denak. Abiapuntuko kokapena eta helburuko kokapena agertzen ez direnean erabiltzen da (modua adierazten duten aditzekin du lotura)	
goal_reached	TRUE/FALSE/UNCERTAIN (egiazkoa/faltsua/zehaztu gabea)	
pathID	bidea	
motion_signalID	Mugimendu-harremanari bide edo modu informazioa ematen dion seinalea	
Adib.: “[Jonek _{sc2}] [kotxez _{ms3}] [bidaiatu _{m2}] zuen” trigger=m2; mover=sc2; motion_signalID=ms3		

10. taula: *MOVELINK* etiketaren atributu eta balioak

MLINK (Measure Link)

Harreman-etiketa honek bi helburu ditu. Batetik, kokapenen arteko distantzia atzemateko erabil daiteke eta, bestetik, objektu bakarraren dimentsioa deskriba dezake, kokapena, espazio-entitatea zein gertaera izan.

11. taulako atributu eta balioen bidez zehazten da harremana:

MLINK

Atributuak	Balioak	Lotura aktibatzen duen elementua
id (identifikatzailea)	ml1,ml2,ml3...	Neurri-egitura (<i>measure</i>); zehaztua nahiz zehaztu gabea
trajector	Erlazioa duen leku, entitate edo gertaera.	
landmark	Zein leku, bide, entitate edo gertaerarekiko duen harremana.	
reltype	DISTANCE, LENGHT, WIDTH, HEIGHT edo GENERAL_DIMENSION (distantzia, luzera, zabalera, altuera edo dimentsio orokorra)	
val	NEAR, FAR, TALLER, SHORTER edo MEASURE tag (hurbil, urrun, altuagoa, motzagoa edo neurria)	
endPoint1	Helburuko kokapena 1	
endPoint2	Helburuko kokapena 2	
<p>Adib.: “[Futbol zelaiak_{se2}] [100 metro_{me3}] ditu luzeran” reltype=LENGTH; trajector=se2; landmark=se2; val=me3</p>		

11. taula: *MLINK* etiketaren atributu eta balioak

METALINK

Maiz, hizkuntza naturaleko testuetan, entitate bereberei behin baino gehiagotan egiten zaie erreferentzia; edo, kasu batzuetan, entitate multzo bati egiten zaio erreferentzia, aurrez bere osagaiak bakarka edo azpialat gisa aipatu direnean. Multzo eta osagaien arteko harreman hori, berez, ez da kokapenezkoa naturan edo errealitatean baina, hala ere, garrantzitsua da jarraipena egitea. Helburu hori du, hain zuzen ere, harreman-etiketa honek.

Funtsean, objektuen artean espaziozkoak edo kokapenezkoak ez diren harremanak zehaztea du helburu erabilera anitzeko etiketa honek, 12. taulako atributu eta balioen bidez:

METALINK

Atributuak	Balioak
id (identifikatzailea)	meta1,meta2,meta3...
objectID1	Erlazioa duen leku, entitate edo gertaera.
objectID2	Zein leku, bide, entitate edo gertaerarekiko duen harremana.
reltype	COREFERENCE, SUBCOREFERENCE, SPLITCOREFERENCE (Korreferentzia, azpikorreferentzia, korreferentzia banatua)
<p>Adib.: “Bi [auto_{se1}] daude kalean. [Horietako_{se3}] [batek_{se2}] ezkerrera egin du” - objectID1=se3; objectID2=se1; reltype=COREFERENCE - objectID1=se2; objectID2=se1; reltype=SUBCOREFERENCE</p>	

12. taula: *METALINK* etiketaren atributu eta balioak

3 Ereduaren zailtasunak

Eredua ulertzeak eta erabiltzeak hainbat motatako zailtasunak sortu dizkigu. Batzuk etiketa jakin batzuen erabilerari lotuak (*non-consuming tag* eta *non-motion event* etiketak, kasu), eta beste batzuk, ereduaren garapenean izandako bertsio-aldaketak direla eta, etiketen izendapenetan gertatutako aldaketekin dute zerikusia (maiz, azalpenetan, adibideetan eta MAE tresnan emandakoak ez baitatoz bat).

Non-motion event

Ereduak eskaintzen dituen etiketen artean, konplexuena *non-motion events* izan dela esan genezake. Izan ere, hizkuntza jakin baten (kasu honetan ingelesaren) gramatikari erabat loturiko baldintzek bideratzen dute horrelakoen etiketatzea eta, beraz, egokitzapena puntuz puntu egitea egokitu zaigu. Funtsean, ereduak eskaintzen duen kasuistika zabala ulertzea eta testuan identifikatzea izan da zailtasuna. Etiketatutako ingelesezko corpusean adibide gehienak aditzei lotuak dira:

a) *So we didn't want to spend some more time there.*

b) *Along the valley Hondol at the right side we found an abandoned mine [...]*

a) eta b) adibideetako testuinguruetan, aditzek, mugimendurik adierazi ez arren, espazioarekin harremanetan daude (*there* eta *at the right side* elementuekin, hain zuzen ere, hurrenez hurren) eta, beraz, argi dago *non-motion event* etiketa jaso beharko luketela. Ingelesezko corpusean bertan, ordea, izan da aditzen etiketatzerik zalantzazkoa iruditu zaigunik:

c) *After crossing the border to Romania we took the road to Arad [...]*

c) esaldiko *took* aditzak, gure ustez, mugimendua adierazten du argi eta garbi, esaldiko adierazpena “bidea hartu” (*took the road*) den heinean.

Non-consuming tag

Antzerako zerbait gertatu zaigu elipsian dauden elementuak azaleratzeko helburua duen *non-consuming tag* etiketarekin ere. Izan ere, eskuartean izan ditugun azalpen eta adibideetan ez da guztiz argi geratzen elementu horiek azaleratzea noiz den garrantzizkoa eta noiz ez. Adibide gisa, etiketatzearen zergatiari buruzko zalantzak sortzen dizkiguten honako hauek (ingelesetik, bere horretan):

a) *A woman and a child are walking over the square* [Ø_{pt3}] [Ø_{pt4}] [Ø_{pt5}] [Ø_{pt6}].

Badirudi, adibide horretan egileek hasierako eta helburuko kokapen bana esleitzen dietela agertzen diren bi espazio-entitateei (*woman* eta *child*), bide “bera” eta “elkarrekin” egin ez dutelako interpretazioa atzeman nahian, antza.

b) *Departing Copala, the road **climbs** to over 6,000 feet in 30 miles, and then continues to **climb** while hugging an impressive cliff-lined ridgeline literally called ‘the spine of the devil.’ [Øpl12].*

Kasu honetan, berriz, helburuko kokapen bakarra azaleratu da (6,000 oin igo ondoren dagoena) eta, gure iritziz, ildo berean, “eskalatzen jarraitu” ondoren dagoena ere maila berean legoke, hau da, beste kokapen bat ere azaleratu zitekeen. Gainera, kasu honetan, biek ala biek *climb* aditza dute, beraz, bereizketarik behar ez delakoan gaude.

Horrelako adibideek etiketa honen erabileran sakondu beharra zenbaterainokoa den zalantzan jartzera eramán gaituzte. Horrez gain, kontuan izan behar dugu elipsi mota ezberdinak aurki ditzakegula (Estarrona, 2014) eta sailkapen horretan hizkuntza bakoitzak (kasu honetan, aditzek) bere joera erakutsiko duela. Hori horrela, hizpide dugun etiketari eta bere erabilerari buruzko azalpenak ez dira bere horretan euskarara ekartzeko modukoak. Ereduan etiketa honek duen garrantziaz jabetzeko zailtasunak direla eta, etorkizunean euskarazko lanetarako mugatu edo zehazteko kontua litzateke eta aipatu sailkapenari jarraiki egitea da gure proposamena¹¹:

1- Elipsi testuala

Mikel Indiara joan zen oporretan eta [Indiara] iritsi bezain pronto, damutu egin zen.

Testuari loturiko elipsia denez, esplizituki ageri ez den elementua, erreferentziakidetasuna baliatuz, testutik bertatik berreskura daiteke. Adibideko esaldiaren hasierari erreparatuta, *iritsi* zena *Mikel* dela ondoriozta dezakegu, eta baita *Indiara* (“nora”; helburuko kokapena) iritsi zela ere.

2- Elipsi pragmatikoa

Epaileak txartela atera dio.

Munduaz dugun ezagutza dela eta, hainbat argumentu ulertutzat ematen ditugu eta, beraz, ohikoena da horiek esplizituki ez agertzea, hots, elipsian dauden elementuak izatea (adibidean, txartela “nondik” eta “nora” atera duen epaileak) . Horrelakoek adiera berezi edo mugatuei eman ohi die bide, argumentuak ere jakinak direlako. Hori dela eta, maiz aurkituko ditugu halakoak testu mota jakin bat – corpus berezitu bat, alegia – aztertzean (adibidean, kirolaren testuinguruan).

3- Infraespezifikazio-elipsia

Ez du ondo abesten.

Honelakoetan, elipsia elementu bat espreski azpimarratu nahi ez izateagatik gertatzen da edo, bestela esanda, beste elementuren bat nabarmendu nahi delako.

Sailkapen horren harira, *non-consuming tag* etiketa elipsia testuala den kasuetarako soilik erabiltzea proposatzen dugu, elementuen arteko lotura zehaztea ahalbidetzen digun bakarra delako (eta, funtsean, hori da espazio-egituren etiketatzearen xedea), lotura duten elementuak testuan bertan aurki daitezkeen heinean. Izan ere, beste bi kasuetan elementu eliptikoak ez du ageriko erreferentziarik izango eta, beraz, ez dugu loturak egiteko beharrezko informaziorik emango.











¹¹ Azalpen eta adibideak sailkapen-proposamenaren iturritik hartuak (Estarrona, 2014)

Spatial_signal

Maila apalagoan bada ere, *spatial_signal* etiketa jasotzen duten elementuen mota semantikoak (*semantic_type*) bereizteak ere zailtasunak edo zalantzak sortu dizkigu. Esaterako, *edalontzia mahaiari dago* esaldian, ereduaren arabera harreman topologikoa zein norabidezkoa identifikatu behar litzateke (*DIR_TOP*). Izan ere, *edalontziak mahaiarekiko* kontaktu zuzena duela adierazten den heinean, informazio topologikoa ematen da; baina, era berean, kontaktu zuzena izate hutsak ez du adierazten *edalontzia mahaiari* gainean dagoen (aurreikus daitekeen harremana) edo beste punturen batean (ez da ohikoena, baina gerta daiteke).

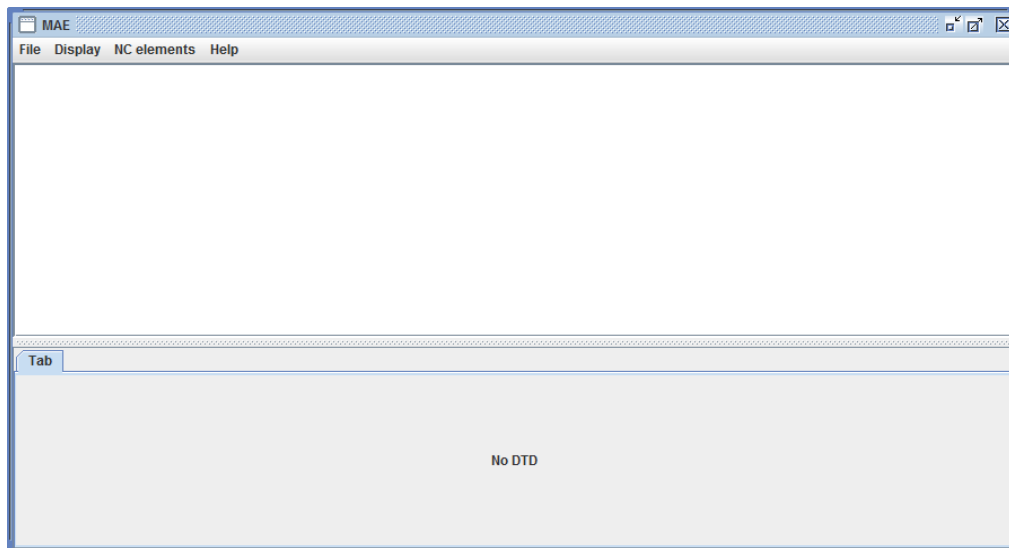
4 Etiketatzeko tresna: MAE

Etiketatzeko, bai ingelesezko ereduak eta bai guk, Amber Stubbs-ek sortutako MAE (*Multi-purpose annotation environment*) tresna erabili dugu. Egilearen orrialdean (<http://pages.cs.brandeis.edu/~astubbs/>) “MAE” atalera jota eskura daiteke, fitxategia dagoen lotura osagarrian (<http://code.google.com/p/mae-annotation/>). Horrela, “MAE_v0.9.6.zip.” izeneko zip fitxategia eskuratuko dugu. Behin bertatik fitxategiak erauzita, izen bereko karpeta izango dugu eta, bertan, besteak beste, “mae_v0.9.6.jar”, exekutatzeko fitxategia.

 samples	09/04/2012 22:40	Carpeta de archivos	
 etiketatzeko lagina	03/05/2016 13:06	Archivo TXT	1 KB
 etiketatzeko lagina	02/06/2016 18:45	Documento XML	3 KB
 ISO-Space-1.15.dtd	03/05/2016 12:51	Archivo DTD	8 KB
 license_GPL3	05/04/2012 10:06	Archivo TXT	35 KB
 MAE_User_Guide_0.9.5	10/04/2012 11:47	Archivo PDF	139 KB
 mae_v0.9.6	10/05/2012 20:56	Executable Jar File	2.675 KB
 README	10/05/2012 18:49	Archivo TXT	4 KB
 README.txt~	10/05/2012 18:49	Archivo TXT~	3 KB
 tag	02/06/2016 18:44	Data Base File	5 KB

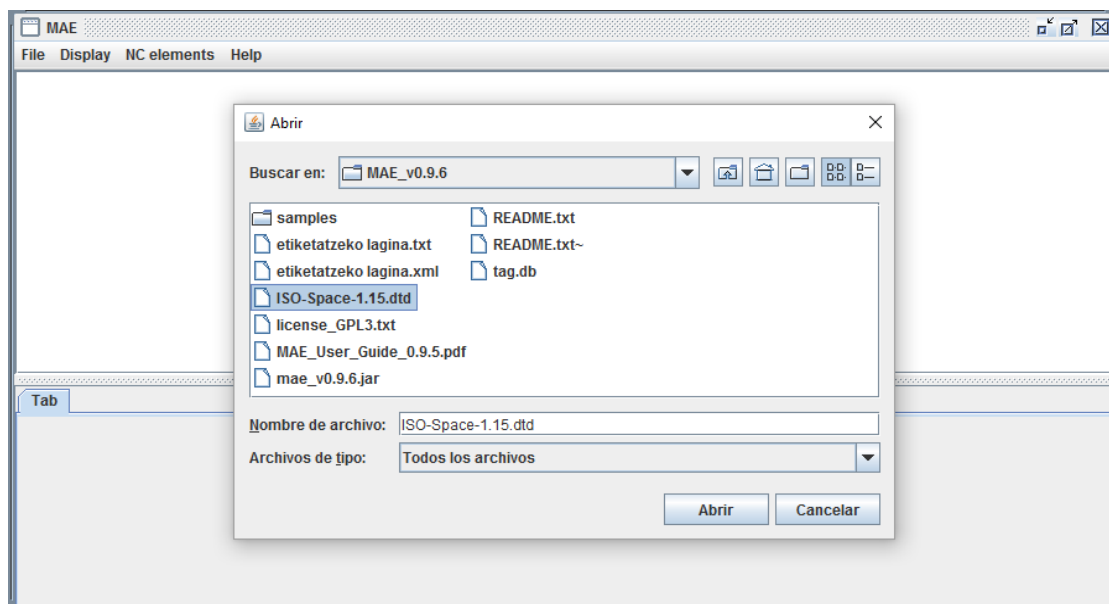
4.1 Tresnaren funtzionamendua

Behin dena kargatuta dugula, 1. irudian erakusten dugun moduko leiho bat zabalduko zaigu.



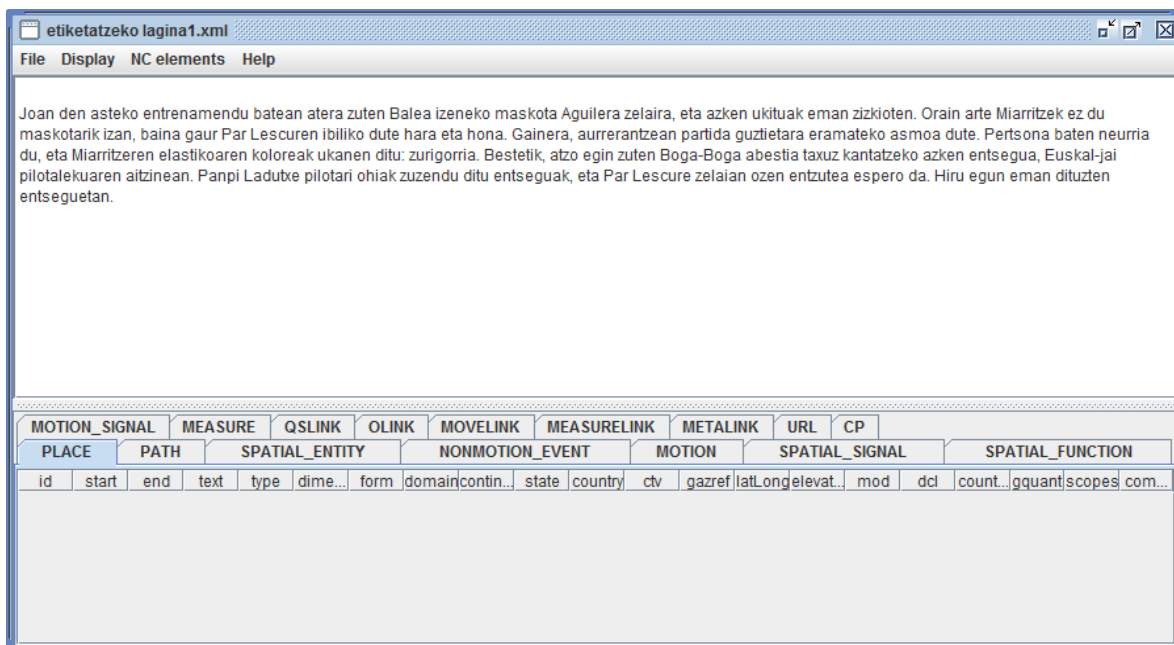
1. irudia: MAE tresna zabaltzean ikusiko dugun lehen leihoa.

Bertan kargatu behar dira etiketatu nahi diren testu edo dokumentuak; baina, etiketatzea egiteko, beharrezkoa da lehenengo “ISOSpace.dtd” eskuratzea, hots, ereduaren etiketa eta loturei buruzko informazio guztia biltzen duen fitxategia. Behin hori izanda, tresnaren barruan *file > load DTD > ISOSpace.dtd* urratsak jarraituko ditugu, 2. irudian erakusten den bezala:



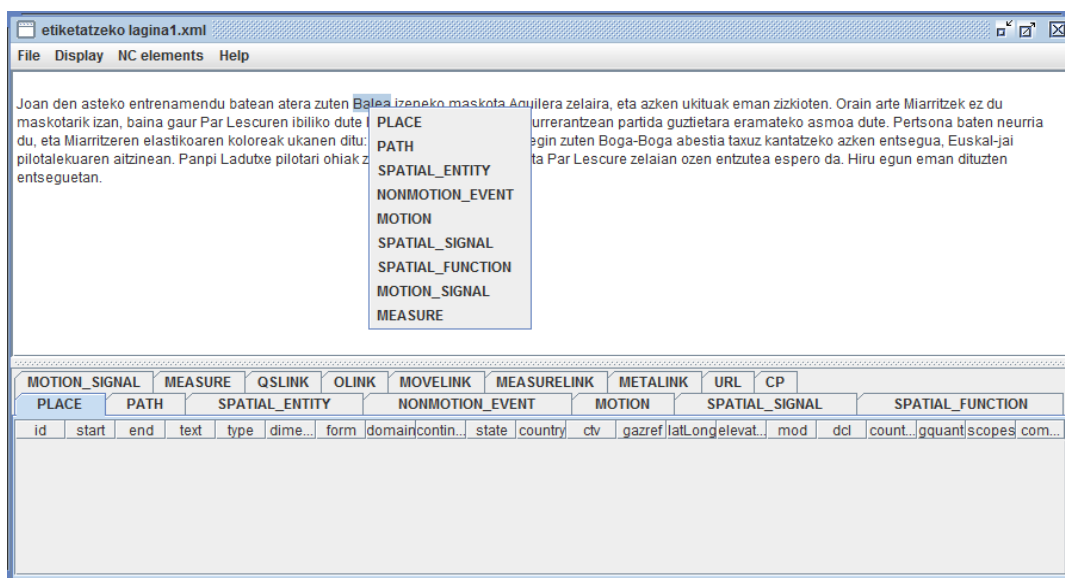
2.irudia: ISO-Space ereduaren dtd-a tresnan kargatzeko urratsak

Horren ondoren, etiketatu nahi dugun testu edo fitxategia kargatu (*file > load file > etiketatu nahi dugun fitxategia*) eta etiketatzeari ekitea besterik ez da geratzen; hori bai, etiketatu beharreko dokumentuak '.xml' formatukoa behar du izan tresnan erabili ahal izateko. Hori guztia kontuan izanda, 3. irudian erakusten dugun moduko zerbait izango dugu:



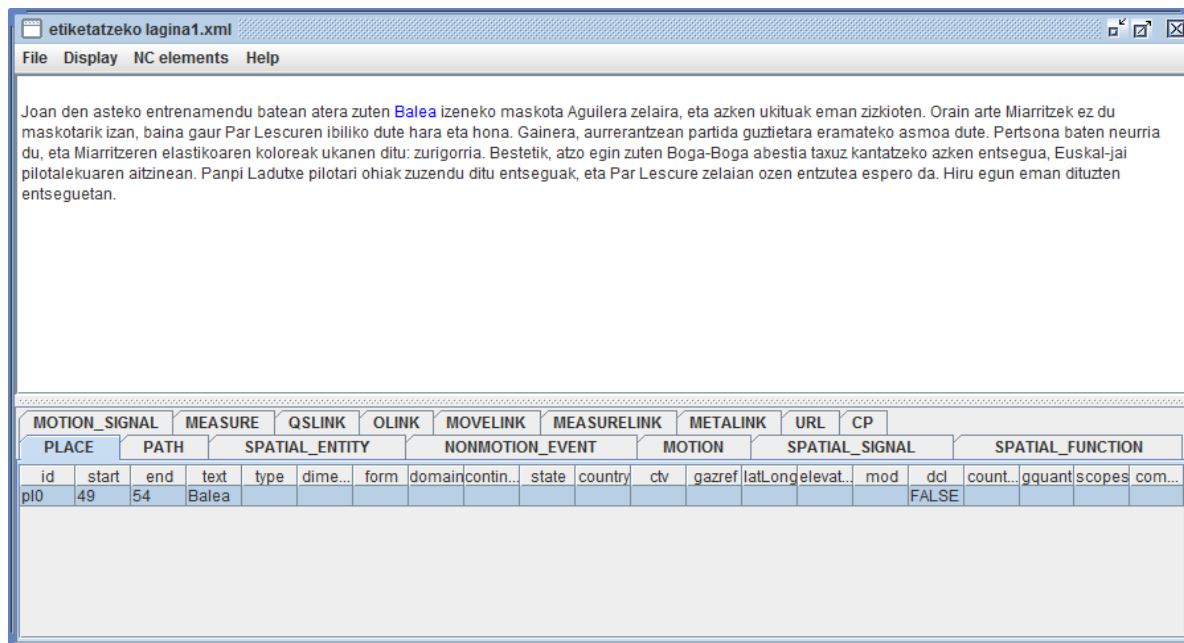
3.irudia: etiketatu nahi dugun lagina, tresnan kargatuta

4.3 atalean zehaztuko dugun bezala, ereduak proposatzen duen metodologiarekin bat egin dugu guk ere. Beraz, lehenengo pausua oinarritzko elementuak (*place, path, motion, spatial_signal...*) identifikatzea da. Horretarako, 4. irudian erakusten den bezala, etiketatu nahi den hitza aukeratu eta eskuineko botoia sakatuz, aukera guztiak biltzen dituen zerrenda agertuko zaigu, horien artean aukera dezagun:



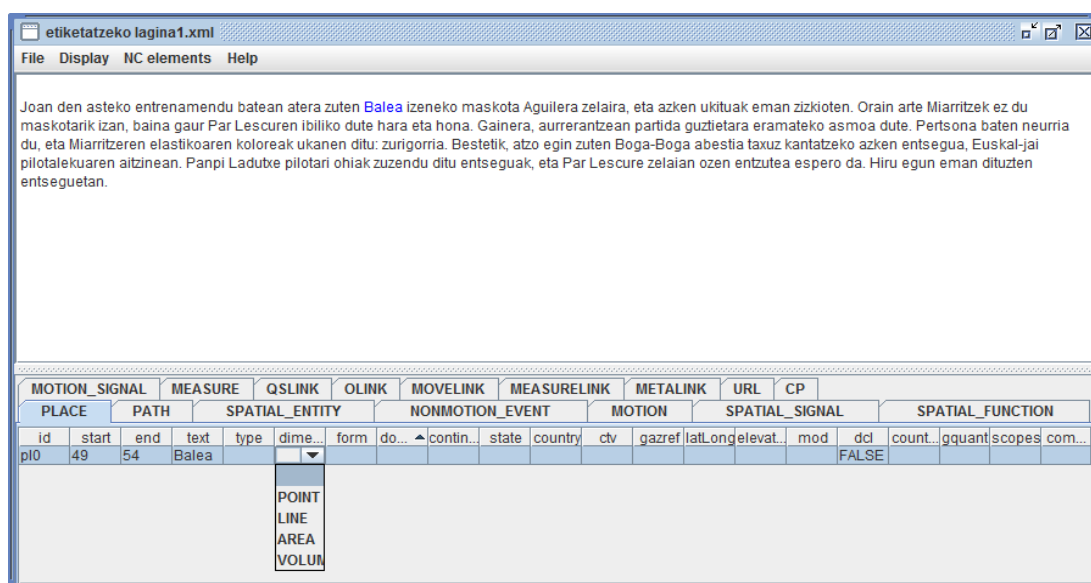
4. irudia: oinarritzko elementuen etiketatzea

Oinarrizko elementuak etiketatzean, bakoitzari dagokion etiketaren azpiko taulan agertuko dira, eta testu-hedadurak, etiketa motaren arabera, kolorez bereiziko dira automatikoki, 5. irudian ikusten den modura.



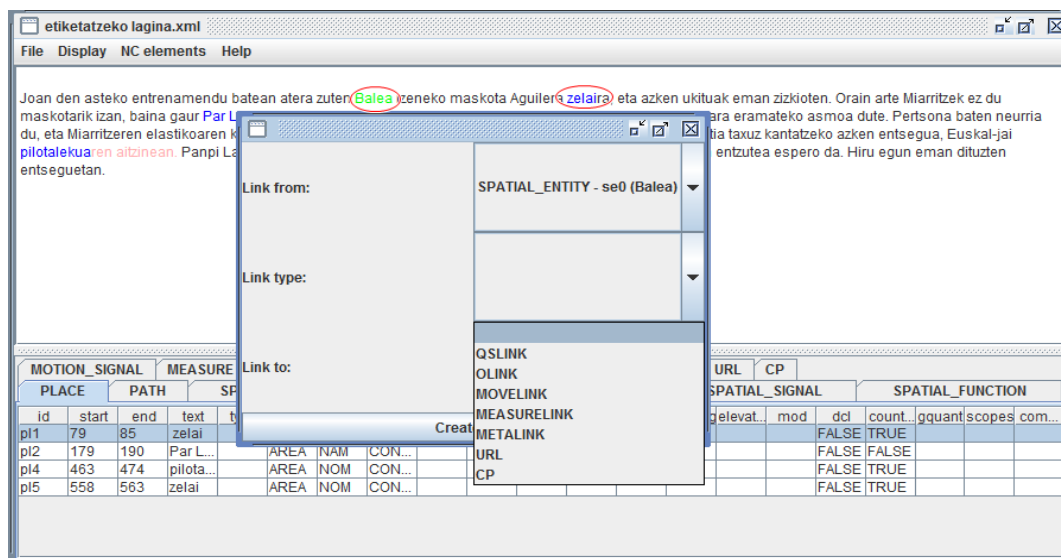
5.irudia: etiketatutako oinarrizko elementuek hartzen duten itxura eta informazioaren agerpena

Etiketa bakoitzaren baitako taulan, identifikatzailearen atala (*id*) automatikoki zehazten du tresnak baina gainerako atributu gehienak eskuz bete beharra dago. Dagokion atributuaren zutabearen, hutsik dagoen laukian klik egin eta, kasu gehienetan, balio posibleen zerrenda agertuko zaigu hemen ere, 6. irudian ikus daitekeen gisan:



6. irudia: oinarrizko elementuen atributu eta balioak zehazteko urratsak

Behin oinarrizko etiketak – euren atributu eta balio guztiekin – etiketatuta, euren arteko loturak ezartzea da hurrengo urratsa. Horretarako, lotu nahi den elementu bakoitza aukeratu eta 'Ctrl+ezker' eginez klik egingo dugu euren gainean. Hori eginda, lotura mota eta dagokion informazioa zehazteko aukera agertuko zaigu, 7. irudian agertzen den modura:



7.irudia: lotura-etiketak ezartzeko eta zehazteko urratsak

Aurreko urratsak jarraituz, oinarrizko elementuak zein euren arteko lotura guztiak zehaztuta ditugunean, amaitutzat emango dugu gure espazio-egituren etiketatzea (gure etiketatzearen emaitza eranskin gisa aurkeztuko dugu, '.xml' formatuan).

4.2 Zailtasunak

Atal honetan, proiektuaren alderdi praktikoa (hau da, etiketatze-prozesuan bertan) sortu zaizkigun arazo edo zailtasunak azalduko ditugu.

Batetik, elementu eliptikoen azalera aipatu beharrean gara atal honetan ere. Izan ere, etiketatze-lanetan ez dugu *non-consuming tag* delakoak nola sortzen diren jakiteko aukerarik izan, ez azalpen teorikoetan, ezta tresnan bertan ere (ez baita beste etiketa guztiekin batera eskaintzen). Aukera hori ez izateak lotura edo harremanak zehazteko garaian eragin nabarmena duela uste dugu. Egia da, halaber, ingelesezko corpus etiketatuan bertan ere ez dugula etiketa honen erabilerarik topatu (teoriaren eta praktikaren garapena modu paraleloan gauzatzen ari ez diren beste adibide bat izan daiteke).

Dena den, gure laginean espazio batekiko harremana izan dezakeen elipsiaren adibide bakarra aurkitu dugu:

Joan den asteko entrenamendu batean atera zuten Balea izeneko maskota Aguilera zelaira.

Adibide horretan, *atera* aditzaren helburuko kokapena esplizituki ageri da (*Aguilera zelaia*), ez ordea hasierako kokapena (“nondik” *atera* duten, alegia). Informazio hori testutik bertatik

interpretatzeko modukoa ez denez – hots, elipsia testuala ez denez – 3. atalean azaldu dugun proposamenari jarraiki, elipsian dagoen elementua ez genuke azaleratuko kasu honetan (hau da, ez genuke *non-consuming tag* etiketarik sortuko).

Halaber, gure laginean zalantzak sortu dizkigun *gertaera* (*non-motion event*) kasu bat ere aurkitu dugu:

Atzo egin zuten Boga-Boga abestia taxuz kantatzeko azken entsegua [...]

Izan ere, gertaera bati buruz ari dela argi izan badugu ere, etiketa zein elementuri (*egin* aditzari edo *entsegua* izenari) esleitu erabakitzea izan da zailtasuna. Etiketa hau azaldu dugun puntuan ikusi bezala, izenek testuinguru jakinetan agertu behar dute markatuko badira. Zentzu hertsian hartuz gero, beraz, goiko esaldiko *entsegua* izena ez genuke markatu behar. Alabaina, aipatu testuinguruetan agertzen den egituretako batekiko (3- IZENA + denborazko adierazpena + *gertatu zen*) hurbiltasun edo antzekotasuna dela eta, kasu honetan, izena markatzea izan da gure erabakia.

Bestalde, MAE tresna erabiltzean, eskuartean izan ditugun azalpenetan ageri ez ziren edo azalpen argirik ez duten hainbat etiketa, atributu edo balioekin egin dugu topo; jarraian zerrendatu ditugunak, besteak beste:

- *domain* atributua, *content* eta *structure* balioekin.
- *spatial_function*, *URL* eta *CP* etiketak.
- *elevation* atributua zenbait etiketatan (*spatial_entity* eta *non-motion event* etiketetan, esaterako)
- *Place* etiketaren *type* atributua.
- *gquant* eta *scopes* atributuak.

Horren aurrean, egitasmo honetan azaldu eta euskarazko eredurako proposatu ditugun erabilerak baino ez ditugu izan kontuan gure etiketatze-lanetan eta gainerakoak hutsik utzi ditugu, ezinbestean. Hala ere, etorkizuneko lanetarako (batez ere MAE tresna erabiltzeko hautua egiten bada) argitu beharreko berrikuntzak direla uste dugu.

4.3 Etiketatze-prozesua: metodologia, kopurua eta denbora

Lehen saiakera honetan, etiketatze-sistema praktikara eramateko, EPEC-Euskararen Prozesamendurako Erreferentzia Corpora (Aduriz et al., 2006) corpuseko lagin labur bat hautatu dugu, ausaz. Ereduko etiketen sailkapenari jarraituz, egileen proposamenarekin bat, esan bezala, etiketatzea bi fasetan gauzatzea proposatzen dugu guk ere: 1) oinarrizko elementuak etiketatzea; 2) euren arteko loturak ezartzea.

Etiketatu dugun lagina zazpi esaldikoa da eta horiek etiketatzeko, esaldiko 5 minutu inguru eman ditugu, batez beste.

5 Ondorioak

Egitasmo hau burutu ostean, ondoriorik bistakoena edo azpimarragarriena eredu egokitzeko aukera da. Hau da, aztergai izan dugun eredu, hizkuntza ezberdinetarako modukoa izatea (zailtasunak zailtasun). Aipatutako koreerarako egokitzapenaren artikuluan, txinerarako egokitzapena ere aurkeztu zen (hori dela eta da *multilingual*, hots, eleanitza), tipologikoki ezberdinak diren hizkuntzetan erabilgarria den eredu dela frogatuz (Lee et al., 2011).

Bestalde, proiektu honetan ereduaren egileek etiketatzeko proposatutako metodologiarekin – hots, espazio-egituren etiketatzea bi fasetan (lehenengoan oinarrizko etiketak eta bigarreanean harremanak) egitearekin – bat egin dugu. Aitortu behar dugu, etiketatu beharreko testuaren lehen irakurraldian, errazagoa dela espazioen arteko harremanak identifikatzea; eta horren arabera topatzea harreman horietan parte hartzen duten elementuak. Praktikan, ordea, partaideak definituta izateak harremanen etiketatzea errazten du eta hori izan da metodologiarekin bat egiteko arrazoi nagusia.

Halaber, nabaria da aztergai dugun arloa hastapenetan dagoela. Oraindik orain, kontzeptu askoren esanahia topatzea zaila da teoria edo azalpenetan eta garatu diren teoretako informazio guztia ez da beti islatzen praktikan. Horrez gain, etiketatzeko tresnaren (MAE) garapena ez doala garapen teorikoarekin batera ikusi dugu; are gehiago, egileek gidalerroen txostenean aurkezten duten tresnaren bertsioa eta eskuragarri jartzen dutena ere ez datoz guztiz bat.

Esan beharra dago arlo honetan euskaraz egin den lehen urratsa izan dela gure proiektu hau, eta beraz, filosofia orokortzat har badezakegu ere, etorkizuneko aplikazio edo egitasmoen helburu edo nolakotasunaren arabera izango dira beharrak. Hori horrela, atzeman nahi den informazio linguistikoaren arabera, hemen hasitako egokitzapena gehiago sakontzeko beharra eta, batez ere, aipaturiko zailtasunak aztertu beharra suerta daiteke.

Bibliografia

- Aduriz I., Aranzabe M., Arriola J., Atutxa A., Díaz de Ilarraza A., Ezeiza N., Gojenola K., Oronoz M., Soroa A. and 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. In Andrew Wilson, Paul Rayson and Dwan Archer ed. *Corpus Linguistics Around the World*. Book series: Language and Computers. Vol. 56, 1-15. ISBN 90-420-1836-4. Rodopi. Netherlands.
- Estarrona, A. *EPEC corpusa predikatu-mailan etiketatzeko oinarriak: EPEC-RolSem, BVI eta e-ROLda*. Euskal Hizkuntza eta Komunikazioa Saila (UPV-EHU). Informatika Fakultatean aurkeztua 2014ko irailaren 4an. Donostia. Tesiker. ISBN: 978-84-8438-564-6.
- Lee K., Fang A.C and Pustejovsky J. (2011). Multilingual Verification of the Annotation Scheme ISO-Space. *Semantic Computing (ICSC), 2011 Fifth IEEE International Conference*, 449 – 458.
- Kolomiyets O. , Kordjamshidi P. , Bethard S. and Moens M-F. (2013). SemEval-2013 Task3: SpatialRoleLabeling. *In Proceedings of the seventh international workshop on semantic evaluation (SemEval 2013)*, 255-266. ACL.
- Kordjamshidi P., Bethard S. and Moens M-F. (2012). SemEval-2012 Task3: SpatialRoleLabeling. *In Proceedings of the 6th International Workshop on Semantic Evaluation (SemEval)*, 365-373.
- Kordjamshidi P., VanOtterlo M. and Moens M-F (2010). SpatialRoleLabeling: Task Definition and AnnotationScheme. *Proceedings of the Seventh conference on International Language Resources and Evaluation (LREC'10)*. 413-420. ISBN 2-9517408-6-7. European Language Resources Association (ELRA).
- Mani I., Doran C., Harris D., Hitzeman J., Quimby R, Richer J., Wellner B., Mardis S. and Clancy S. (2010). SpatialML: annotation scheme, resources, and evaluation. *Language Resources and Evaluation*, Volume 44, Issue 3, 263-280. Springer Science+Business Media.
- Pustejovsky J. , Moszkowicz J. and Verhagen M. (2012). A Linguistically Grounded Annotation Language for Spatial Information. *Traitement Automatique des Langues (TAL)*, Volume 53 - n° 2/2012, 87-113.
- Pustejovsky J., Kordjamshidi P., Moens M-F. , Levine A. , Dworman S. and Yocum Z. (2015) SemEval-2015 Task8: SpaceEval. *In Proceedings of the 9th International Workshop on Semantic Evaluation. Association for Computational Linguistics (ACL)*.
- SpaceEval Annotation Guidelines (2014). <http://jamespusto.com/blog/category/news/>

Eranskina: etiketatutako lagina

Sarreran aipatu bezala, eranskin honetan, aurkeztu dugun ereduari jarraiki etiketatutako lagina eskainiko dugu, xml formatuan.

```
<?xml version="1.0" encoding="UTF-8" ?>
<ISOSpaceTaskv1.15>
<TEXT><![CDATA[
Joan den asteko entrenamendu batean atera zuten Balea izeneko maskota
Aguilera zelaira, eta azken ukituak eman zizkioten. Orain arte Miarritzek
ez du maskotarik izan, baina gaur Par Lescuren ibiliko dute hara eta
hona. Gainera, aurrerantzean partida guztietara eramateko asmoa dute.
Pertsona baten neurria du, eta Miarritzeren elastikoaren koloreak ukanen
ditu: zurigorria. Bestetik, atzo egin zuten Boga-Boga abestia taxuz
kantatzeko azken entsegua, Euskal-jai pilotalekuaren aitzinean. Panpi
Ladutxe pilotari ohiak zuzendu ditu entseguak, eta Par Lescure zelaian
ozen entzutea espero da. Hiru egun eman dituzten entseguetan.
]]></TEXT>
<TAGS>
<PLACE id="pl1" start="79" end="85" text="zelai" type=""
dimensionality="AREA" form="NOM" domain="" continent="" state=""
country="" ctv="" gazref="" latLong="" elevation="" mod="" dcl="FALSE"
countable="TRUE" gquant="" scopes="" comment="" />
<PLACE id="pl2" start="179" end="190" text="Par Lescure" type=""
dimensionality="AREA" form="NAM" domain="" continent="" state=""
country="" ctv="" gazref="" latLong="" elevation="" mod="" dcl="FALSE"
countable="FALSE" gquant="" scopes="" comment="" />
<PLACE id="pl4" start="463" end="474" text="pilotalekua" type=""
dimensionality="AREA" form="NOM" domain="" continent="" state=""
country="" ctv="" gazref="" latLong="" elevation="" mod="" dcl="FALSE"
countable="TRUE" gquant="" scopes="" comment="" />
<PLACE id="pl5" start="558" end="563" text="zelai" type=""
dimensionality="AREA" form="NOM" domain="" continent="" state=""
country="" ctv="" gazref="" latLong="" elevation="" mod="" dcl="FALSE"
countable="TRUE" gquant="" scopes="" comment="" />
<SPATIAL_ENTITY id="se0" start="49" end="54" text="Balea" type="PERSON"
dimensionality="VOLUME" form="NAM" domain="" latLong="" elevation=""
mod="" dcl="FALSE" countable="TRUE" gquant="" scopes="" comment="" />
<SPATIAL_ENTITY id="se1" start="243" end="250" text="partida"
type="DYNAMIC_EVENT" dimensionality="POINT" form="NOM" domain=""
latLong="" elevation="" mod="" dcl="FALSE" countable="TRUE" gquant=""
scopes="" comment="" />
<MOTION id="m0" start="37" end="42" text="atera" domain="" latLong=""
elevation="" motion_type="PATH" motion_class="MOVE_EXTERNAL"
motion_sense="LITERAL" mod="" countable="FALSE" gquant="" scopes=""
comment="" />
<MOTION id="m1" start="262" end="271" text="eramateko" domain=""
latLong="" elevation="" motion_type="PATH" motion_class="MOVE"
motion_sense="LITERAL" mod="" countable="FALSE" gquant="" scopes=""
comment="" />
```



```
<SPATIAL_SIGNAL id="s0" start="474" end="488" text="ren aitzinean."
cluster="" semantic_type="DIRECTIONAL" comment="" />
<SPATIAL_SIGNAL id="s1" start="563" end="565" text="an" cluster=""
semantic_type="TOPOLOGICAL" comment="" />
<SPATIAL_SIGNAL id="s2" start="190" end="191" text="n" cluster=""
semantic_type="TOPOLOGICAL" comment="" />
<MOTION_SIGNAL id="ms0" start="85" end="87" text="ra" adjunct_type="PATH"
comment="" />
<MOTION_SIGNAL id="ms1" start="256" end="261" text="etara"
adjunct_type="PATH" comment="" />
<MOVELINK id="mv12" fromID="pl1" fromText="zelai" toID="pl1" toText="
zelai" trigger="atera" source="" goal="zelai" midPoint="" mover="Balea"
landmark="" goal_reached="YES" pathID="" motion_signalID="" comment="" />
<MOVELINK id="mv11" fromID="se1" fromText="partida" toID="m1"
toText="eramateko" trigger="eramateko" source="" goal="partida"
midPoint="" mover="" landmark="" goal_reached="NO" pathID=""
motion_signalID="" comment="" />
</TAGS>
</ISOSpaceTaskv1.15>
```