

Mugarri: bigarren
hizkuntzako ikasleen
hizkuntz ezagutza
eskuratzeko sistema
anitzeko ingurunea

Egilea: **MONTSE MARITXALAR ANGLADA**
Urtea: 1999
Zuzendaria: ARANTZA DIAZ DE ILARRAZA SANCHEZ
Unibertsitatea: UPV/EHU
ISBN: 978-84-8438-219-5

Hitzaurrea

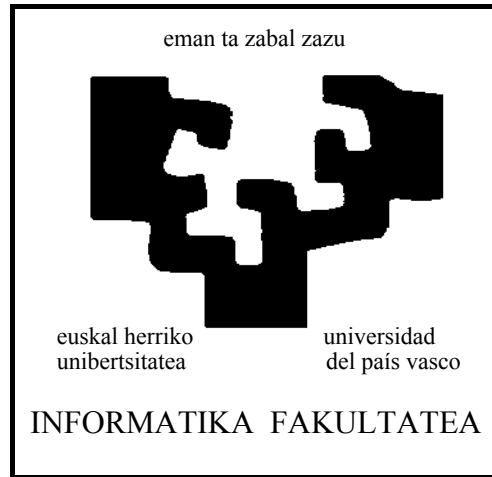
Euskararen tratamendu automatikoak hizkuntzen irakaskuntza arloan ekar zitzakeen aldaketak aztertzea izan tesi honen hasierako motibazioa. Ordenagailuek ikasleak nola lagun zitzaketen izan genuen lehen azterbidea. Ondoren, konturatu ginen bai irakasleak bai ikerlariak laguntzeko tresna lagungarriak ere sor litezkeela hizkuntzaren tratamendu automatikoa integratzen bagenuen gure tresna informatikoetan. Une hartan hizkuntzaren tratamendu automatikoa eta hizkuntzen ikaskuntza arloek ez zutela elkar ezagutzen esan dezakegu. Informatika arloan irakasteko tutoreak sortzen hasiak ziren, baina honelako gutxik erabiltzen zuten hizkuntzaren tratamendu automatikoa. Guk, adimen artifizialeko ikerlerro honetan ematen ziren aurrera pausuak handiagoak izan zitezkeela aurrikusten genuen, beti ere baldin eta hizkuntzaren tratamendu automatikoarekin osatzen bagenituen. Hau izan zen une hartako ikerbidea, baita gaur egun jarraitzen duguna ere.

Une honetan, egitasmo berriei ere ekin diegu, konkretuki, eduki didaktikoen sorkuntza automatikoa dugu aztergai. Hizkuntza bat ikasteko ariketak automatikoki sortzen ditugu euskaraz idatzitako testuak ariketa-iturriak izanik. Ikerlerro honi garrantzia ematen diegu egun dagoen behar bati erantzuten badio. Irakasleek askotan ariketak asmatu beharrean dabilta, eta beraiek sortutako ariketak “artifizialak” izaten dira askotan. Testu errealetan oinarrituz gero, berriz, errealitate bati erantzuten dion materiala sortzen ari gara. Gainera, sorkuntza automatikoaren bitartez dirua eta denbora aurrezteko aukera handia dugu.

Honetaz gain, tesia garatu zeneko garaian hasitako hainbat ikerlerroetan ere lanean dihardugu gaur egun, hots, erroreen tratamendu automatikoa, idazketa-prozesurako tresna lagungarriak eta irakasleentzako baliabideak ikasleen ikaste-prozesuaren jarraipena egin ahal izateko. Besteak beste, idazlanen ebaluazioa ere aztergai dugu gure ikerketa proiektuen artean. Garbi izan behar dugu ordenagailuak ezin dituela detektatu gizakiak adina ezaugarri idazlan batean, baina makinak lana erraz diezaioke irakasleari zein ikasleari idazlan baten ebaluazio prozesuan. Esaterako, *-lako* egitura behin eta berriro erabiltzen duenean ikasle batek, gramatika aberasteko gomendioak eman ditzake makinak *bait-*, *-z gero* eta honelakoak proposatuz.

Azkenik, aipatu beharra dago euskara nagusiki aztertzen dugun arren, gero eta sistema eleaniztun gehiago garatzen ari garela: ariketen sorkuntzaren kasuan ingelesarekin ere bagabiltza eta idazlanen ebaluazioan Kubako ikerketa talde batekin ari gara elkarlanean hango gazteleraz idatzitako idazlanak ebaluatu nahian.

LENGOAIA ETA SISTEMA INFORMATIKOEN SAILA



MUGARRI: BIGARREN HIZKUNTZAKO IKASLEEN HIZKUNTZ EZAGUTZA ESKURATZEKO SISTEMA ANITZEKO INGURUNEA

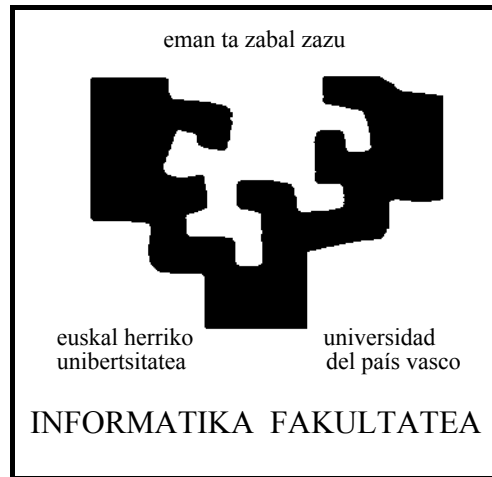
Montse Maritxalar Angladak

Informatikan Doktore titulua eskuratzeko aurkezturiko

TESI-TXOSTENA

Donostia, 1999ko uztaila.

LENGOAIA ETA SISTEMA INFORMATIKOEN SAILA



MUGARRI: BIGARREN HIZKUNTZAKO IKASLEEN HIZKUNTZ EZAGUTZA ESKURATZEKO SISTEMA ANITZEKO INGURUNEA

Montse Maritxalar Angladak Arantza Díaz de Ilarrazaren zuzendaritzapean egindako tesiaren txostena, Euskal Herriko Unibertsitatean Informatikan Doktore titulua eskuratzeko aurkeztua.

Donostia, 1999ko uztaila.



Nire amari eta nire aitari

“Elementuak banan banan hartuz gero, ez dira biak bat izango, baina bai osotasunean; ez azalean, baina bai muinean, argia baita espazioaren bihotza eta ilargiak argi irreal berberaz jazten baititu bi espazioak.”

(Consummatum est, J.M. Irigoien)

ESKERRAK EMANEZ

Arantzari, lan honen hasierako cimiento eta tartekoak jarri izanagatik.

Iñaki, Isabel, Kepa, Xarregi eta Xartolari, besteak beste, lan honen idazketan beraiengandik jasotako oharrengatik.

Maiteri, pantaila aurrean MUGARRI errealitate bihurtzen laguntzeagatik. Honegatik Aitor, Koldo eta Nereari ere eskerrak.

IXAkide guztioi (Izaskun, Jose Mari, Ruben, ...) edozein unetan behar txiki handi horiei erantzuteko prest egoteagatik.

Idoiari, emandako bultzadagatik.

Hizkuntzen ikaskuntza munduan hainbeste erakutsi didazutenoi: Pika, ILAZKIk irakasleak, eta Edinburgoko Unibertsitateko Linguistika Aplikatua sailaren eta HABEren liburutegiak.

Anari, gauza askorengatik.

Itziarri, pixukideari. Desanbigua ezan!. Eta ex-pisukideei.

Aipatu ditudan eta ez ditudan fakultatekide guzti horiei, indarrak emateagatik.

Nire familiari, hor egoteagatik.

Eta, azkenik, Carlosi.

I. Motibazioa eta aurkezpena.	9
I.1 Motibazioa eta helburuak.	9
I.2 Metodologia: diziplinartekoa eta definizio-diseinu-inplementazio- ebaluazio zikloan oinarritua.	12
I.3 MUGARRI ingurunearen aurkezpena.	14
I.3.1 Ingurunearen definizioa.	14
I.3.2 Ingurunearen arkitektura.	16
I.3.3 Hiru sistemen arteko komunikazioa.	18
I.4 Txostenaren eskema.	21
II. Hizkuntzen ikaste-prozesua lantzeko sistemen azterketa.	23
II.1 Sarrera.	23
II.2 Azterketa bibliografikoa.	24
II.2.1 Ordenagailuz Lagundutako Hizkuntzen Ikaskuntza tradizionala.	24
II.2.2 Sare bidezko hizkuntz ikaskuntza.	26
II.2.3 Multimedia tekniken aplikazioa hizkuntzen ikaskuntzan.	27
II.2.4 Ordenagailuz Lagundutako Hizkuntzen Ikaskuntza Adimenduna.	28
II.2.5 Lengoaia Naturalaren Prozesamendua integratu duten sistemak.	29
II.3 Ondorioa: diziplinarteko garapena.	31
III. Psikolinguistika arloan egindako lan esperimentalaren deskribapena.	33
III.1 Helburuak.	33
III.2 Tartehizkuntza eta fosilizazioa.	34
III.2.1 Tartehizkuntza.	34
III.2.2 Fosilizazioa.	36
III.3 Desbideratze kontzeptua. Sailkapenak.	37
III.3.1 Erroreen garrantzia hizkuntza baten ikaste-prozesuan.	37
III.3.2 Desbideratze kontzeptua.	38
III.3.3 Desbideratzeen sailkapena.	41
III.3.3.1 Aztertutako corpusak.	41
III.3.3.2 Azaleko versus sakoneko taxonomiak.	43
III.3.3.2.1 Azaleko maila hutsezko desbideratzeak.	44
III.3.3.2.2 Desbideratze linguistikoak.	46
III.3.3.2.3 Sakoneko desbideratzeak.	47
III.4 Esperimentuetan erabilitako metodologia.	49
III.4.1 Burututako esperimenduak.	49
III.4.2 Azterketa esperimental.	50
III.4.2.1 Helburuak.	50
III.4.2.2 Metodologiaren zehaztapena.	51
III.4.2.3 Azterketarako aukeraturako ikasleak zehazteko irizpideak.	52
III.4.2.4 Egitura linguistikoaren aukeraketaren irizpideak.	53
III.5 Ondorioa: sistema erdiautomatiko baten beharra.	54
IV. Tartehizkuntza modelizatzeko sortu eta egokitutako tresna linguistiko-konputazionalak.	57
IV.1 Tartehizkuntzaren modelizazioa corpusetatik.	58
IV.2 Domeinu-ereduaren adierazpidea.	59

IV.2.1 Tartehizkuntz ereduak. Printzipioak.....	59
IV.2.2 KADS metodologia.....	62
IV.2.3 Tartehizkuntz ereduaren deskribapena.....	63
IV.3 Ikaskuntza-mailako ereduaren eraikuntzarako erabilitako metodologia.....	74
IV.3.1 Goitik beherako metodologia.....	74
IV.3.2 Ezagutza finkoa eta aldakorraren identifikaziorako irizpideak.....	76
IV.4 Ikasle eta ikaskuntza-mailako ereduaren hizkuntz ezagutza detektatzeko tresnen deskribapena.....	77
IV.4.1 Tartehizkuntzaren erregela eta fenomeno linguistikoak detektatzeko tresnen deskribapena.....	79
IV.4.2 Ikaskuntza-mailari dagozkion tresnen ezagutzaren eguneraketa sistemaren autoikasketa dela medio.....	88
IV.5 Analizatzaile morfologikoaren egokitzapena.....	90
IV.6 Inplementazio-egoera eta landu beharrak.....	93

V. HITES: bigarren Hizkuntzaren Ikaste-prozesuan bereganatzen den Tartehizkuntz Eredua modelizatzeko Sistema adimenduna.

V.1 Sarrera.....	97
V.2 HITES, IDAZKIDE eta IRAKAZI: elkarrekin harremanetan dauden sistema adimendunak.....	98
V.2.1 Sistemen definizioa.....	98
V.2.2 Ingurunearen arkitektura.....	99
V.2.3 Hiru sistemen arteko komunikazioa.....	101
V.3 HITESen domeinuaren definizioa.....	102
V.3.1 Tartehizkuntz ereduak.....	102
V.3.2 Ikaslearen ereduak.....	103
V.4 HITESen arkitektura.....	105
V.4.1 HITESen arkitekturaren osagaiak.....	105
V.4.2 Ezagutza-baseak.....	107
V.4.2.1 Lan-memoria.....	107
V.4.2.2 Helburuak eta atazak.....	111
V.4.3 Arrazonatze- eta erabakitze-modulua.....	120
V.4.4 Interfazearen kontrolatzailea: psikolinguistarekiko elkarrekintza.....	123
V.5 Ikaskuntza HITESen.....	125
V.5.1 Ikaskuntza-elementua HITESen.....	125
V.5.2 Ikaskuntza-ataza: adibidea.....	126
V.6 HITES: ikaskuntzarako ahalmena duen agente bat.....	127
V.6.1 HITES agenteen ikuspegitik.....	127
V.6.2 COSY agenteanitzaren arkitektura.....	129
V.6.3 COSY: HITESi egoki dakiokkeen arkitektura.....	130
V.7 Objektuei zuzendutako programazioan oinarritutako inplementazioa.....	131
V.7.1 Ikaskuntza-inguruneke objektuen hierarkia: garapena HITESen.....	131
V.7.2 Exekuzio-adibide bat HITESen.....	134
V.8 Inplementazio-egoera eta landu beharrak.....	141

VI. IDAZKIDE eta IRAKAZI:	143
idazte-prozesurako eta diagnosirako laguntza-sistemak.	143
VI.1 Sarrera.....	143
VI.2 Ikasle-eredu komuna.....	144
VI.2.1 Ezagutza-baseak.....	144
VI.2.2 Ikasle-ereduaren gunea.....	145
VI.2.3 Diagnosi-ezagutza.....	147
VI.3 IDAZKIDE: idazte-prozesurako sistema lagungarria.....	148
VI.3.1 IDAZKIDEn arkitektura.....	148
VI.3.2 Antolatzailea, aholkularia eta gainbegiratzailea.....	150
VI.3.3 Tresna linguistikoaren integrazioa IDAZKIDEn.....	151
VI.3.4 Ikasle-eredua.....	153
VI.3.4.1 Ikasle-ereduak sistema adimendunetan.....	153
VI.3.4.2 Ikasle-eredua IDAZKIDEn.....	155
VI.3.5 Ikaste-estrategiak.....	159
VI.3.5.1 Ikaste-estrategiak sistema informatikoetan.....	159
VI.3.5.2 Ikaste-estrategiak IDAZKIDEn.....	163
VI.3.6 IDAZKIDEn interfazea.....	167
VI.4 IRAKAZI: diagnosirako sistema lagungarria.....	170
VI.4.1 IRAKAZIren interfazea.....	172
VI.5 Inplementazio-egoera eta landu beharrak.....	175
VII. MUGARRI ingurunearen garapenean zeharreko ebaluazioa.	177
VII.1 Sarrera.....	177
VII.2 Lan enpirikoaren ebaluazioa eta emaitzak.....	179
VII.2.1 Hitz-mailako azterketa ikaskuntza-mailaren arabera.....	179
VII.2.2 Esaldi-mailako azterketa ikaskuntza-mailaren arabera.....	184
VII.2.3 Ezagutza linguistikoaren egonkortasunaren azterketa.....	190
VII.2.3.1 Goiko ikaskuntza-maila altutik goiko maila baxura.....	190
VII.2.3.2 Ariketa-motaren eragina.....	192
VII.3 Ingurune informatikoaren emaitzak.....	193
VII.3.1 Erabilitako corpusak.....	194
VII.3.2 Aplikatutako metodologiaren berrespena.....	195
VII.3.3 Tartehizkuntzaren azterketa automatikoa.....	200
VII.3.3.1 Fenomeno morfosintaktikoen identifikazio erdiautomatikoa.....	200
VII.3.3.2 Hiztegiaren aberastasuna.....	201
VII.3.3.3 Katgoria gramatikalen erabilera.....	204
VII.3.3.4 Esaldien konplexutasunaren azterketa.....	206
VIII. Ondorioak eta zabaldutako ikerlerroak.	209
VIII.1 Emaitzak eta ekarpenak.....	209
VIII.2 Ondorioak.....	212
VIII.2.1 Diziplinarteko lanaren ondorio metodologikoak.....	212
VIII.2.2 Bigarren hizkuntzaren ikaste-prozesua ikertzeko MUGARRI ingurunearen garapenerako irizpideak.....	213
VIII.3 Zabaldutako ikerlerroak.....	215
VIII.3.1 Zabaldutako ikerlerro nagusiak.....	215
VIII.3.2 HITES, IRAKAZI eta IDAZKIDE sistemetan zein tresna linguistikoetan landu beharrekoak.....	217

BIBLIOGRAFIA	219
IRUDIEN ZERRENDA	225
TAULEN ZERRENDA	229
ERANSKINAK	231
A. Azaleko maila hutsezko desbideratzeen sailkapena.....	231
B. Desbideratze linguistikoen sailkapena.....	237
B.1 Hitz-mailako desbideratze linguistikoen sailkapena.	239
B.2 Esaldi-mailako desbideratze linguistikoen sailkapena.	242
C. Desbideratzeen sakoneko arrazoen sailkapena.	245
D. Landa-lanean ikasleei jarritako probetako bat.....	251
D.1 Irakasleen hipotesien zerrenda.....	253
D.2 Hipotesi-zerrenda lantzeko ikasleei egindako azterketa.....	254
D.3 Suertatu diren eta ez diren hipotesiak azterketen emaitzetan.	256

I. Motibazioa eta aurkezpena.

I.1 Motibazioa eta helburuak.

Adimen Artifizialaren hainbat eremutan (sistema adituak, sistema tutoreak, ikaslearen modelizazioa...) emandako aurrerapenak, Lengoaia Naturalaren Prozesamenduan (LNP) garatutako teknikak eta psikolinguistikan burutu diren aurrerapausoak elkartu eta koordinatu beharrekoak dira OLHIA eremuan. Orain dela bost urtetik hona hainbat korrontek (Chanier, 94) Ordenagailuz Lagundutako Hizkuntzen Ikaskuntza Adimenduna (OLHIA) arloan hurbilpen diziplinartekoa aldarrikatu du.

UPV/EHUko Lengoaia eta Sistema Informatikoak sailean azken hamabi urteetan lanean aritu den IXA ikertaldeak LNP arloan euskararen tratamendu automatikorako hainbat tresna lortu du, besteak beste, Urkiaren (1997) eta Alegriaren (1995) tesietan azaltzen den euskararako analizatzaile morfologikoa. Tresna hauen berrerabilgarritasuna posible zela ikusita, tresna horien egokitzapena hizkuntzen ikaskuntza arloan landu dugu. Honetaz gain, aipatu sailean sistema tutoreak landu dituen taldea ere badugu. Gai honen inguruan hainbat tresna sortu izan da azkenaldian burututako tesiak (Arruarte, 98; Elorriaga, 98) direla eta.

Hasiera batean Sistema Tutore Adimendunen helburua irakaslearen papera betetzea zen. Hizkuntza bat irakasteko sistemen artean lehenengoa Weischedel-ek (1978) eraikitako German Tutor izenekoa dugu. Honek alemaneraz idatzitako testua aurkeztzen zion ikasleari hainbat galderarekin batera. Ikasleak galdera hauei erantzun ondoren sistemak erantzunei buruzko laguntza ematen zion ikasleari. Ikasleak gaizki idatzitakoa detektatzeko gai ere bazen sistema hau. Ordutik hona tresna asko garatu dira, batez ere gramatikan oinarritutakoak, besteak beste (Pulman, 84; Barchan et al. 86; Schuster, 86; Menzel, 88; Schwind, 90; Yazdani, 90).

Aipatutako testuingurua oinarri izanik, tesi honen hasierako motibazioa bigarren hizkuntza ikasten ari den ikasle batentzako sistema tutorea garatzea izan zen, hau da **irakaste**-sistema bat. Geroago, ordea, ikaslearen ikaste-prozesua aztertzeke ingurune baten diseinua eta garapena izan genuen helburu, hau da **ikaste**-sistema bat. Abiapuntua irakaslearen ikuspuntutik ikaslearen ikuspuntura aldatu zen.

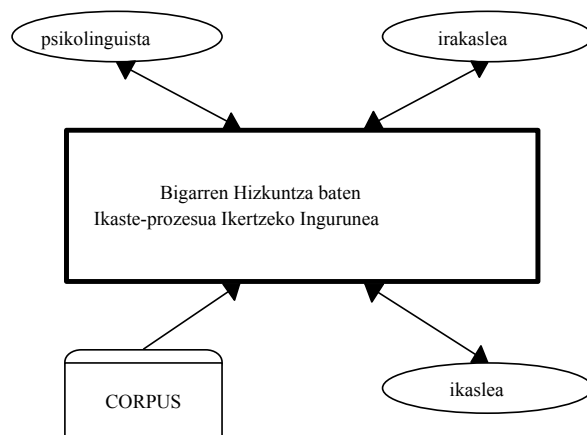
Hortaz, ikaslearen ikaste-prozesua aztertzeke ingurune bat garatzea erabaki genuen, eta hau izango da ondoko kapituluetan zehar azalduko duguna. Konkretuki, bigarren

hizkuntzako ikasleen hizkuntz ezagutzaren eskuratze-prozesua ikertzeko, MUGARRI izeneko sistema anitzeko ingurunea garatu dugu.

Aurkeztuko dugun lanaren helburua bikoitza izango da: teorikoa eta praktikoa.

- Helburu teorikoa honela definitzen dugu:
Psikolinguista baten laguntzaz, eta ikasle batek bigarren hizkuntzan idatzitako testuak oinarriztat hartuz, ikaste-prozesuan zehar ikasleak berak bereganatzen doan ezagutza errepresentatzea. Ezagutza honi tartehizkuntza deituko diogu.
- Helburu praktikoa bi azpichelburutan banatzen dugu:
 - Irakasleari, bere ikasleen ikaste-prozesuaren historia jarraitzen, desbideratzeak diagnostikatzen eta zuzentzen laguntzea.
 - Ikasleari, bigarren hizkuntzari buruzko ezagutza berria eskuratzen laguntzea.

Ondorioz, lortu nahi dugun ingurunean hiru giza agente egongo dira elkarrekintzan sistemarekin: psikolinguista, irakaslea eta ikaslea (ikus I.1 irudia).



I.1 irudia. MUGARRI ingurunearen giza agenteak.

Psikolinguistak tartehizkuntzaren errepresentazioan aditua denak, sistemari lagunduko dio tartehizkuntza adieraziko duen ezagutza-basearen eraketan. Corpus bat aztertuz sortuko da tartehizkuntzaren errepresentazioa, bai ikasle konkretu batena bai ikaskuntza-maila berean dagoen ikasle-multzo osoarena izango dena. Ikaskuntza-maila euskaltegietan antolatzen diren taldeen arabera definituko da. Batzuetan ikasle desberdinek ikaskuntza-maila berean egon arren, jakintza-maila desberdina izaten dute, hizkuntzari dagokionez. Gure sisteman, ikaskuntza-maila berean daudenak ikaskuntza-maila bera dutela esango dugu, nahiz eta jakin batzuetan hizkuntz maila desberdina izaten dutela. Geroago sistemak ikasleekin izango duen elkarrekintzak lagunduko digu perfila edo hizkuntz maila hori

pertsonalizatzen. Hortaz, hemendik aurrera, ikasle-talde baten ezagutza lantzean ikaskuntza-mailaren ezagutzari buruz arituko gara.

Irakasleak, ikasleak ezagutzen dituenak, ikasle bakoitzari buruzko informazioa emango dio sistemari: ikaste-prozesuan eragina izan dezaketen ezaugarriak (esaterako, ikaskuntza-maila konkretu bat lortzeko ikasleak behar izan duen urte-kopurua, ama-hizkuntza, ...), ariketak egitean ikasleak erabiltzen dituen ikaste-estrategiak, etab. Dena den, sistema lagungarri ere gertatuko zaio irakasleari aldi berean, ikasleak erabiltzen dituen egitura linguistikoen diagnostikoa egiteko, – psikolinguistaren laguntzaz eratutako tartehizkuntza ezagutza-basea oinarri hartuko baitu –.

Ikasleak izango duen ikuspuntua bestelakoa izango da. Ikasleak idazte-prozesuan lagungarri gertatzen zaion sistema ikusiko du bere aurrean. Sistema ikaslearen ikaste-prozesuaren osagarria izango da. Ikasleak idazten dituen testuak lantzeko aukera izango du; adibidez, zuzentasuna landu, hitz konkretuei buruzko informazioa bilatu hiztegian, hitz baten deklinabide kasu guztiak, analisiak, ... Eta, etorkizunean modu erraz batean ariketak eransteko aukera izango du sistemak. Sistema, bere aldetik, ikaslearekiko elkarrekintza behatuz tartehizkuntzaren errepresentazioa fintzeko informazioa jasoko du.

Hiru ikuspuntu horiek abiapuntu bezala hartuz, hiru sistema desberdin definituko ditugu MUGARRIn:

- **HITES:** tartehizkuntz egiturak modelizatzeko sistema.
- **IRAKAZI:** irakasleari orientatutako sistema, ikaslearen ikaste-prozesuan ematen diren aurrerapenak aztertzen lagunduko diona.
- **IDAZKIDE:** bigarren hizkuntza ikasteko OLHIA ingurunea.

Esan bezala, diseinatu dugun ingurunearen helburu nagusia hau da: ikaslearen ikaste-prozesuari buruzko informazioa jasotzea. Konkretuki tartehizkuntzaren eraketa dugu aztergai, ikaslearen konpetentzia gramatikala, hain zuzen. Sistema informatikoa irekia izateaz gain, independentea izango da ikasten den hizkuntzarekiko. Horretaz gain, sistema hedagarria izango dugu, hau da, diseinuan aurreikusi dugu etorkizunean aukera izango dugula ikaslearen konpetentzia komunikatiboa ere aztertzeko. Honetarako, bai multimedia arloan bai ahots tratamendurako ikerkuntzan garatzen ari diren tresna informatiko berriak ingurunean gehitzea interesgarritzat jotzen dugu.

Esan bezala, tesi honetan bigarren hizkuntza ikasteko garatuko dugun MUGARRI izeneko ingurunearen helburu nagusia tartehizkuntzaren eraketa-prozesuaren azterketa da. Etorkizunean, tesi honetan garatu diren ondoko gaietan sakontzea interesgarri deritzogu:

- Ikaskuntza-maila bateko ikasleek duten hizkuntz jarioetasuna.

- Bigarren hizkuntza ikasten duenarentzako laguntza.

Badira ere, tesi honetan landu ez diren arloak non guk garatutako tresnak erabil baitaitezke:

- Hizkuntz desberdinak ikasten duenaren aberastasuna hizkuntza bakoitzean.
- Hizkuntza bat ikasteko programagintzaren garapena ikaskuntza-mailen zehar.

Dena den, erabilpenei buruz VII. kapituluan arituko gara zabalago. Kapitulu honetan, berriz, jarraitutako diziplinarteko metodologia eta MUGARRI ingurunearen aurkezpena egingo dugu: ingurunearen mailak eta hauen arteko komunikazioa azalduko ditugu ingurunearen ikuspegi orokorra emateko asmoz. Tesian zehar, berriz, maila horietako osagaiei buruzko azalpenak emango ditugu beste hainbat kapitulutan.

I.2 Metodologia: diziplinartekoa eta definizio-diseinu- implementazio-ebaluazio zikloan oinarritua.

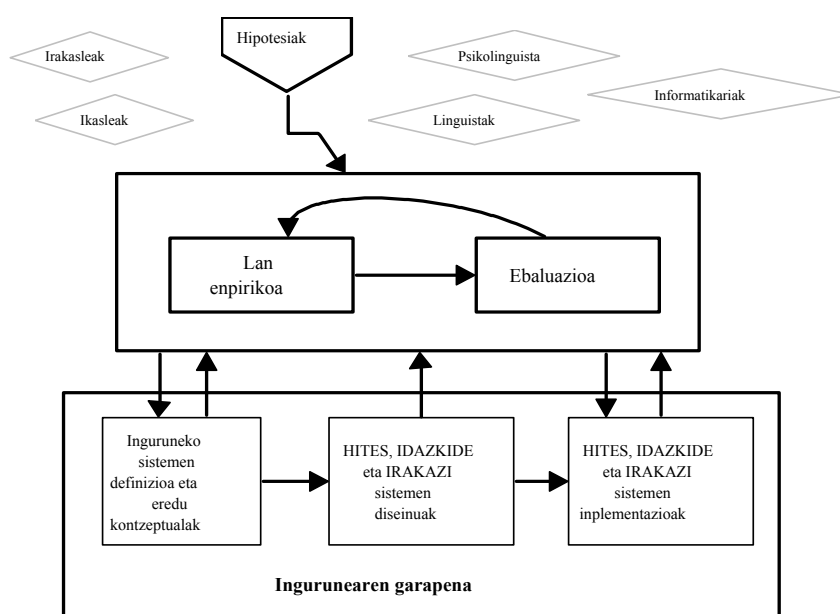
Aurreko atalean azaldu dugunez, lau agente nagusi agertzen zaizkigu MUGARRIn. Alde batetik, ingurune informatikoa bera, eta bestetik hiru giza agente: psikolinguista, irakaslea eta ikaslea. Egoera honek, nahitaez, agente guztien partehartzea eskatzen du ingurunearen definizioan. Honetaz gain, esan dugun bezala, adimen artifizialean, lengoia naturalaren prozesamenduan eta psikolinguistikan emandako aurrerapausoak elkartu eta koordinatu behar dira OLHIA eremuan. Hau dela eta, aurkezten dugun tesian bi ikuspegi landu dugu batez ere. Alde batetik, Adimen Artifizialean (AA) ikertutako hainbat teknika (hala nola, sistema tutoreak, eta ikaslearen modelizazioa) eta lengoia naturalaren prozesamenduko zenbait tresna; eta, beste aldetik, psikolinguistikan landutako tartehizkuntza kontzeptua (Selinker, 72), eta honen ikerketarako sistema erdiautomatiko baten diseinua, aipatu arloan garatutako teknikak integratuz.

Hau guztia aurrera eramateko ezinbestekoa dugu aipatutako agenteen arteko elkar ulertzea lantzea. Hori dela eta, lan honen hasieran, aurrez aurre aurkezten zitzaizkigun baliabideak aztertu genituen. Lengoia naturalaren prozesamenduari zegokionez dagoeneko esperientzia genuen informatikariok eta linguistak elkarlanean; psikolinguistika arloan, berriz, aurretik ez genuen proiektu diziplinartekorik izan, informatikariok eta psikolinguistek zein hizkuntz irakasleek elkarrekin lan eginez. Horregatik, hasiera batean hainbat saiakera egin genuen ILAZKI euskaltegiko irakasle batzuekin elkarriketak antolatuz. Elkarriketa haietatik ateratako ondorioak zirela medio sistemaren definizioari ekin genion, informatikariok, psikolinguistak, linguistak eta hizkuntz irakasleak elkarlanean aritzeko metodologia zehaztuz.

Behin elkar ulertzearen mugak gaindituak izanda, ingurunearen garapenerako lan-metodoa finkatu genuen. I.2 irudian agertzen denez, ezinbestekotzat jo dugu ingurunearen garapenean zehar lan enpirikoak burutzea. Lan enpiriko hauen helburuak bi motatakoak izan dira: batzuek ingurunearen definizioaren eta diseinuaren zehaztapenean lagundu dute, eta, besteek eramandako bidea birbideratu dute askotan ebaluazio partzialak eginez.

Hortaz, laburbilduz, ebaluazioa MUGARRI ingurunearen garapenearen eragilea izan da, eta ebaluazioan parte hartu duten agenteak jakintza-arlo desberdinekoak izan dira.

Ondoko irudian ingurunearen garapenearen eskema azaltzen da. Bertan islatzen da ebaluazioak prozesuan izandako garrantzia.



I.2 irudia. Ingurunearen garapenearen eskema: ebaluazioa hobekuntzaren eragile gisa.

I.2 irudian ikusten denez MUGARRIn sortu diren sistemen garapena prozesu zikliko gisa planteatu da, non ebaluazioa ziklo bakoitzeko fase bat den; hau da, prozesu lineal honetan ebaluazioa ez dugu uzten bukaerarako. Lan-metodo honetan hiru ziklo nagusi agertzen dira:

- Sistemen definizioa.
- Sistemen diseinua.
- Sistemen inplementazioa.

Sistemen definizioa lehenengo zikloa dugu. Lehen ziklo honen helburu nagusia beharren identifikazioa izan zen, konkretuki, bigarren hizkuntzaren ikaskuntza arloan zeuden beharren artean zeintzuk landu zitezkeen informatikaren teknikak erabiliz. Fase honetan irudian agertzen diren agente guztiek hartu zuten parte linguistek izan ezik, hau da, irakasleek, ikasleek, psikolinguistak eta informatikariek. Beharren identifikazioa nagusiki 1992/93

ikasturtean burutu zen. Diziplinarreko bileren ondoren¹ identifikatutako beharretatik abiatuta sistemen definizioa eta eredu kontzeptuala zehaztu zen. Horretaz gain irakasleen ustetan ikasleek idazketan izaten dituzten arazoak corpusen azterketaren bidez berretsi genituen.

Sistemen diseinua sistemaren definizioa zein eredu kontzeptualak zehaztu ondoren hasten da garatzen. Bigarren ziklo honen helburu nagusia da, behin beharrak identifikatuta, hauek asetzeko sistema informatikoa diseinatzea. Lan enpirikoa ere egiten da fase honetan. Prozesu honetan linguista konputazionalak ere hartu zuten parte, batez ere, egitura linguistikoen tratamendu informatikoaren bideragarritasuna aztertzeari zegokionean. Horretan hainbat muga topatu genuen diseinuan: alde batetik, lengoia naturalaren tratamenduak mugak jartzen zizkigun garapen informatikoan, eta, beste alde batetik, euskara hizkuntzan, normalizazio-prozesuan dagoenez, oraindik ere hainbat egitura zehatz-mehatz definitu gabe daude. Hortaz, askotan, diseinua lantzean, gertatu izan zaigu definizio-fasera itzuli beharra, aipatutako arazo linguistikoengatik, edo informatika-tekniken mugengatik, edo irakasleek identifikatutako beharren zehaztasun faltagatik ere.

Sistemen implementazioaren helburua definizioan zehaztutako beharrei erantzuteko sistema informatikoa gauzatzea da. Honetan bi ebaluazio-mota bereizi ditugu: sistemen funtzionamendua eta erabiltzailearen asetzeta. Funtzionamenduaren fasean parte hartzen duen agente nagusia informatikaria izaten da; honek emaitzak konparatzen ditu lan enpirikoan jasotako hainbat datuarekin, edo sistemarekin zerikusia duen adituari galdetuz. Asetzeari dagokionez sistema bakoitzaren erabiltzailea izaten da agente nagusia: IRAKAZIren kasuan irakaslea, HITESen psikolinguista eta IDAZKIDEn ikaslea. Ebaluazio-mota hau zikloaren amaieran kokatzen da gehienbat: gure esperientzian oraingoz gutxi landua izan da, iritzi batzuk jaso baditugu ere. Edozein kasutan, erabiltzaileen asetzearen ebaluazioa etorkizuneko lanerako utzi dugu, hain zuzen ere, hiru sistemen arteko integrazioa burutuko denerako. Beraz, komentatuko ditugun emaitzak (§ VII.3) sistemen funtzionamenduari buruzkoak izango dira. Emaitza hauek aipatutako sistemen erabiltzean psikolinguistek edo irakasleek atera ditzaketen ondorio-motak ematen dizkigute. Honek esan nahi du aipatutako atalean azalduko duguna ez dela tesiaren frogapen bat, baizik eta tesi honetan garatu diren sistemak nola erabil litezkeen erakusteko saiakera bat.

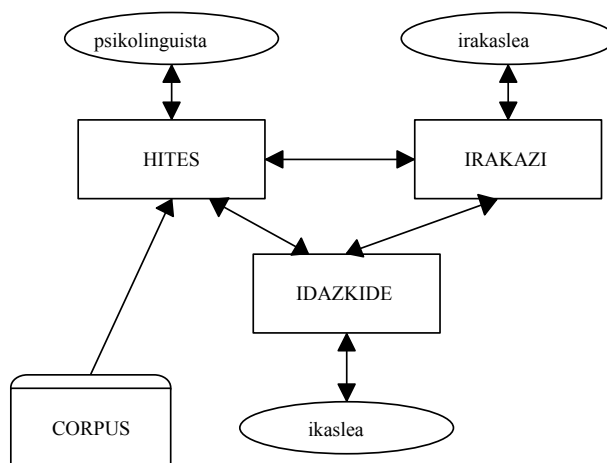
¹Aipatzekoa da irakasleen artean eszeptizismoa izaten zela nagusi informatika-teknikek egin zezaketenaz ikaste-prozesuen azterketan.

I.3 MUGARRI ingurunearen aurkezpena.

Atal honetan tesian proposatzen dugun ingurune informatikoa definitzeaz gain ingurunearen arkitektura azalduko dugu. Besteak beste, ingurunea osatzen duten mailen arteko komunikazioan arreta jarriko dugu.

I.3.1 Ingurunearen definizioa.

Tesi honetan bigarren hizkuntza ikasteko aurkezten dugun ingurunean ondoko agenteak topatzen ditugu: ikaslea, irakaslea, psikolinguista eta HITES, IDAZKIDE eta IRAKAZI sistema informatikoak (ikus I.3 irudia).



I.3 irudia. Hiru sistemen arteko komunikazioa.

Aipatu den bezala, MUGARRIren helburu nagusia ikaste-prozesuko zenbait unetan dagoen hizkuntzaren ezagutza aztertzea da, hain zuzen ere, honako kasu hauetan: ikasle bakoitzarena batetik, eta ikaskuntza-maila berean dagoen ikasle-taldeena bestetik. Helburua noski ez da maila jakin batean izan behar luketen ezagutzaren berri ematea, egiazki duten ezagutza azaltzea baizik. Ezagutza honek euskaltegiek beren programazioetan helburu gisa finkatzen dituzten ezagutzekin ez du zergatik guztiz bat etorri beharrik.

Hurrengo lerrootan aipatu sistema informatikoak banaka-banaka definitzeari ekingo diogu.

HITES bigarren hizkuntza ikasten ari diren ikasleek duten tartehizkuntzaren kompetentzia gramatikalaren modelizaziorako sistema adimenduna da. Sistemak ikasle jakin baten tartehizkuntza aztertzeaz gain, ikaskuntza-maila berean dagoen ikasle-talde baten tartehizkuntza ere aztertzen du. Esan liteke beraz, HITESen saio bat hauxe dela:

- ikasle bakar bati dagozkion testu-multzoen azterketa, helburua ikasle jakinaren tartehizkuntz eredu eguneratzea izanik.

edo/eta

- ikaskuntza-maila berean dagoen zenbait ikasleren testu-multzoa aztertzea ikaskuntza-maila jakinari dagokion tartehizkuntz eredu eguneratzeko asmotan.

IRAKAZI ikaslearen ikaskuntzan gertatzen diren prozesuen diagnostikoa egiteko irakasleari lagungarri zaion sistema bat da. Horretarako HITESek sortutako tartehizkuntz ereduak erabiltzen ditu.

IDAZKIDE idazte-prozesuan zehar ikasleari laguntza ematen dion laguntzaile elkarreragilea da, eta HITESek sortutako ikaskuntza-maila baterako tartehizkuntz ereduetan nahiz IRAKAZI emandako informazioan oinarriturik dago.

I.3.2 Ingurunearen arkitektura.

MUGARRIren osagaiak hiru maila nagusitan bildu daitezke: sistemak, tresna linguistikoak eta ezagutza-baseak. Hiru maila hauen arteko komunikazioa zuzenean gertatzen da, konkretuki bi komunikazio-mota definitzen dugularik, hau da, kontsultak eta eguneraketak (ikus I.4 irudia).

Tresna linguistikoek honelako harremana dute beste bi mailekin:

- Ezagutza-baseen artean ezagutza linguistikoari bakarrik egiten dizkio kontsultak.
- Sistemen zerbitzurako daude diseinatuak.

I.4 irudiak dagerzkigun tresna linguistikoak moduluan, besteak beste, honako hauek aurkitzen ditugu: analizatzaile morfosintaktikoa, proposatzaile morfologikoa, lexikoaren informazioari buruzko kontsulta-interfazea, lexikoaren definizioei buruzko kontsulta-interfazea, hitz-mailan akatsen detektatzaile/zuzentzailea eta deklinabideen sortzailea.

Sistema bakoitzean tresna linguistikoetatik behar den erabilpen propioa definitzen da. Adibide gisa, analizatzaile morfosintaktikoari dagokionez, IDAZKIDEk hitz baten analisi batetik unitate esanguratsuak detektatzea nahi izaten du ikasleari laguntza eskaintzeko asmoz; HITESek, berriz, analisisan aplikatutako erregela linguistikoei buruzko informazioa behar izaten du tartehizkuntz egiturak osatzeko.

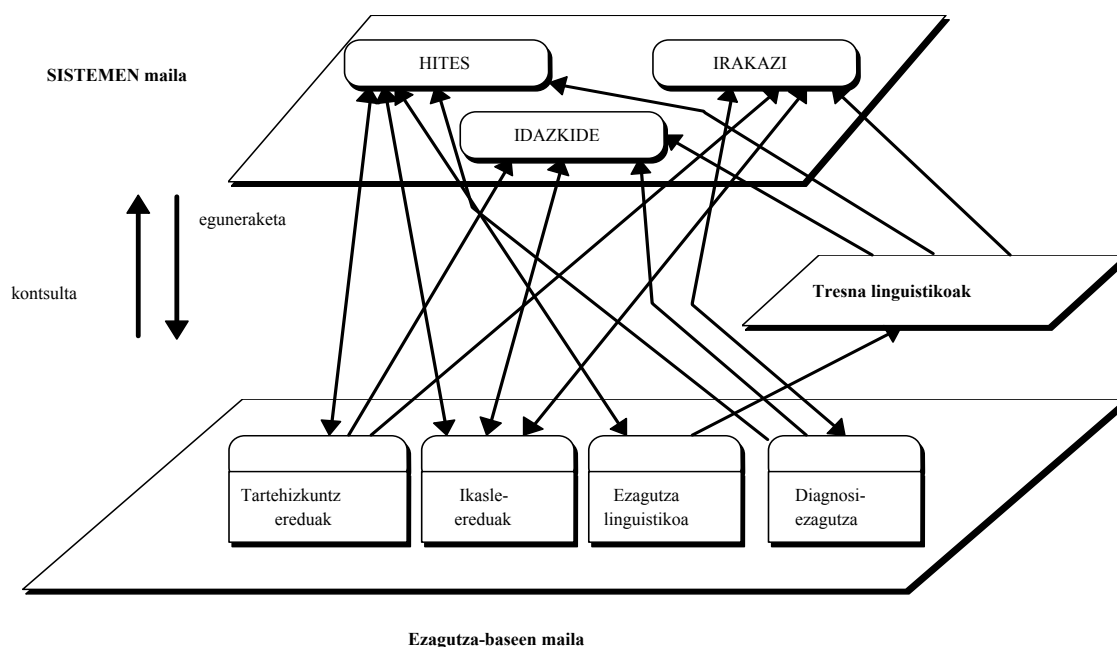
Sistemek beste bi mailekiko komunikazioa honela definitzen dute:

- Ezagutza-baseei dagokienez bai kontsultak bai eguneraketak egiten dituzte sistemek, kasu guztietan izan ez arren (ikus I.4 irudia).
- Tresna linguistikoei deia egiten diete ezagutza linguistikoari buruzko kontsultak egin behar dituztenean.

Ezagutza-baseekiko eguneraketak direla eta, ikaslearen ikaste-prozesua aztertzeko informazio zehatzagoa lortzen da. Bada, eguneraketa hauek direla medio, hiru sistemen arteko komunikazioa lortzen da zeharka. Adibidez, HITES sistemak tartehizkuntz ereduak eguneratzen dituen momentutik IDAZKIDEk ikasleari emango dion laguntza hobea izango da.

Ezagutza-baseak agente pasiboak dira beste bi mailekin konparatuz, hau da, inoiz ez dute hasten komunikazioa:

- Sistemek egiten dizkiete kontsultak edo eguneraketak.
- Tresna linguistikoek ezagutza linguistikoaren kontsulta egiten dute.



I.4 irudia. MUGARRI ingurunearen arkitektura.

Inguruneko hiru maila nagusiak aipatu ondoren sistemen maila eta ezagutza-baseen mailak zabalago deskribatuko ditugu:

Sistemen maila

HITES, IDAZKIDE, eta IRAKAZIk alde batetik, elkarrekiko zuzeneko komunikazioa dute eta, beste aldetik, zeharkako komunikazioa burutzen dute komunean dituzten ezagutza-baseak direla medio. Hiru sistemek konpartitzen dituzten ezagutza-baseak ondokoak dira: ikaskuntza-maila bakoitzari dagokion ezagutza linguistikoa, tartehizkuntz ereduak, diagnosi-ezagutza, eta ikasle bakoitzari dagokion ikasle-eredua. I.3 irudian, dagoeneko sistemen arteko zuzeneko komunikazioa errepresentatu dugunez, I.4 irudian zeharkako komunikazioa

irudikatu dugu gezien bidez. Bertan hiru sistemek beren ataza egoki burutzeko erabiltzen dituzten tresna linguistikoak ere agerian jartzen zaizkigu. Aipatu tresnak sistemen maila eta ezagutza-baseen mailaren artean irudikatu ditugu.

Ezagutza-baseen maila

I.4 irudian agertzen denez, sistemen maila eta ezagutza-baseen arteko komunikazioa dago. Ezagutza-baseen kontsulta zein eguneraketaz arduratuko litzatekeen modulua ere beharrezkoa litzateke ezagutza-baseen mailan, honi buruz ordea ez gara arituko.

Esan bezala, arkitekturaren maila honetan lau ezagutza-base definitu ditugu: ezagutza linguistikoa, tartehizkuntz ereduak, diagnosi-ezagutza eta ikasle-ereduak. Horiei buruzko zenbait zehaztasun emango dugu jarraian.

Ezagutza linguistikoa ikaskuntza-mailaren arabera antolaturik dago. Ezagutza kolektiboaren zeharkako errepresentazioa da. Zeharkako errepresentazioa dela esaten dugu tartehizkuntz egiturarik ez duen arren, – tartehizkuntz ereduak egiten baitute hau –, tresna linguistiko-konputazionalentzat beharrezkoa den ezagutza linguistikoa definitzen duelako. Ezagutza linguistikoaren errepresentazioa aipatutako tresnen diseinu eta inplementazioan erabilitako formalismoaren menpe dago.

Tartehizkuntz ereduak ikaskuntza-mailaren baitan eratzen dira eta maila bakoitzean dauden ikasleek duten ezagutza komuna adierazten dute. Ereduak ez dira ikaskuntza-maila horietan ikasleek lortu beharko luketen ezagutza, haatik ikasle jakinek duten ezagutzaren isla baizik. Tartehizkuntz ereduak batik bat HITESek sortuak dira, nahiz eta ikaslea IDAZKIDEREkin elkarrekintzan denean emandako informazioa ere eredu hauen osaketarako aski lagungarri suertatzen den.

Diagnosi-ezagutza moduluak ikaskuntza-maila baten ikasleek bere tartehizkuntz komunaren egitura linguistiko desbideratuak erabiltzeko eduki ditzaketen arrazoiak gordetzen ditu. Ezagutza hau HITESek erabiltzen du IRAKAZIk eskaintzen duen ezagutza dela medio. IDAZKIDE ikaste-sistema lagungarriak ere diagnosi-ezagutza erabiltzen du ikaslearekiko elkarrekintzan.

Ikaslearen ereduak ikaslearen beraren tartehizkuntz ereduaren zein bere ikaskuntzaren ezaugarrien berri ematen duen eredu generiko bat genuke. Ezaugarri hauek (hizkuntz estrategiak, ikaskuntza-historia, beste hizkuntza batzuen ezagutza, etab.) irakasleak (IRAKAZI) zein ikasleak (IDAZKIDE) emandakoak dira. Eredu hau, alabaina, IDAZKIDEn testu-idazketaren laguntza-prozesuan erabiltzen da gidaritzarako.

I.3.3 Hiru sistemen arteko komunikazioa.

HITES, IRAKAZI eta IDAZKIDE elkarren osagarri diren helburuak dituzten hiru sistema dira, eta zenbait ezagutza-base ere komunak dituzte. Hala nola ezagutza-base hauek direla medio lortzen da zeharkako komunikazioa, I.4 irudian adierazi dugun bezala. Zeharkako komunikazioaz gain, zuzeneko komunikazioa ere topatzen dugu HITES, IDAZKIDE zein IRAKAZIren artean.

Hortaz, aipatu bi komunikazio-motei buruz arituko gara ondoko lerrootan, hauek dira:

- Zeharkako komunikazioa: ezagutza-baseen eguneraketan bitartez.
- Zuzeneko komunikazioa: sistemek elkarri bidaltzen dizkieten mezuen bidez.

Zeharkako komunikazioa

HITES, IDAZKIDE eta IRAKAZI elkarren osagarri diren sistemak dira. Hirurek ezagutza-baseak konpartitu arren, sistematarik bakoitzak badu bestelako baliabide-motarik ere. Hau dela eta, bakoitzaren baliabide nagusienak aipatutako ditugu labur, eta, ondoren, komunean dituzten baliabideak direla medio, zeharkako komunikazioaz arituko gara.

HITESek tartehizkuntz ereduaren eratze-prozesuan psikolinguistak hartutako erabakiak erabiltzen ditu. HITESek testuetatik abiatuz fenomeno linguistikoak eta beren testuinguruaren erabilpena ondorioztatzen du, eta psikolinguistaren laguntzaz tartehizkuntzaren egiturak eratu eta osatu ere egiten ditu.

IDAZKIDEk hara non hartzen duen kontuan ikasleak bere idazte-prozesuan hartzen dituen aukerak: ikaslearen jarduna jasotzen du eta ariketez gain iradokizunak, zuzenketak eta ezagutzaren kontsulta-tresnak erabiltzeko proposamenak baditu.

IRAKAZIren kasuan, irakasleak sistemari ikasleek erabilitako egituren diagnostikoari buruzko datuak eskaintzen dizkio.

Sistema bakoitzaren baliabide propioak aipatu eta gero hiru sistemek konpartitzen dituzten ezagutza-baseak komentatuko ditugu zeharkako komunikazioari dagokionez.

Alde batetik, IRAKAZIk diagnosi-ezagutzan eguneratutako informazioa bai HITESek bai IDAZKIDEk erabiliko dute. Beste alde batetik, HITES dela medio, IRAKAZIk berak zenbait datu eskuratzen du, konkretuki psikolinguistaren laguntzaz tartehizkuntzak izeneko ezagutza-basean eguneratutako tartehizkuntz egiturak jasotzen ditu. Azkenik, IDAZKIDE bere aldetik bai tartehizkuntzak bai ezagutza linguistikoaren ezagutza-baseetatik ikaskuntza-mailari buruzko informazioa lortzen du ikasleekiko elkarrekintza gidatzeko; aipatutako bi ezagutza-baseen eguneraketaz HITES arduratzen delarik.

Zuzeneko komunikazioa

Bai IDAZKIDEk bai IRAKAZIk egitura linguistiko interesgarriak detektatzen dituztenean HITESi bidaltzen dizkiote, honek azter ditzan ea tartehizkuntzaren parte diren. Kontuan izan behar dugu honelako egiturak direla corpus bidez detektatzerik izan ez dutenak. Modu berean momentu konkretu batean HITESek berak egitura ez osatuak identifikatzean bai IDAZKIDERI bai IRAKAZIRI eska diezaioke laguntza egitura horiek osatzeko. Honelako informazio-trukaketa ezagutza-baseen bidez ez egitea proposatzen dugu, honelako komunikazio-motak sistemen arteko mezu-trukaketa eskatzen baitu. Honi buruz maila teorikoan arituko gara.

Hiru sistemen arteko komunikazio-helburuak teorikoki planteatzen ditugun arren diseinu pausu batzuk emanak dira V. eta VI. kapituluetan ikusiko dugunez. Dena den, garapen erreal batean komunikazio-protokoloen azterketa zehatza beharko genuke, eta hau, VIII. kapituluan komentatuko dugun bezala, etorkizuneko ikerlerro nagusi gisa proposatzen dugu. Hortaz, hiru sistemen arteko komunikazio zuzenaren helburuak definituko ditugu.

Kasu guztietan bi zentzuko komunikazioa egon daiteke, hau dela eta kasu bakoitzeko bina helburu nagusi aipatuko ditugu.

HITES eta IDAZKIDERen arteko komunikazioa:

- HITESek IDAZKIDERI informazioa eskatzean helburua ondokoa da:

Ikasleari idazketan aholku ematerakoan IDAZKIDERen erabaki-hartzeak hobetzea.

IDAZKIDEk ikaslearekiko elkarrekintzan egitura linguistikoak lantzeko ariketak proposatuko dizkio HITESek eskatuta, azken honek corpusetan detektatutako zenbait egituraren ulermena lortzeko, hau da, tartehizkuntz egiturak osatu ahal izateko.

- IDAZKIDEk HITESi informazioaren berri ematean helburua ondokoa da:

HITESen tartehizkuntz ereduaren eraketa osatzea.

HITESek IDAZKIDEtik jakingo du ikasleak bere testuen zuzenketan zein aukera egin duen, eta honek bide emango digu aztertzen ari garen ikaslearen ikaskuntza-mailaren tartehizkuntz ereduaren IDAZKIDEk detektatu dituen egitura linguistiko berri horiek gehitzeko.

HITES eta IRAKAZIren arteko komunikazioa:

- HITESek IRAKAZIRI informazioa eskatzean helburua ondokoa da:

HITESen tartehizkuntz ereduaren eraketa osatzea.

HITESek IRAKAZItik jakingo du ikasleak ikasgelan erabiltzen dituen egitura linguistikoak corpusean agertzen ez direnak, eta honek bide emango digu aztertzen ari garen ikaslearen ikaskuntza-mailaren tartehizkuntz ereduaren irakaslearen bidez detektatu diren egitura linguistiko berri horiek gehitzeko.

- IRAKAZIK HITESi informazioaren berri ematean helburua ondokoa da:

HITESek tartehizkuntzaren ezagutzari buruzko informazioa xehetasun handiagoz hornitzea.

IRAKAZIK egitura horiek gauzatzen diren testuinguruari buruzko informazioa are xehetasun handiagoz emango dio, gero HITESek tartehizkuntzari buruzko ezagutza egunera dezan.

IDAZKIDE eta IRAKAZIren arteko komunikazioa:

- IRAKAZIK IDAZKIDERI informazioa eskatzean helburua ondokoa da:

IRAKAZIK erabiltzen duen diagnosi-ezagutza fintzea.

IDAZKIDEEK ikaslearekiko elkarrekintzan egitura linguistikoak lantzeko ariketak proposatuko dizkio IRAKAZIK eskatuta, azken honek irakasleak proposatutako zenbait egituraren ulermena lortzeko, hau da, irakaslearen diagnostikoa fintzeko.

- IDAZKIDEEK IRAKAZIRI informazioaren berri ematean helburua ondokoa da:

IRAKAZIK erabiltzen duen diagnosi-ezagutza fintzea.

IDAZKIDEEK bere saioetan ikaslearekin izandako prozesuaren trazek egitura desbideratuen diagnosirako gakoak eman ditzakete. Hortaz, IDAZKIDEEK informazioa hori bidaliko dio IRAKAZIRI honen bidez irakasleak diagnostiko zehatza egin dezan.

I.4 Txostenaren eskema.

Hasierako motibazioa eta tesi honi hasiera batean jarritako helburuak aurreko ataletan azaldu ditugu. Bigarren hizkuntza ikasten ari den ikaslearen tartehizkuntza lantzeko MUGARRI ingurunea izan dugu aztergai. Ingurune honetan hiru sistema garatu ditugu: HITES, IRAKAZI eta IDAZKIDE. Sistema hauen definizioa, diseinua eta inplementaziorako jarraitu dugun metodologia diziplinartekoa izan da, sistemaren garapenean zehar hainbat esperimentu burutu delarik, sisteman sortzen ziren tresnak baliagarriak ziren ala ez ikusteko asmoz.

Lehenbizi hizkuntzen ikaskuntza munduan sortu diren tresna automatiko edo erdiautomatikoaren ikuspegi orokor bat (II. kapitulua) eman ondoren, diseinatutako ingurunea azaltzeari ekingo diogu.

III. kapituluan, hizkuntza baten ikaste-prozesuaren ikerketarako sistema informatiko adimendun baten oinarri psikolinguistikoa aztertuko dugu. Besteak beste, ondoko kontzeptuak landuko ditugu aipatu kapituluan: desbideratzeak, tartehizkuntza eta fosilizazioa. Honetaz gain, jarraitutako diziplinarteko metodologia ere komentatuko dugu.

IV. kapituluan, corpusaren azterketarako garatu ditugun tresna linguistiko-konputazionalak definituko ditugu. Honetaz gain, ikaskuntza-maila bereko ikasleek komunean duten tartehizkuntz ereduen osaketa-prozesurako garatu dugun metodologia azalduko dugu. Eta, gainera, domeinuaren adierazpena (tartehizkuntzaren adierazpena) deskribatuko dugu, hau baita lortu beharreko azken helburua. Hortaz, kapitulu honetan sistema osoaren bukaerako irteera deskribatuko dugu. Honela, emaitzak nolakoa behar lukeen ezagututa, hurrengo bi kapituluetan deskribatuko diren sistemen helburuez jabetzeko moduan izango gara.

V. kapituluan, HITES sistema deskribatuko dugu: bigarren hizkuntza ikasten ari diren ikasleek duten tartehizkuntzaren konpetentzia gramatikalaren modelizaziorako sistema adimenduna da. Tartehizkuntza bigarren hizkuntza ikasten ari den ikasleak bere ikaste-prozesuan zehar eskuratzen duen egitura linguistikoen multzoa izango litzateke. HITESek helburu bikoitza du: batetik, ikaskuntza-maila berean dauden ikasleek jaso duten ezagutza komunaren errepresentazioa lortzea du xede (ezagutza kolektiboa); eta bestetik, ikasle jakin baten tartehizkuntzaren errepresentazioa ere eskuratzea (ezagutza berezitua).

VI. kapituluan, IDAZKIDE eta IRAKAZI sistemak azalduko ditugu. IDAZKIDE idazte-prozesurako ingurune lagungarria da; ikaslearekin egiten du lan eta bai HITESekin bai IRAKAZIekin komunikazioa du. IRAKAZIk ikaslearen ezagutza linguistikoaren diagnosian eskaintzen dio laguntza irakasleari. Edozein kasutan, kapitulu honetan gehienbat IDAZKIDE azalduko dugu garatuagoa baitago eta ekarpen gehiago baititu. IRAKAZIri buruz, ordea, ideia orokor batzuk emango ditugu. Hau dela eta, Ikaste-estrategiei buruzko azterketa, ikasle-ereduaren errepresentazioa, diagnosi-ezagutza eta tresna linguistikoen integrazioaz ere arituko gara IDAZKIDE sistema azaltzean.

VII. kapituluan, tesian zehar bigarren hizkuntzaren ikaste-prozesua aztertzeke aurkeztu dugun ingurunearen ebaluazioari buruz arituko gara. Azpimarratzekoa da ebaluazioa ez dela izan bukaeran egindako azterketa, baizik eta inguruneke sistemen garapena gidatu duen eragilea. Ebaluazioak planifikazioan lagundu du egindakoa hobetuz, produktu hobe ateratzeko helburuarekin. Kapitulu honetan hainbat esperimenteren emaitzak, besteak beste, emango ditugu. Honetaz gain, kapituluaren azken atalean komentatuko ditugun emaitzak sistemen funtzionamenduari buruzkoak izango dira. Emaiza hauek aipatutako sistemen erabilpenean psikolinguistik edo irakasleak atera ditzaketen ondorio-motak emango dizkigute.

Eta, azkenik, VIII. kapituluan, tesi honetatik atera ditugun ondorioak aurkeztuko ditugu. Honekin batera, lanari bukaera emateko asmoz etorkizuneko lanari ekingo diogu.

II. Hizkuntzen ikaste-prozesua lantzeko sistemen azterketa.

Informatika eta Hizkuntzen Ikaskuntza terminoak batera aditzen ditugunean, ikasle bat pantaila baten aurrean ikusten dugu ariketak egiten, edo hizkuntzan trebatzeko jokoren batekin. Aipatutako terminoen azpian dagoen kontzeptzioa, berriz, zabalagoa da. Alde batetik, ikasleek erabil ditzaketen tresna informatikoak ditugu, bai metodo autonomo moduan, bai euskaltegian landutako gaiak gainbegiratzeko asmotan, baita ikasgelan bertan ariketa osagarri gisa; beste alde batetik, irakasleek materiala prestatzeko dituzten aukera informatikoak ere gero eta ugariagoak dira; azkenik, psikolinguistika arloan ikerketak egiteko tresna lagungarriak pixkanaka sortzen hasiak dira. Beraz, ikasle, irakasle eta psikolinguistentzako produktu informatikoak aztergai ditugu hemen. Informatikak hizkuntzen ikaskuntza arloari eskaintzen dizkion posibilitateen ikuspegi orokor eta zabala emango dugu hain zuzen.

Kapitulu honetan, ordenagailuek eskaintzen dituzten aukeren artean ikasleentzat, irakasleentzat edota psikolinguistentzat erabilgarri izan daitezkeenak aipatuko ditugu. Bost multzo nagusi bereiziko ditugu: Ordenagailuz Lagundutako Hizkuntzen Ikaskuntza (OLHI)¹ tradizionala, sareak, multimedia, Ordenagailuz Lagundutako Hizkuntzen Ikaskuntza Adimenduna (OLHIA)², eta Lengoia Naturalaren Prozesamendua (LNP)³ integratu dituzten sistemak. Paraleloki Euskal Herriaren egoera zein den ikusten joango gara. Azkenik, hausnarketa moduan informatikak euskararengan, eta zehazki euskararen ikaskuntzan, eduki dezakeen eraginaren ildotik, ondorioak komentatuko ditugu.

II.1 Sarrera.

Aukera asko dira informatikak hizkuntzen ikaskuntza arloan ematen dizkigunak. Batzuk esplizituki hizkuntza bat ikasteko metodo edota osagarri gisa aurkezten dira, hots, Didaktiker eta HABEren artean ateratako *Ikastxip* eta *Benetan* sistemak, ordenagailuz euskara ikasteko programak, hain zuzen. Beste batzuk ere, nahiz eta hizkuntza bat ikasteko izenburupean ez saldu, erabilgarri gerta daitezke hizkuntza baten ikaste-prozesuan zehar, besteak beste,

¹CALL: Computer Assisted Language Learning.

²ICALL: Intelligent Computer Assisted Language Learning.

³NLP: Natural Language Processing.

ELHUYAREK egindako *Gipuzkoa Ezagutuz CD-ROMa* ikasgelan ariketa osagarriak antolatzeko aukera ematen du. Hau dela eta, ondoren egingo dugun egoera-azterketa ez dugu mugatuko ikasteko programak izenaren azpian agertzen diren eta, izan ere, ikuspegi zabalagoa hartuz, ordenagailuek ikasgelan eta ikasgelatik kanpo opatzen dizkiguten baliabideak aipatuko ditugu.

Hala eta guztiz ere, azterketari ekin baino lehen, material espezifiko eta autore-paketeak kontzeptuen arteko desberdintasuna argi uztea garrantzizkoa zaigu. Material espezifikoetan egitura zein edukiera linguistiko eta pedagogikoa aurrerabakia datorkigu, beste modu batean esanda, programaren barrualdea ez dago aldatzerik inola ere. Aurretik aipatutako Ikastxi eta Benetan honen adibideak dira. Autore-paketeetan, berriz, diseinu-mailan egitura aurre mugatua etorri arren, erabiltzaileak edukiera definitzeko aukera izaten du. Esate baterako, honelako paketeen bidez euskaltegi bateko irakasleek nahi lituzketen ariketak presta zitzaketan, beren uneko beharretara egokituz. Programa hauen bidez ariketa-mota desberdinak diseina daitezke, hots, hutsuneak bete, erantzun-zerrenda batetik zuzena aukeratu, ulermen-ariketak, e.a. Irakasleak ariketak prestatu ondoren ikasleek lan egin dezakete programa hauekin.

Beraz, laburbilduz, ondokoa hartu behar dugu kontuan kapituluarekin jarraitu baino lehen: lehenik, hizkuntza baten ikaste-prozesuan lagungarri izan daitezkeen tresna automatikoak ez dira ikasteko programak izenburupean datozenak bakarrik; eta, bigarrenik, tresnen erabiltzaileek, hau da irakasleek, ikasleek edo psikolinguistek, diskriminatu beharko dute beren material propioa antolatzeko aukera ematen dizkieten sistemak eta aukera hori ematen ez dizkietenak. Bai batzuk bai besteak oso baliagarri dira: aurretik genioen bezala gakoa erabileran dago.

II.2 Azterketa bibliografikoa.

Esanak esan, kapitulu hasieran aipatutako bost multzo nagusien ikuspegi orokorra emango dugu, hauek dira:

- Ordenagailuz Lagundutako Hizkuntzen Ikaskuntza (OLHI) tradizionala.
- Sare bidezko hizkuntz ikaskuntza.
- Multimedia tekniken aplikazioa hizkuntzen ikaskuntzan.
- Ordenagailuz Lagundutako Hizkuntzen Ikaskuntza Adimenduna (OLHIA).
- Lengoaia Naturalaren Prozesamendua (LNP) integratu dituzten sistemak.

II.2.1 Ordenagailuz Lagundutako Hizkuntzen Ikaskuntza tradizionala.

OLHI tradizionala diogunean testu-prozesadoreak, pakete integratuak, CD-ROMak (Read Only Memory) eta hipertestuak uler ditzakegu. Batzuk besteak baino ezagunago egiten zaizkigu, dudarik gabe. Eta, askotan ez dira erlazionatzen hizkuntzen ikaskuntza munduarekin. Hauetaz gain, aipagarriak ere badira OLHI tradizionalan sarreran azaldutako autore-paketeak. Gaur egun, pakete hauek beste arlo batzuetan ere erabiltzen diren arren, hots, multimedia eremuan, beren sorrera Ordenagailuz Lagundutako Ikaskuntzarako hasierako produktuetan gertatu zen.

Testu-prozesadoreak bukaerarik gabeko tresna dugu, batez ere, maila altuko ikasleen artean. Testu-prozesadore arruntenek zuzentzaile ortografikoak integratuta izaten dituzte; askok thesaurusak ere eskaintzen dituzte, hauek dira, sinonimo eta antonimoen hiztegiak; beste batzuek hizkuntz anitzeko hiztegiak gehitzen dituzte; zenbaitek perpausen itzulpengintzarako laguntzak edota testuaren ideiak antolatzeko tresnak erantsita dituzte; eta, kasu batzuetan konkordantzia-bilduma egiteko aukera izaten da (konkordantzia-bilduma: hitz baten agerpen guztien zerrenda bere testuinguruetan, frekuentzien informazioa ere agertzen delarik, besteak beste). Zuzentzaile ortografikoak oso lagungarri gerta dakizkieke ikasleei beti irakasleari galdezka aritu beharrean: Euskal Herrian XUXEN euskararako zuzentzaile ortografikoa (Agirre et al., 92) dugu eskura, testu-prozesadoreetatik kanpo ere lan egiten duena. Sinonimo eta antonimoen thesaurusak idazlanak aberasteko edota testuen ulermenerako oso egokiak izaten dira, adibidez, liburu-hiztegi batean nekagarri izan litezkeen bilaketak oso erosoak izaten dira aipatutako hiztegietan. Azken hauek irakaslearentzat ere erabilgarri izan daitezke ariketak prestatzeko. Euskaraz gaur egun merkatuan ez dugu eskuragarri honelako hiztegirik, ikerkuntzan lantzen ari badira ere. Edozein kasutan merkatuan badago, adibidez, Atzekoz Aurrera izeneko hiztegia, edozein hitz bukaera emanda hitz-zerrenda ematen duena. Esaterako, -tasun eta -keria atzizkiez bukatutako ariketak prestatu behar ditugunerako, edo bertso-ariketak antolatu nahi ditugunerako baliagarria da. Etekin handia ateratzea dago honelako programei, zalantzarik gabe, hori bai, betiere ikaskuntza-jardueretan egokiro integratuz. Euskararengan arlo honetan pixkanaka pauso txiki handiak ematen ari dira.

Pakete integratuetan testu-prozesadoreak, grafikoak, datu-baseak eta komunikaziorako baliabideak, besteak beste, izaten ditugu batera erabilgarri. Honelako paketeak hizkuntzaren arlo konkretuak (adibidez, turismo-bulego bat) eta idazketa-teknikak lantzen dituzten ikastaroetan gehienbat izan daitezke aproposak.

CD-ROMen kasuan tresna erabilgarriak ditugu, irakasleentzat batez ere: ikaskuntza-materiala prestatzeko aukera aparta ditugu hauek. Hiztegiak, entziklopediak, egunkariak eta

multimedia paketeak topatzen ditugu CD-ROM hauetan. Informazio asko batera gordetzen dute eta bilaketa-modu ugari eskaintzen dizkigute. Ikasleek ere erabil ditzakete beren ikaskuntza-aktibitateetan, hala ere, gure ustez, irakasleak aktibitatea prestatua eta gidatua antolatu beharko luke. CD-ROMetan informazio gehiegi atzigarri dagoenez ikasleak galduta aurki lezake bere burua. Ingeleserako existitzen diren CD-ROMak franko dira, eta, honegatik, iturri inportantea dute eskoletan ikaskuntza-materiala prestatzean. Gure egoera bestelakoa bada ere, pixkanaka sortzen ari den materiala gure ikaskuntza-jardueretan kontuan izatea komeniko litzaiguke, hots, Euskal Herriko entziklopedia elektronikoa.

Hipertestuak azkenekoak ditugu aztergai. Hauetan testu-informazioa sortu eta era ez-sekuentzial eta asoziatiboan informazioa atzi dezakegu. Testuan zehar nabigatu; esaldi erdian zaudela hitz bati buruzko informazioa bilatu edo gramatikari buruzko informazioa eskatu eta gero testura itzuli zuzenean zinen tokira; ariketa baten zehar edozein unetan aurretik egindako mota bereko ariketak kontsultatu eta abar luze bat egin daiteke hipertestuekin. Buckinghameko unibertsitatean, adibidez, proiektu bat burutu zen hipertestuen bidez hizkuntza bat ikasteko materiala sortzeko. Helburua irakurketa lantzea zen. Testu berdinak frantsesez, gaztelaraz eta ingelesez prestatu zituzten, hiru maila desberdinetan. Ikasleak maila eta hizkuntza hautatzeko aukera zuen. Irakurketan zehar ikasleak informazio-iturri desberdinak eska zitzakeen: gramatika, semantika eta nahi zuen testu-zatia beste hizkuntzetan nolakoa zen ikusi. Honelako materiala sortzeko ez da informatikari izan behar, baizik eta hizkuntz irakaslea. Ez da inolako ezagutza informatikorik behar, tresna erabiltzeko behar denaz aparte.

II.2.2 Sare bidezko hizkuntz ikaskuntza.

Azpiatal honetan, oso laburki, aipatuko dugu gaur egungo sareek hizkuntzen ikaskuntza arloan izan dezaketen eragina. Komunikazio-sareak eta erabiltzaile anitzeko sistemak gero eta ezagunago zaizkigu. Komunikazio-sareen bidez ordenagailu desberdinak elkar konektatzen dira. Honelako sistemen bidez hitz idatziaz gain, ahotsa, bideoa... gero eta errazago bidal daitezke munduan zehar.

Munduan zehar badira milaka proiektu arlo honetan, batzuk hizkuntzen ikaskuntzaren inguruan garatuak. Esate baterako, ELNET(European Learning NETwork) proiektuan, Frantzia, Alemania eta Britainia Handiaren artean bosna eskolak hartu zuten parte. Aktibitateen artean honelakoak zeuden: negozioien informazio-bulegoa simulatzen zuen konferentzia, gai konkretuei buruzko eztabaidak, e.a. Proiektuaren helburua zen negozioa eta hizkuntza domeinuetan denbora, distantzia eta kultura desberdinez osatutako taldeen ikaste-estrategiak bilakatzea. Beauvois-ek (1997) ere sare lokal batean integratutako DIWE izeneko sistema garatu du. Honek denbora errealean egiten du lan, eta ikasle-talde bat zein irakaslea

jartzen ditu elkarrekintzan. Hala ere, sare bidez komunikatzen diren sistemetan arazo zailena laguntza berezituarena izaten da, hau da, nola integratzea ikasle-ereduak sarean (Weber et al., 97).

Kontuan izanik munduan zehar dauden ingeles hiztunak eta ingelesa ikasteko eskolak, erraz uler daiteke komunikazio-sistema hauek euskararen ikaskuntzan egon daitekeen erabiltzaile-kopuruarekin ez dagoela konparatzerik. Euskara ikasteko proiektuen artean sare bidez antolatutako lehen saiakera egin da Hezinet izeneko proiektuan (Perez et al., 99): proiektu honek EGA azterketako prestakuntza eskaintzen du. Honetaz gain, euskara ikasteko ikastaroak bereziki ez izan arren, Euskal Herriko Gipuzkoako Ikastolen Elkarte (Gereka, 98) ere sare bidezko ikastaroak antolatzen hasia da.

II.2.3 Multimedia tekniken aplikazioa hizkuntzen ikaskuntzan.

Multimedia-produktuak modu desberdinetan informazioa eta ideiak antolatzeko eta erabiltzeko tresna informatikoak dira. Multimedia aplikazioetan bideoa, argazkiak, musika, soinuak, animazioa, testu idatziak, grafikoak eta abar luze bat batera integratuta topatzen ditugu. Informazioa era ez-sekuentzial eta asoziatiboan atzitzeko aukera dugunean hipermedia terminoa erabiltzen da.

Multimedia aukerak ematen ditu bai era autonomoan bai ikasgelaren barruan aplikazioak erabiltzeko. Aurreko azpiataletan aipatutako aplikazioekin konparatuz, oraingo hauek errazago bete dezakete ikaskuntza-helburu bat beren aldetik, bestelako ikaskuntza-jardueretan integratu gabe. Ordenagailu aurrean jarri, eta, ikasleak turismo-bulego batean, esaterako, murgilduta aurki dezake bere burua. Mota honetako aplikazioa dugu FLUENT, Schoelles et al.ek (1997) aurkezten digutena. Sukalde baten animazioa aurkezten zaio ikasleari, eta honi sistemak hainbat agindu bidaltzen dio. Sistema eta ikaslearen arteko elkarrizketa analizatzeko lengoia naturalaren teknikak erabiltzen ditu sistemak.

Bideoan elkarrizketa bat ikusi, eta nahi izanez gero, pertsonaiek diotena idatzita azaltzea eska dezake ikasleak multimedia sistema batzuetan; edozein momentutan elkarrizketa gelditu eta uneko esaldiari buruzko informazio gramatikala bila dezake; edo, elkarrizketa gelditu, eta atzera egin aurreko esaldia ez delako ondo ulertu; istorioaren pertsonaia baten papera ikasleak berak betetzea ere badu, eta ordenagailuak, irakaslea balitz, ikasleari, honek egin dituen partehartzei buruzko oharra egingo dizkio. Ikasle askorentzat oso motibagarriak dira multimedia aplikazioak, batez ere, ikasgelan antzerki bidez egiten diren egoera-simulazioak gustatzen ez zaizkien horientzat.

Euskara arloan hainbat saiakera egon dira dagoeneko multimedia arloan. Lehenik, EGA azterketa prestatzeko ariketak lantzen dituen aplikazioa garatu da (Perez et al., 99). Honetan, hainbat ariketa-mota, pelikulak, eta abar luze bat eskaintzen zaio ikasleari. Bigarrenik, Bai & by izeneko sistema multimedia garatu da. Honek bi arlo nagusi jorratzen ditu: autoebaluaketarako tresnak eskaintzen dizkio ikasleari (ikaste-erritmoa, aprobeixamendua, ...) eta ikasten irakasten dio. Honetarako estrategia bikoitza erabiltzen du, alde batetik, ikasketa egituratua eskaintzen zaio ikasleari, eta, beste aldetik, aholkularitza didaktikoa ematen zaio irakaslearen bidez, honek DSIS izeneko sistema informatikoa erabiltzen duelarik. Hirugarrenik, sistema aipagarria CLIC (Busquets, 95) izeneko multimedia aplikazioa dugu. Honetan hezkuntzaren arlo desberdinetarako hainbat aktibitate-mota garatu da (gurutzegramak, esplorazio-aktibitateak, animazioak, bideo-sekuentziak, etab.); besteak beste, euskara hizkuntza ikasteko aktibitateak ere integratu dira. Eta, azkenik, aipatzekoa da Zuberoako Kantuen Bilduma. Hau ikasteko sistemak izenaren azpian ez, baina ikaste-jardueretan integra daitekeen multimedia aplikazioa dugu.

Multi/hipermedia aplikazioak sortzeko autore-paketeak gero eta garatuagoak dira, HyperCard, Genesis eta ToolBook, besteak beste. Hizkuntz irakasle eta psikolinguistek begira sortutako programak ditugu hauek. Autore-paketeen bidez zenbait software bilakatu da, Estatu Batuetan batez ere, eta azken urteetan Britainia Handia eta beste zenbait herritan ere. Aholkatzea ikasteko MacEnglish izeneko ikaste-programa eta atzerriko herrialde batean moldatzeko Learn to Speak programa aipatutako autore-paketeen bidez sortutako aplikazioak ditugu.

II.2.4 Ordenagailuz Lagundutako Hizkuntzen Ikaskuntza Adimenduna.

Adimenduna terminoa aipatzen dugunean ordenagailuaren laguntza-motari buruz ari gara. Azken urteetan Adimen Artifizialaren Sistema Tutore Adimendunak (STA) eta Linguistika Konputazionalaren (LK) ekarpenak OLHIA arloan erabili ohi dira. Hau dela eta, atal honetan OLHIA arloan garatutako sistemak komentatuko ditugu. Dena den, tesi honek linguistika konputazionalaren tekniken integrazioa aurkeztea helburu duenez, linguistika konputazionalan sortutako teknikak integratzen dituzten sistemei atal berezia eskaini diegu, § II.2.5-n.

Hasiera batean Sistema Tutore Adimendunen helburua irakaslearen papera betetzea zen. Hori zela eta, lau modulu nagusi definitzen zituzten honelako sistemetan: bat Ezagutza-Basea, ikasi beharreko bigarren hizkuntza gordeko lukeena; bi, Irakaslearen Eredua, ikasleari buruzko informazioa bilduko lukeena; hiru, Modulu Pedagogikoa, irakaslearen rola ordezkatuko lukeena; eta, azkenik, Interfazea, sistema eta irakaslearen arteko elkarrekintzaz

arduratuko litzatekeena. Honelako sistemak, besteak beste, programatzen ikasteko (Fernández, 88; Díaz de Ilarraza, 90) garatu izan dira, baita hizkuntzen irakaskuntzarako ere.

Hizkuntza bat irakasteko sistemen artean lehenengoa 1978an Weischedel-ek eraikitako German Tutor izenekoa dugu. Honek alemaneraz idatzitako testua aurkezten zion ikasleari hainbat galderarekin batera. Ikasleak galdera hauei erantzun ondoren sistemak erantzunei buruzko laguntza ematen zion ikasleari. Ikasleak gaizki idatzitakoa detektatzeko gai ere bazen sistema hau. Pulman-en (1984) sistemak, ordea, ez zuen kontuan izaten ikaslea gaizki idatz zezakeenik. Geroztik, Schuster-en (1986) VP2 sistema ingelesaren phrasal verbs direlakoak itzultzen irakasteko sortu zen. Honek egiten zuen diagnostia beti ama-hizkuntzaren interferentziaren ondorio kontsideratzen zuen. Schwind-en (1990) tutore adimendunak, alemanera ikastekoa zenak, hainbat ariketa aurkezten zion ikasleari, honek aldi berean sistemari galderak egin ziezazkiokeelarik. Sistemak, bere aldetik, ikasleari egiten zizkion galderak ikasleak erroreren bat egiten zuenean. Orain arte aipatutako sistemetan tratatzen zen ezagutza mugatua eta aurreikusia zen. Beste sistema batzuetan, hots, Barchan et al. (1986), Yazdani (1990) eta Menzel-en (1988) kasuan, testu arruntak tratatzen ziren.

Aipatu berriko sistema adimendun hauek gramatikan oinarritutako sistemak ziren, beste batzuk, berriz, semantikan oinarriturik aurkitzen genituen. Azken hauen artean LINGO (Language Instruction Graphic Operations) (Kramsch et al., 85) programak gela baten errepresentazio grafikoan oinarritzen zen. Ikasleak objektuei buruzko informazioa eskatzeko aukera izateaz gain, objektuak mugitzeko aginduak eman zitzakeen. Antzeko ideia inplementatu zen Adventure Game Program (Culley et al., 86; Mulford, 89) izeneko aplikazioan. Azkenik, semantikan oinarritutako sistemen artean, SWIND (Zock, 92) eta CODA-X (Frederiksen et al., 92) ere aipagarriak dira. Biek kontzeptuen arteko errepresentazioa zuten oinarri: ama-hizkuntza eta xede-hizkuntza erlazionatzen zuen errepresentazioa hain zuzen. Azken aldian, semantikan oinarrituta garatu diren sistemen artean aipatzekoa da Morante et al.ek (1998) aurkezten dutena. Honetan EuroWordNet-en erabilera aztertzen dute gaztelera eta katalana ikasten ari direnentzat.

Ikusi dugunez, Ordenagailuz Lagundutako Hizkuntzen Ikaskuntza Adimenduna arloan hasiera batean garatu ziren sistemek irakaskuntzaren ikuspuntua zuten. Hau dela eta, irakaslearen papera betetzea ezinezkoa zela segituan konturatu ziren, zenbait kasutan tresna interesgarriak sortu baziren ere. Beraz, irakaslearen ikuspuntua utzita, ikaslearen partehartze aktiboa lantzea izan zen hurrengo pausua. Hau dela eta, azken hamar urteetan ikaslearen ereduaren garapena dezente aztertu da bai sistema tutoretan (Self, 92) bai hizkuntzen ikaskuntzarako sistemetan (Bull et al., 97).

II.2.5 Lengoaia Naturalaren Prozesamendua integratu duten sistemak.

Orain dela bost urtetik honako korranteak (Chanier, 94) OLHIA arlo diziplinartekoa aldarrikatzen du. Adimen artifizialean (sistema adituak, sistema tutoreak, ikaslearen modelizazioa...) eman aurrerapenak, linguistika konputazionalarenak eta psikolinguistikan burutu diren aurrerapausoak elkartu eta koordinatu behar ditugu OLHIA eremuan. Euskal Herrian pausoak emanak dira hiru diziplinetan: UPV/EHUKo informatika fakultatean sistema tutoreak aztertzeke taldea ari da lanean; euskara irakasteko Ikastxip, Benetan eta bai & by izeneko aplikazioak (Zubiri, 90) DDK sistema adituan oinarrituta daude; linguistika konputazionalan euskararako analizatzaile morfologikoa, lematizatzailea, datu-base lexikala, itzulpengintzarako tresna lagungarriak eta abar luze bat ditugu; eta, azkenik, psikolinguistikan lan ugari ere topatzen ditugu.

Tesi honetan, alde batetik, linguistika konputazionalan garatzen ari diren produktu eta prototipoak hizkuntzen ikaskuntza munduan integratzea proposatzen dugu, adimen artifizialak ematen dituen baliabideak direla medio. Honela, XUXEN zuzentzaile ortografikoa bera ikaslearen mailaren arabera egokitu dugu, betiere alde aurretiko azterketa teorikoa eginez. Beste alde batetik, hiztegiak ere ikaskuntza-material informatiko gisa planteatzen ditugu. Eta, azkenik, euskarazko corpusen azterketarako softwarea erabiltzen dugu azterketa psikolinguistikoetarako. Edozein kasutan, ematen ari diren aurrerapen informatikoak kontuan izanik eta euskararen ikaskuntza arloan dauden hutsuneak eta betebeharrak aztertuz, aipatutako hiru diziplinen arteko koordinazioa gauzatzea da tesi honen helburuetako bat. Gure ustez, honelako diziplinartekotasunak badu zer emanik euskararen ikaskuntzaren bilakaeran.

Filosofia honen ildotik, azken urteetan hizkuntzen ikaskuntza arloan sortu den hainbat sistematan lengoaia naturalaren prozesamenduan garatutako teknikak integratu dira. Eta hauei buruz arituko gara ondoren:

Holland et al.ek (1995) LNP tekniken ahalmena aztertzen dute hizkuntza bat ikasteko sistemetan integratu ahal izateko. Besteak beste, irakaskuntza-printzipioen eragina sistemen diseinuan eta ikasle-motaren araberrako sistemen diseinu-egokitzapena lantzen dute.

Lengoaia naturalaren prozesamenduan garatutako sistemen erabiltzaileak linguistak izaten dira gehienbat, eta, noski, emaitzen irteera-formatuak hauei begira diseinatuak dira. Beraz, hizkuntza bat ikasteko sistemetan egin behar den lehen gauza irteera-formatu hauen egokitzapena dugu: (Hamel et al., 98)-n, esaterako, arlo hau lantzen da gehienbat.

Honelako sistematan irteera-formatua ez da egokitu behar den gauza bakarra, sistemen arkitektura bera ere definitu behar da eta. Alde batetik, hizkuntzari buruzko ezagutzaren errepresentazioa; beste alde batetik erabiltzaile konkretuari laguntza berezia emateko aukera

eskaintzen duen ikasle-eredua; honetaz gain, ikaslearekin aritzeko sistemaren estrategiak; eta, azkenik interfazea. Sistema guztiek, noski, ez dituzte atal hauek guztiak garatzen. GLOSSER (Nerbonne et al., 96) sistemak hainbat hizkuntzaren irakurketan laguntzen du. Honetan interfaze berean ondokoa integratu da: analizatzaile eta desanbiguatzaile morfologikoa eta bilaketa-tresnak hiztegi elebidunetarako zein corpusetarako. GLOSSERen, beraz, ikaslearekin aritzeko estrategia sinple batzuk landu dira interfaze berean hainbat tresna integratuz. Filosofia bereko sistema, hau da, ikasle-eredurik gabekoa, aurkezten da (Werner, 94)-n. Sistema honek serbo hizkuntzaren idazketa-prozesurako laguntza eskaintzen du. Dimitrova et al.ek (1997) terminologia ikasteko ITELS izeneko sistema proposatzen du. Konkretuki aztertzen duena ondokoa da: lexikoiek, hitzak osatzeko erregelek eta termino semantikoek nola eragin dezakete ikaslearen ezagutzaren diagnosian zein ikaslearekiko saioaren gidan. Sistema honetan ere ez da ikasle-eredurik aurkezten. Tesi honen VI. kapituluaren aurkeztuko dugun IDAZKIDE sisteman berriz ikasle-eredua definitzen dugu. Horretaz gain ikaskuntza-maila berean dauden ikasleen ezagutza kolektiboaren eredua ere erabiltzen dugu. Schneider et al.ek (1998) ere honelako eredu kolektiboa gordetzen duen sistema aurkezten dugu. Sintaxi mailako erroreen tratamenduak esaldien analisi-kopurua nabarmenki igotzen duela eta, kopuru hau gutxitzeko asmoz, ikasle-populazio baten aurreikuspenak gordetzen dituen eredu kolektiboa erabiltzen du.

Lengoaia naturalaren prozesamendua integratzen dituzten sistemen munduan, sare bidezko komunikazio arloa ezin dugu utzi aipatu gabe. NLP teknikak integratzeaz gain hizkuntzaren edozein input-a tratatzen duten sistemak ere garatzen hasiak dira hizkuntzen ikaskuntza arloan. Hauen artean dugu japonieran pasiboa idazten ikasteko sistema (Yang et al., 97). Sistema honek hizkuntzaren arlo oso konkretua (esaldi pasiboak) landu arren ikaslearen edozein esaldi onartzen du. Azkenik, esan beharra dago tesi honetan garatu den ikaskuntza-ingurunearen helburua dela hizkuntzaren edozein input-a tratatzea.

II.3 Ondorioa: diziplinarteko garapena.

Informatikak hizkuntzen ikaskuntza munduan ekar dezakeenaren ikuspegi orokorra eta zabala eman ondoren, ondorio gisa diziplinarteko garapena proposatzen dugu.

Tresna informatikoen eraketa-prozesuan askotan informatikaren mugek arazo larriak planteatzen dituzte. Produktu informatikoen diseinu eta inplementaziorako beharrezko baliabideak asko izaten dira. Errekurtso ekonomikoez gain, diziplina desberdinen arteko aditu edota arituen elkarlana beharrezkoa zaigu. Ikasleentzako tresna linguistiko lagungarriak (zuzentzaile ortografikoak, zuzentzaile gramatikalak, definizio-hiztegiak, sinonimo-hiztegiak...) eraiki baino lehen, beharrak non dauden aztertzeko irakasle eta

psikolinguisten iritzia beharrezkoa zaigu. Aipatutako tresna linguistikoak diseinatzeko linguisten laguntza ere ezin da ahaztu, batez ere, hizkuntza normalizazio-prozesuan dagoen herri honetan. Irakasleentzat zenbait material era eroso batean prestatzeko tresnak diseinatzean diziplinarteko koordinazioa ere nahitaezkoa da. Eta azkenik, psikolinguistika arloan ikerkuntzan lagun ditzaketen tresnak osatu nahi baditugu ere koordinazioa beharrezkoa zaigu.

Beraz, psikolinguistikan, linguistika konputazionalan, informatikan eta euskaltegietan aritzen garenon artean eraman beharreko gaia dugu gure aurrean (Maritxalar & Díaz, 94). Eta hau da hain zuzen tesi honetan landu dena. Alde batetik psikolinguistika arloan garatutako ideiak izan genituen aztergai hasiera batean (III. kapitulua); ondoren, euskaltegietako hainbat irakaslerekin lan esperimentalak burutu genituen (III. eta VII. kapituluak); hurrengo pausu batean linguistika konputazionalako tekniken egokitzapena landu genuen (IV. kapitulua); eta, azkenik, HITES, IRAKAZI eta IDAZKIDE (V. eta VI. kapituluak) sistema informatikoez osatutako MUGARRI ingurunearen garapena burutu dugu, lan esperimentalak paraleloki burutu direlarik (VII. kapitulua).

III. Psikolinguistika arloan egindako lan experimentalaren deskribapena.

Hirugarren kapitulua hizkuntza baten ikaste-prozesuaren ikerketarako sistema informatiko adimendun baten oinarri psikolinguistikoa aztertzen dugu. Sistema informatikoaren helburu nagusia izango da: ikaslearen ikaste-prozesuari buruzko informazioa jasotzea; ikaslearen tartehizkuntza aztertzea; tartehizkuntzaren osaketan egon daitezkeen arrazoen hipotesiak lantzea; eta azkenik, ikaste-estrategiek tartehizkuntzaren osaketan duten eragina ikertzea. Honetarako, oraingoz, tartehizkuntzaren konpetentzia gramatikala hartu dugu aztergai, eta, konkretuki, egitura desbideratuetara murriztu dugu ikerketa. Edozein kasutan, azken helburua tartehizkuntza bere osotasunean lantzea izango denez, tartehizkuntzarako diseinatuko dugun eredu kontzeptuala irekia izateaz gain, independentea izango da ikasten den hizkuntzarekiko. Beste alde batetik, sistema inkrementala izango dugu, hau da, etorkizunean aukera izango dugu, tresna informatiko berriak sistema berean gehituz, ikaslearen konpetentzia komunikatiboa aztertzeko. Hau guztia dela eta, kapitulu honetan ondoko kontzeptuak aztertuko ditugu sakonki: desbideratzeak, tartehizkuntza eta fosilizazioa, azkenean azaleko mailako eta sakoneko desbideratzeen sailkapenak lortuko ditugularik. Ikaste-estrategiei buruzko azterketa VI. kapituluan landuko dugu; han gaia hobeto egokitzen da, ikasleentzako hizkuntz ikaste-sistema lagungarria azalduko baitugu.

III.1 Helburuak.

Azken bi hamarkadetan Bigarren Hizkuntzaren Ikaskuntza (BHI) arloan ikerketa ugari topatzen ditugu. Aldi berean linguistikan, psikolinguistikan, psikologia kognitiboan eta adimen artifizialean aurrerapen handiak burutu dira. Alde batetik, lau ikerketa arlo hauek BHI arloan lan egiteko ingurunea eskaintzen digute, eta, beste alde batetik, BHIren inguruan lortzen diren ondorioak lagungarri izan daitezke ere aipatutako ikerketa arloetan.

Aurrean agertzen zaizkigun ideiak, teoriak, esperimentuak eta abar ugariak dira; Krashenen Bigarren Hizkuntzaren norbereganatze-prozesuari buruzko teoria (Krashen, 77, 78); honi egindako kritika gogorrak (Gregg, 84; Long, 85; McLaughlin, 78; Taylor, 74); ikasleen erroreak aztertzeko errore-analisiaren prozedura (Corder, 74); honen baliotasunaren kritikak (Schachter & Celce-Murcia, 77 — *oso zaila da ziur jakitea ikasle bat zein errore-mota ari den egiten eta zergatik egiten ari den errore hori* —); **Tartehizkuntza** (Selinker, 72; Adjemian, 76; Tarone, 79) kontzeptuaren inguruan sortutako teoria; Chomsky-ren Gramatika Unibertsala (Chomsky, 65); eta abar luze bat. Gure lanaren helburua ez da izango

ez Linguistika, ez Psikolinguistika, ez Psikologia Kognitiboa, ez BHI arloan, ezta Adimen Artifizialaren arloan ere teoria berriak osatzea edo daudenak kritikatzeko. Gure kasuan, bigarren hizkuntzaren ikaste-prozesuaren azterketarako MUGARRI izeneko ingurune informatikoa diseinatu eta inplementatuko dugu. Hizkuntza bat ikasten duen pertsona baten ikaste-prozesua ikertzeko, ordenagailuz lagundutako sistema adimenduna lortu nahi dugu hain zuzen ere.

Hurrengo kapituluetan garatuko dugun sistemak edozein hizkuntzaren ikaste-prozesua lantzea izango du helburu. Dena den, bigarren hizkuntzatzen euskara aukeratu dugunez, euskararen problematikak hainbat erabaki bideratu du ikerketa honetan. Esaterako, hizkuntza estandarizazio-prozesuan egonik hainbat zalantza izaten da hizkuntza menperatzen dutenengan ere.

Interesgarritzat jotzen dugu azpimarratzea tratamendu orokorrerako sistema diseinatu dugula, hau da, testu errealean tratamendu automatikoranzko bidea hartu dugu, laborategi-adibideak baztertuz ikerketaren helburu gisa. Hizkuntz ezagutzari dagokionez ere askotan hainbat fenomenoren azterketa alhora utzi behar izan dugu euskararako egun dauden tresna linguistiko-konputazionalak ezartzen dizkiguten mugen arabera.

III.2 Tartehizkuntza eta fosilizazioa.

Bigarren hizkuntzaren ikaskuntza arloan "ez-ikastea" problema behin eta berriro agertzen zaigu. Ikasleak ikaste-prozesuan zehar bere tartehizkuntza osatzen du. Tartehizkuntza aldakorra izan arren zenbait egitura egonkortu egiten da, askotan xede-hizkuntzaren arauetatik urruti daudelarik. Egitura bat tartehizkuntzan egonkortzen denean **Fosilizazio** fenomenoaren aurrean gaude. Selinkerrek (1992) fenomeno hau ikasleen ez-ikaste problema aztertzeko bide egokia dela dio. Ikasleak pixkanaka bere tartehizkuntza osatzen du eta berarekin batera bigarren hizkuntzaren gramatika propioa. Gure sistemak ikaslearen kompetentzia gramatikala eta kompetentzia horren garatze-prozesua aztertzeko balioko du. Dena den, sistemarekin esperimendatzeko ikasleak ikasi nahi duen xede-hizkuntzaren gramatikarekiko egitura desbideratuak hartuko ditugu oinarritzat, bai **desbideratze** fosilizatuak bai ez-fosilizatuak. Beraz, azken helburua ikaslearen kompetentzia gramatikala eta honen jasotze-prozesuaren azterketa izan arren, esperimenduak egitura desbideratuetara murriztu ditugu.

III.2.1 Tartehizkuntza.

Ikasle batek bigarren hizkuntza ikasten ari denean, pixkanaka, ikasten duen heinean, tartehizkuntza propioa sortzen du. Tartehizkuntza hori aldatzen joango da xede-hizkuntza

"lortzen"¹ duen arte. Hasiera batean xede-hizkuntzari buruz ez daki ezer, eta, esan bezala, ikaste-prozesuaren ondorioz aipatutako tartehizkuntza sortzen du. Beraz, ikasle baten ikaste-prozesuan laguntza eman nahi badiogu beharrezkoa gertatzen da bere tartehizkuntza ezagutzea, tartehizkuntza hori ikaste-prozesuaren isla baita.

Selinkerrek (1972) eta Corderrek (1971b) diotenez, ikasleak duen egitura psikologiko ezkutuak tartehizkuntzaren elaborazioa gidatzen du. Bi hauek eta Nemser-ek (1971) definitzen dituzten tartehizkuntzak (Selinker, 72), dialekto trantsizionalak (Corder, 71b) eta sistema hurbilduak (Nemser, 71) ondoko hiru ezaugarri komunak dituzte:

- a) Ikasle baten diskurtsoa H1 (ama-hizkuntza) eta H2rekiko (bigarren hizkuntza) independentea da eta barnetik egituratua den sistema linguistiko baten emaitza da.
- b) Sistema linguistiko hori, ikaste-prozesuan zehar aldakorra da.
- c) Egoera zehatz batean, maila bereko ikasleen sistema linguistikoa oso antzekoa da, ikaste-esperientziaren ondorioz sortutako diferentzia batzuk kenduta.

Corderrentzat (1971b) dialekto idiosinkratiko baten gramatikak beste gramatika estandarren erregelak izateaz gain, beste dialekto estandarretan ez dauden erregela propioak izaten ditu. Dialekto idiosinkratikoa ez da dialekto soziala, hau da, ez du nahitaez talde sozial baten jokaera linguistikoaren isla izan behar. Dialekto idiosinkratikoa poeta batena da, afasiko batena, haurren hizkuntza eta bigarren hizkuntza ikasten dutenena. Dena den, berarentzat azken hau oso aldakorra denez, dialekto trantsizionala dela dio. Edozein kasutan, bigarren hizkuntza ikasten ari den ikasle baten uneko tartehizkuntza, dialekto idiosinkratikoa edo sistema hurbildua sistematikoa, erregelek gobernatuta eta komuna da ikaskuntza-maila bera duten ikasleentzat, ikaste-esperientziaren ondorioz sortutako diferentzia batzuk kenduta. Honegatik, diseinatu dugun sisteman ikaslearen tartehizkuntza adierazteaz gain, maila bereko ikasleen tartehizkuntz ereduak ere agertzen dira. Azken hauek maila bereko ikasleen perfilak errepresentatzen dituzte.

Selinkerren (1972) ustez, aipatutako tartehizkuntza ondoko bost prozesu psikologikoen eraginez sortzen da: gainjeneralizazioa, irakaskuntza transferentzia, ikaste-estrategiak, komunikazio-estrategiak eta interferentziak. Berak honela definitzen ditu prozesu hauek: **gainjeneralizazioa** egiten denean, erregela ez dagozkion inguruneetara zabaltzen da, edo erregelaren murriztapenak ez dira zuzenki aplikatzen. **Irakaskuntza Transferentziaren** iturburu arrazoi pedagogikoetan datza: irakastean erabilitako materialak transferitutako input-a dugu hau. **Ikaste-estrategiak** ikasleak subkontzienteki burutzen omen ditu. **Komunikazio-estrategiak** ikaslea, hizkuntzaren zuzentasuna nolabait baztertuz,

¹ Ikasleen %5ak hizkuntza zeharo menperatzea lortzen du (ik. Interlanguage. Selinker, 72)

komunikatzen saiatzen denean sortzen dira. Eta, azkenik, **Interferentziak** beste hizkuntza baten jakitearen ondorioz sortutakoak izaten dira, batez ere, ama-hizkuntzagatik, edo inguruan indar handiena duen hizkuntzaren eraginez.

Richards (1971), Selinkerrek definitutako bost prozesuetan oinarritzen da ikasleek egindako errorean azterketa egiteko, errorean azpian dauden arrazoiak aipatutako prozesu psikologiko horietan oinarrituz. Guk egingo dugun azterketan bide bera jarraituko dugu eta azaleko mailan agertzen diren desbideratzeen azpian dauden sakoneko arrazoiak topatzeko Selinkerren prozesu psikologikoak abiapuntu gisa hartuko ditugu. Dena den, sakoneko desbideratzeen sailkapena egitean aipatutako bost prozesu psikologikoak berrantolatuko ditugu, bai O'Malley & Chamot-en (1990) ikaste-estrategien sailkapena landuz, bai ikasle konkretu batzuen ikaste-prozesuaren azterketan bere irakasleek emandako sakoneko arrazoiak sailkatuz (Andueza et al., 96).

Tartehizkuntzaren ezagutzan sakontzeko ikasleak egiten dituen desbideratzeak zeresan handia dute. Beraz, desbideratzeak aztertuz, tartehizkuntza hobeto ezagutzeko aukera izango dugu. Honen ondorioz, ikaslearen ikaste-prozesuari buruzko informazio zehatzagoa edukiko dugu.

Desbideratze kontzeptuaz zabalago hitz egingo dugu § III.3-n eta tartehizkuntza kontzeptua zehatzago azalduko dugu hurrengo kapituluan. Han tartehizkuntzarako definitu dugun eredu kontzeptuala ikusiko dugu.

III.2.2 Fosilizazioa.

Psikolinguistikan, konkretuki bigarren hizkuntzaren ikaste-prozesuan, fosilizazio fenomenoak agertzen da. Batzuetan, ikasleak bere ikaste-prozesuan osatzen ari den tartehizkuntzaren alderdi bat lantzeari uzten dio, eta alderdi hori fosilizatu egiten da. Honegatik izan daiteke, batzuetan euskara nahiko ondo menperatzen duen ikasle baten idazlan batean ondoko esaldia topa dezakegula: “*Etxeko jabeak argi utzi **zidan** nire eginkizunak*”. Euskararen ikaskuntzan maila altua duten ikasleen artean zenbait egitura fosilizatuta topa dezakegu, ikasleek jarraitu duten ikaste-prozesua zeharo desberdina izanagatik. Irakasle askok intuitiboki probabilitate handiz esan dezakete zeintzuk diren maila altuko ikasle askok fosilizatzen dituzten egiturak (adib. ergatiboa, zenbait aditzen nominalizazioa, e.a.). Guk zenbait irakasleren iritziak kontutan hartuz, eta bi hizkuntz eskola¹ desberdinetan ikasten aritu diren ikasleen idazlanak aztertu eta gero, zenbait egitura fosilizatu topatu dugu.

¹IRALE eta ILAZKI

Bai ama-hizkuntzaren ikaskuntza, bai bigarren hizkuntzaren ikaskuntza aztertzen ari diren ikertzaileek egitura desbideratuak erabiltzea ikaste-prozesuaren baldintza beharrezkotzat jotzen dute, baina desbideratze horiek zenbateraino diren ikaste faltagatik edo zenbateraino landu behar diren ez dago hain garbi. Batzuetan ikasle batek xede-hizkuntzaren egitura bat ondo ezagutzen duela dirudien arren gaizki erabiltzen du gehienetan, eta, gainera, ez du inolako hipotesi probarik egiten egitura oker horretan xede-hizkuntzarantz hurbiltzeko asmoz. Egitura horren ikasketa-prozesua geldirik geratu da une batean, ez du aurrera egin eta, egitura ezagutu arren, ez du behar bezala bereganatu. Fenomeno hau, fosilizazio fenomenoa hain zuzen, oso nabarmena izaten da euskararen ikaskuntzan.

Fosilizatuta geratu diren desbideratzeak zuzentzeak askotan ez du inolako eraginik. Beraz, gure helburua ez da izango desbideratze horiek zuzentzea, baizik eta desbideratze horien fosilizazioa nola gertatzen den lantzea. Fosilizazio fenomenoek ikaskuntza eza aztertzeko bide ona eskaintzen digute (Selinker, 92:251); honegatik, aurkeztuko den sistema ez-ikaste prozesuari buruzko informazioa jasotzeaz ere arduratuko da.

Azken urtetan Euskalerrriko euskaltegietan eztabaidak izan dira ikaskuntza-maila bakoitzean jarraitzen diren programak direla eta. Irakasleen esperientziak dio hainbat egitura fosilizatu egiten direla ikasle gehienengan. Programak aldatu beharko lirateke, edo ...? Honek azterketa sakona eskatzen du. Guk uste, lan honetan aurkeztu dugun sistema modu bat izan daiteke honelako azterketa sakona lantzen hasteko.

III.3 Desbideratze kontzeptua. Sailkapenak.

Atal honetan, hizkuntza baten ikaste-prozesuan erroreak eduki dezaketen garrantziari buruzko zenbait ideia aipatuko dugu eta errore kontzeptua bera definitzen saiatuko gara ere. **Errorea** edo **desbideratzea**? Errore terminoak automatikoki zuzenketa hitza dakarkigu burura. Zenbait metodoren arabera, ikasle batek errore bat eginez gero, zerbait gaizki egin duela esan nahi du. Kontzepzio honetan ikaslearen ikaste-prozesu progresiboa baztertzen da nolabait. Guretzat erroreak ez dira erroreak izango, desbideratzeak baizik. Desbideratzeak ikasleak ikaste-prozesuan zehar osatzen duen tartehizkuntzaren parte izango dira, ikaslearen ikaste-prozesuaren islaren osagaia, eta ez gaizki egiten duen zerbait. Desbideratzeak zuzentzea batzuetan komenigarri izan daiteke, baina beste batzuetan ez da batere komenigarri izaten. Beraz, desbideratze terminoa erabiliko dugu, eta ez errore terminoa. Dena den, azterketa bibliografikoan errore terminoa erabiliko dugu, liburuetan topatu izan dugun bezala. Azterketa honetatik aparte, ordea, desbideratzei buruz hitz egingo dugu, hauek ikaste-prozesuaren islaren osagaia direla garbi utzi nahian.

III.3.1 Erroreen garrantzia hizkuntza baten ikaste-prozesuan.

Erroreak ikasbide positiboa ala negatiboa diren aztertuko dugu.

Hizkuntza baten ikaste-prozesuan zehar ikasle batek huts edo erroreak egitea ezinbestekoa da. Erroreak prozesuaren parte bat dira, dudarik gabe, baina, erroreen aurrean hartzen den jarrera oso desberdina izan daiteke irakaslearen arabera.

Hizkuntzaren irakaskuntza metodologiaren historian zehar erroreen aurrean zenbait joera desberdin topatzen dugu, adibidez, metodo estrukturala edo entzutezkoetan ikasleari ez zaio uzten erroreak egiten, metodologia honetan garrantzitsuena esaldiak gramatikalki zuzenak izatea da, nahiz eta esangurarik gabekoak izan. Honelako joera hartzen denean ikasleak hutsak egiteko beldurra izan dezake, irakasleak, ikaskideak edota jatorrizko hiztunek irrigarri utziko dutelakoan. Honen ondorioz, ikaslea inhibitu egiten da eta Stevick-ek (1976) dioen bezala, ikasketa defentsiboari ekiten dio. Ikasleak bere mezuaren edukia ahazten du, egitura gramatikalean arreta jarriaz, akatsik ez egiteko. Errorea akatsa da, ikasbide negatibotzat jotzen da. Garrantzi handia ematen zaio erroreari, eta ikaslearen ikaste-prozesuan oztopoa besterik ez da bihurtzen. Metodo kognoszitiboek berriz normaltzat jotzen dituzte ikaslearen erroreak eta ikasbide positibo gisa tratatzen dituzte.

Badirudi, erroreei garrantzia kendu behar zaien arren, hizkuntzaren ikaste-prozesuan ezinbesteko osagaia izan daitezkeela. Corderrek (1967) dio erroreak ikaslearen hipotesi zein estrategien ebidentzia direla. Erroreak ezin ditugu ikusi akatsak balira bezala, ikasbide positibo gisa baizik.

Corderrek (1971b) dioen bezala, ikasleak egiten dituen esaldi idiosinkratikoak bere dialekto trantsizionala edo tartehizkuntzaren gramatikaren isla dira eta ez ditugu ikusi behar erroreak edo esaldi agramatikalak bezala. Guk ere esaldi idiosinkratikoak tartehizkuntzaren gramatikaren isla bezala ikusiko ditugu. Dena den, beharrezkoa iruditzen zaigu desberdintzea ikaslearen ikaste-prozesuaren une bakoitzean esaldi idiosinkratiko horien artean xede-hizkuntzatik gertu daudenak eta zeintzuk diren ikasleak landu beharko lituzkeenak. Bere maila baino altuagoko egiturak ez ditugu sartuko landu beharreko horien artean. Ikasleak mailaren arabera menperatuta eduki beharko lituzkeen egitura gramatikalak izango dira aztertuko ditugunak.

Beraz, desbideratzeak, eta ez erroreak, ikaslearen tartehizkuntzaren osagai gisa ikusiko ditugu eta ikaste-prozesuaren barruan normaltzat joko ditugu.

III.3.2 Desbideratze kontzeptua.

Honez gero, azterketa batzuk egin ditugu IRALE eta ILAZKI euskaltegietako zenbait ikasleren idazlanetan oinarrituta. Testuak aztertzeaz gain, ILAZKIk irakasle batzuekin ikasleen idazkera landu dugu. Honen ondorioz, egitura desbideratuen kontzeptua argitu dugu. Desbideratze kontzeptua inon definiturik topatu ez dugun arren, Corder-ek (1971a) *superficially deviant* egiturak aipatzen ditu, eta Chanierrek (1994) ere aipamena egiten du *deviant form* bezala; azken honek desbideratze forma horiek errore kontzeptuaren ikuspegi positibo baten moduan definitzen ditu. Gure ustez, definizioa motz geratzen da "errore"en ikuspegi positiboa psikolinguistikan aldeztu aurretik landutako ikuspuntua baitugu; gu ikuspegi positiboa baino areago goaz desbideratze kontzeptua definitzean. Hortaz, egitura bat desbideratzea dela esango dugu ondoko hiru arrazoietakoa bat betetzen denean:

- Egitura linguistikoren bat "gaizki" erabili da, hau da, beti errore, akats... deitu izan dena.
- Egitura bat maizegi erabili da beste egitura bat edo batzuen ordez, azken hauek ekidinez (hots, eta partikula gehiegitan erabili *ordea*, *honegatik*... erabili beharrean). Kasu honetan lau konbinazio desberdin kontsidera daitezke ordezkatutako egitura zein ordezkapen-egitura gramatikalki era zuzenean edo okerrean erabiltzearen arabera.
- Egitura bat ekidin da, ordezkapen-egiturarik erabili gabe (esaterako, *nor da? nor ote da?* esan behar zenean).

Historian zehar, erroreak, akatsak, hutsegiteak, lapsusak... termino asko erabili izan dira. Termino hauen esanahia aztertu eta desbideratze kontzeptuarekin erlazionatzen saiatu gara.

Norrish-ek (1981) hiru hizkuntz jokabide anomalo bereizten ditu: errakuntza, hutsegitea eta lapsusa. **Errakuntza** desbideratze sistematikoari deitzen dio, ikasleak zerbait ikasi ez, eta, ondorioz, gaizki erabiltzen duena. Baina batzuetan ikasleak egitura bera zuzen eta beste batzuetan oker erabiltzen du. Desbideratze aldakor honi **hutsegitea** deitzen dio. Azkenik, kontzentrazio faltagatik, nekeagatik eta honelako arrazoiengatik erabilpen okerra gertatzen denean, **lapsusa** izendatzen du. Lapsusak ez du zerikusirik forma jakin bat irakatsia izatearekin, ez izatearekin ala irakasteko bidean izatearekin.

Edge-k (1989) ere hiru kategoria bereizten ditu: saialdiak, errakuntzak eta lapsusak. **Saialdiak** azalgaitzak diren erabilpen okerre deitzen die. Ikasleak ez du jakin egituratzen esandakoa ezagutza faltagatik, eta irakasleak ez du argi ikusten zein den ikaslea erabiltzen saiatu den egitura. **Errakuntzen** kasuan, ordea, ikasleak zuzentzeko ahalmenik izan ez arren, garbi izaten da zein den ikasleak erabili nahi zuen egitura. Azkenik, **lapsusak**, hau da, ikasle berak zuzen ditzakeen erroreak definitzen ditu.

Guk, ordea, lau kategoria bereiziko ditugu: saialdiak, errakuntzak edo desbideratze sistematikoak, hutsegiteak edo desbideratze aldakorrak eta lapsusak. **Desbideratze** terminoa orokorrean hitz egiteko erabiliko dugu. **Saialdiak**, Julian Edge-ren esanahiaren zentzuan, ikasleak ezagutza faltagatik egiten dituen esaldi ulergaitzak izango dira. Kasu hauetan, irakasleak ez du argi ikusten zein den ikaslea erabiltzen saiatu den egitura. **Desbideratze sistematikoak** John Norrishen errakuntzak izango dira, baina, ez bakarrik ikasleak zerbait ez ikastearen ondorio direnak. Askotan, euskara-maila altua duten euskaldun berriek gaizki erabiltzen dute ergatiboa era sistematiko batean, nahiz eta ergatiboa ondo ezagutu. Beraz, honelako "okerkuntzak", fosilizazio baten ondorioa izan litezkeenak, sistematikoak bihurtzen direnean, desbideratze sistematikoak deituko ditugu ere. Norrish-en hutsegiteei **desbideratze aldakorrak** deituko diegu. Hauek askotan gertatzen dira egitura zeharo asimilatuta ez dagoelako. Batzuetan ikasleak zuzen dezake bere hutsegitea, baina beste batzuetan ez. Fosilizatutako zenbait egitura askotan jokaera aldakor honetan ere islatzen da. Eta, azkenik, **lapsusak** kontzentrazio-falta, nekea eta honelako arrazoiengatik ikasleak egiten dituen "okerkuntzak", oso puntualak izateaz gain, ia beti ikasleak berak zuzen ditzake.

Beste ikuspegi batetik Corderrek (1973) hiru errore-mota definitzen ditu: **errore presistematikoak**, **sistematikoak** eta **postsistematikoak**. Presistematikoak ikasleak ezin ditu zuzendu, ezta errore hori egitearen arrazoiren bat eman ere. Gure kasuan, III.1 taulan agertzen denez, desbideratzeen mota identifikatzeaz gain, hau da, saialdia, desbideratze sistematikoa, desbideratze aldakorra edo lapsusa; hauen tratamenduari nola ekin ere detektatu behar da, hau da desbideratze azalezina, zuzentezina edo zuzengarria. **Desbideratze azalezinak** Corderrek definitzen dituen errore presistematikoak dira. **Zuzentezinak**, azaldu bai baina ikasleak zuzentzerik ez duena da. Azkenik, **desbideratze zuzengarriak** ikasleak azal eta zuzen ditzakeen errore postsistematikoak dira.

Desbideratze-mota versus Desbideratzearen tratamendua	Saialdiak	Desbideratze sistematikoak	Desbideratze aldakorrak	Lapsusak
Desbideratze azalezinak	X	X	X	
Desbideratze zuzentezinak		X	X	
Desbideratze zuzengarriak		X	X	X

III.1 taula. Desbideratze-moten tratamendua.

Beraz, ikaslearengan topatzen ditugun desbideratzeak saialdia, desbideratze sistematikoa, desbideratze aldakorra edo lapsusa diren identifikatu eta gero tratamenduari ekin beharko

diogu. Dena den, tratamenduari ekin baino lehen jakin beharko dugu ea desbideratze azalezina, desbideratze zuzentezina edo desbideratze zuzengarria den. Saialdiak beti desbideratze azalezinak izango dira eta lapsusak desbideratze zuzengarriak, baina desbideratze sistematikoak eta desbideratze aldakorak ezin ditugu aldeztirik identifikatu Corderren sailkapen honetan. Beraz, azken bi kasu hauetan ikaslearekin elkarriketaren bat izatea beharrezkoa gertatzen da desbideratze azalezinak, desbideratze zuzentezinak ala desbideratze zuzengarriak diren esan ahal izateko.

Behin desbideratze-mota identifikatuta izan ondoren, hau da saialdia, desbideratze sistematikoa, desbideratze aldakorra edo lapsusa den, eta desbideratzea azalezina, zuzentezina edo zuzengarria den jakindakoan, ikasleak duen hutsune kontzeptuala identifikatzeko bidean aurkitzen gara, eta, ondorioz, zuzentzeko estrategia erabaki dezakegu. Saialdien kasuetan ezingo dugu ezer egin desbideratzea ezin dugulako inolaz ere arrazoitu. Lapsusen kasuan, bai ikaslea, bai sistema bera, batzuetan desbideratzea azaltzeko eta zuzentzeko gai direnez, sistemak detektatzearen papera bakarrik egingo du, edo agian, ikaslea bere desbideratzeaz konturazazi. Desbideratze sistematikoen kasuan desbideratze azalezina (ikasleak ezin du azaldu) izanez gero, sistemak berak tratamendurako behar duen erabakia hartuko du; desbideratze zuzentezina izanez gero, desbideratze hori egitearen arazoiei buruz ikasleak esan dezakeena galdetzen saiatuko da, ahal izango duen neurrian noski. Azkenik, desbideratze sistematikoa desbideratze zuzengarria balitz, ikasleak berak zuzen dezala saiatuko da, beti zuzenketa gainbegiraturaz. Ikasleak egindako zuzenketa onargarria ez balitz sistemak identifikatutako desbideratze sistematiko zuzengarria ez duela ondo identifikatu esan nahiko luke. Desbideratze aldakorren kasuan, lehenengo ikaslea bere desbideratze aldakorraz konturazaziko du, batzuetan egitura era batean eta besteetan beste era batean erabiltzen duela erakutsiz, gero desbideratze aldakorra azalezina, zuzentezina ala zuzengarria den identifikatzen saiatuko da desbideratze sistematikoen kasuan komentatu dugun tratamendua eginez.

Honen guztiaren helburua ez da "errorea zuzentzea" izango, baizik eta ikasleak erabiltzen dituen egitura desbideratuak aztertzea, tartehizkuntzaren errepresentazioa fintzeko asmoz.

III.3.3 Desbideratzeen sailkapena.

Atal honetan egitura linguistikoren bat gaizki erabiltzearen ondorioz gerta daitezkeen desbideratzeak aztertuko ditugu. Beraz, desbideratze kontzeptuaren definizioan aipatutako lehenengo kategoria bakarrik landuko dugu, hau da, zentzu klasikoan errorea deritzona. Bibliografia aztertzean, kategoria hau izan da gehienbat aztertua topatu duguna. Egiturak

maizegi erabiltzea edota ekiditea, ordea, oso gutxi landutako desbideratzeak dira. Azken hauek ere aztertu dira gure landa lanean erabilitako corpusetan.

III.3.3.1 Aztertutako corpusak.

Corpusa biltzean jarraitutako irizpideak ondokoak izan genituen: hizkuntz eskola desberdinak, hainbat ikaskuntza-maila, ariketaka egindako idazlan sortak (adib. entzundako gai baten laburpena, eskaera-eskutitz bat, ipuin sortua, e.a.), eta, azkenik, hizkuntz ingurugiro diferenteak, hizkuntza ikasteko motibazio anitz eta ikaste-prozesu desberdinak jarraitzen zituzten ikasleak: hori bai, ikasleari eskatu zitzaion ikastarora erregularitasunez etortzeko eta idazlanak egiteko konpromezua.

Corpusa osatzen duten testuak kodetzean jarraitutako idazkeran ondoko informazioa adieraztea erabaki genuen: hizkuntz eskola, ikaskuntza-maila, ikaslea eta ariketa. Gure corpusa mota honetako kodeak dituzten testuek osatzen dute. Dena den, bada beharrezko zaigun zenbait datu, kodeketan adierazterik ez dagoena. Hau dela eta, irakasleak bidalitako ariketei buruzko informazioa zein ikasleen ezaugarri pertsonalak biltzeko IRAKAZI sistema erabili genuen corpusa biltzerakoan. Honen bidez jasotako informazioa ondokoa dugu, besteak beste: idazlanen ordena kronologikoa, idazlana idatzitako unean ikasgelan lantzen ari ziren egitura linguistikoak... Irakaslea egitura linguistiko konkretua lantzen ari bada, egitura hori behin eta berriro erabiltzen saiatzen diren ikasleak topatuko ditugu, eta bai erabiltzea ekiditen dutenak ere. Beraz, unean uneko landutako egitura linguistikoak biltzea garrantzitsua da, irakasleak bidalitako ariketarekin batera¹.

Ikasle batzuek egituren erabilpena behartzen duten bitartean beste batzuek ekidin egiten dute. Beraz, testuak ezin dira aztertu ikasleen ikaste-estrategiak eta zenbait ezaugarri pertsonal kontsideratu gabe. Honegatik, lehen aipatutako IRAKAZI sistemaren bidez honelako ezaugarriak jaso genituen: "zuzentasunari begiratzen dio", "arriskuak hartzen ditu egitura berriak erabiliz", "zalantzen kasuan egiturak ekiditen ditu"... Ikaste-estrategiekin loturik dauden ezaugarri hauez gain ikaslearen adina, ikasteko arrazoiak, bere hizkuntz ingurugiroaren erreferentziak, ama-hizkuntza, ikaste-prozesuaren historia, eta abar ditugu bilduta.

Lehen aipatutako informazioa bai ikasleekin bai irakasleekin izandako elkarrizketetatik jaso genuen (Andueza et al., 96). Aztertutako sistema garatu aurretiko fase esperimentalean erabili genuen corpusa 300 bat testuk osatzen dute. Testuak hizkuntz eskoletako (IRALE eta ILAZKI) ikasleek idatziak dira 1991etik 1997rako urteetan zehar. Corpusa bi azpimultzo

¹ Gure esperimentuetan kasu batzuetan izan ezik ikasgelan landutako egitura linguistikoen eragina idazlanetan aztertu genuen. Etorkizuneko lanetarako horri buruzko azterketa zabalagoa egitea gomendatzen da.

nagusitan banatu da ikaskuntza-mailaren arabera: ikaskuntza-maila ertainak eta ikaskuntza-maila altuak. Testu guztiak ez ditugu era berean aztertu, eta batzuk baztertu egin dira: ikasle beraren testu gutxi, testuei buruz jasotako informazioa ez oso zehatza... Testuen azterketa hiru fasetan banatu dezakegu:

Lehenengo fasea: testua aztertze-unitatea izanik, hasiera batean, hau da 91/92 ikasturtean, 47 testu (6. eta 7. ikaskuntza-mailak) erdiautomatikoki aztertu genituen hitz-mailan, eta hala, hitz-mailako egitura linguistikoen lehen sailkapena lortu genuen (Maritxalar et al., 93). Ondoren, 93/94 eta 94/95 ikasturteetan beste 150 bat testu (7., 8. eta 9. mailak) ere jaso genituen hitz-maila eta esaldi-mailako hainbat egitura linguistiko erdiautomatikoki landuz. Hasierako 47 testuak eta ondorengo 150 testu haietako hogeitaz aztertu zituzten irakasleek ez zituzten ikasleak ezagutzen.

Bigarren fasea: 94/95 ikasturtean lan egin zuten bi irakasletako bakoitzak 150 bat testuetako (9., 10-11. eta 12. mailak) corpusa aztertu zuen. Irakasleek astean bost ordu eskola zuten, eta idazlanak hamabostero, batzuetan astero, idazten zituzten. Kasu honetan irakasleek zuzeneko harremana zuten ikasleekin. Lan experimental hauen emaitzen ondorioz, beharrezko genituen tresna linguistiko-konputazionalak definitu genituen.

Hirugarren fasea: ondorengo ikasturteetan, 95/96 eta 96/97an, bigarren faseko corpusaren azpimultzoak aztertu genituen sortzen ziren hizkuntz tresnak erabiliz. Eta beste irakasle batek emaitzak ebaluatu zituen. Irakaslearen ustez, hainbat kasutan detektatutako egitura linguistikoak testuaren ikaskuntza-mailari egokitzen zitzaizkion, beste hainbat kasutan, ordea, detektatutako desbideratzeak nagitasunez idatzitako testuen lapsustzat eman zituen irakasleak. Irakasle berak 96/97 ikasturtean bere gain zuen ikasle-talde baten 100 bat idazlan (10. maila) jaso eta aztertu genituen. Bigarren faseko idazlanetan agertutako lapsus haiek bere ikasleen idazlanetan ere agertzen ziren, – irakaslearen ustetan bere ikasleen lapsusak ziren –. Bukaerako froga bat egin genuen lapsus hutsak ziren edo ez konprobatzeko (ikus VII. kapitulua): zenbait kasutan lapsusak ez baizik eta desbideratze barkeratuak zirela ondorioztatu genuen, irakasleak uste zuena deuseztatuz.

III.3.3.2 Azaleko versus sakoneko taxonomiak.

Desbideratzeen sailkapenei buruz dagoen zenbait ikerketa oinarritzat hartuz, eta dagoeneko aztertu ditugun corpusak ere kontsideratuz¹, gure desbideratzeen lehenengo sailkapena egin dugu.

¹IRALE 91-92 eta ILAZKI 92-93 eta 94-95

Desbideratzeen tratamendua egin ahal izateko desbideratzeen taxonomia baten beharrean aurkitzen gara. Helburua tresna informatikoa egitea dugunez, argi dago, ordenagailuak sailkapenak egitea eskatzen digula; dena den, hori ez da arrazoi bakarra izan: desbideratzeen tratamenduaren bibliografia aztertzean sistema informatiko baterako egokia iruditu zitzaigun *errore-analisisa prozedura* (Corder, 74) oinarritzat hartu genuen hasiera batetik (Maritxalar et al., 93), – Ellis-ek (1985) dio errore-analisiaren tradizioa 1970aren aurretik datorrela –. Ondoren, analisi kontrastiboaren hazkundearekin batera, errore-analisiaren interesa jaitsi zen. Geroago, ordea, hirurogeiko hamarkadaren bukaeran, berriro landutako prozedura dugu. Azkenean, Corderrek (1974) definitutakoa izango da errore-analisisa deritzona. Erroreak aztertzeko prozedura honetan ama-hizkuntza ez zen eragigai bakarra, analisi kontrastiboan moduan, hala ere, ama-hizkuntzaren transferentzia ez zen baztertu. **Errore-analisisian** bost pausu nagusi definitzen dira, eta hauek dira gure **desbideratze-analisisian** jarraitu ditugunak:

- 1 Hizkuntz Corpusaren aukeraketa.
- 2 Corpuseko erroreen identifikatzea (gure kasuan desbideratzeak).
- 3 Erroreen sailkapena (Gure azterketan **azaleko maila hutsezko desbideratzeak** eta **desbideratze linguistikoak** — *ondoren aztertuko ditugun sailkapenak* —)
- 4 Erroreen azalpena (Desbideratzeen zergati psikolinguistikoaren identifikazioa — hauek dira **sakoneko desbideratzeak** —)
- 5 Erroreen garrantziaren ebaluaketa.

Taxonomia desbideratze idatziei buruzkoa izango da, eta, konkretuki, zazpigarrenetik hamabigarren ikaskuntza-mailarako ikasleek idatzitako testuak aztertzean landutako taxonomia izango da landuko duguna.

Baina, zeren arabera definitu dugu gure taxonomia? Desbideratze bera egin arren, berdina al da hirugarren mailako ikasle batek egindako desbideratzea edo zortzigarren mailako batena? Seguraski desbideratze hori egitearen arrazoiak/k desberdina/k izango da/dira kasu bakoitzean. Bi maila agertzen zaizkigu hemen, **azaleko desbideratzeak** (adib. "haiek zeunden" "haiek zeuden"en ordeaz erabiltzea) eta **sakoneko desbideratzeak** (adib. Zortzigarren mailako baten kasuan hizkuntza beraren barruko interferentzia egon da, bi hitzek antz handia dutelako. Behe mailako batek desbideratze bera egingo balu, agian, arrazoiak pluraleko hirugarren pertsona ez ezagutzea litzateke, eta, horren ondorioz, ezagutzen duen forma bakarra erabili duela pentsa genezake.). Hau dela eta, taxonomia bakarra ez da nahikoa izango, bi taxonomia nagusi behar ditugu: **azaleko taxonomia** eta **sakoneko taxonomia**. Beraz, desbideratze bat aztertzerakoan azaleko taxonomian dagozkion mota-kodeak identifikatuko ditugu, baita sakoneko taxonomian dagozkion mota-kodeak ere. Gainera, kontuan izan behar dugu desbideratze bati ez dagokiola beti azaleko mailako desbideratze-mota bakarra, ezta sakoneko bakarra ere. Desbideratze bera egitearen arrazoiak bat baino gehiago izan daitezke, dudarik gabe.

Jarrai dezagun aztertzen nondik abiatu taxonomia hauek. Lehengo adibidean, zortzigarren mailako ikaslearen kasuan sakoneko mailan ikusten dugu antz handiko bi hitz nahastu dituela. Desbideratze berari, berriz, azaleko mailan, bi ikuspegi desberdin eman dakiok. Ikuspegi linguistiko batetik singularreko bigarren pertsona eta pluraleko hirugarrena nahastu ditu ikasleak, eta azaleko ikuspegi huts batetik bi hitzen arteko ordezkapena egin du. Beraz, azaleko mailan beste bi maila bereiziko ditugu: **azaleko maila hutsa** eta **maila linguistikoa**. Honegatik, hiru taxonomia desberdin definitu beharko ditugula iruditzen zaigu: azaleko mailan bi eta sakoneko mailan bat.

III.3.3.2.1 Azaleko maila hutsezko desbideratzeak.

Honelako taxonomia batek aztertzen du zein modutan ikasleak aldatzen dituen hizkuntzaren azaleko egiturak. Irakurritako bibliografia (Dulay, 82) zein corpusak azertu ondoren guk jarraituko dugun taxonomia azaltzeari ekingo diogu. Taxonomia honetan 4 desbideratze nagusi definitzen ditugu:

A1 Omisioa edo kenketa

Ondo osatutako esaldi batean agertu behar den itema agertzen ez denean, omisioa edo kenketa gertatzen da. Kendutako itema letra, morfema, hitza edo hitz-multzoa izan daiteke: (adib. *asi* → *hasi*).

A2 Gehikuntza

Ondo osatutako esaldi batean agertu behar ez den itema agertzen denean, gehikuntza baten aurrean gaude. Agertutako itema letra, morfema, hitza edo hitz-multzoa izan daiteke.

Hiru gehikuntza-mota bereiziko ditugu (Dulay, 82):

- Marka bikoitza (adib. *Ez dut nahi ezer ez*)
Egitura batzuetan ezaugarri bera bi aldiz markatzen da, behin bakarrik azaldu behar denean.
- Erregularizazioa (adib. *Bizitzen "bizi" esan beharrean*)
Egitura irregular bati legokiokeen erregela erregularra aplikatzen zaionean.
- Gehikuntza sinpleak
Marka bikoitza eta erregularizazioa ez diren gehikuntzak ditugu hauek.

A3 Erakuntza

Letra, morfema, hitza edota egitura ez-zuzena erabiltzen denean erakuntza gertatu dela esango dugu. Kasu honetan lau erakuntza definituko ditugu. Dulayek (1982), berriz, hiru bereizten ditu.

- Erregularizazioa (adib. *seme-alabak* esan beharrean *semeak* hitza erabili)
Marka erregularra erabiltzen da marka irregular baten ordean egitura ez-zuzena lortuz.
- Archiformak (adib. "*hau*" erakuslea erabili "*hau, hori eta hura*" erabili behar den guztietan)
Forma-klase baten elementu guztiak adierazteko klase horietako bat aukeratzen da.
- Forma trukaketa (adib. *zu zeuden* eta *haiak zeunden*).
Forma-klase baten barruan elementu bat bestearen ordean erabiltzen da eta alderantziz.
- Asmaketa (adib. *cotxe* forma erabili *kotxe* formaren ordean)
Xede-hizkuntzan existitzen ez den letra, morfema edo egitura erabili du esaldian integratuz. Erakuntza-mota hau ez dago definituta (Dulay, 82)-n.

A4 Ordena aldaketa

Letra, morfema edo hitzen posizioa aldatzen denean: (adib. *alai ahots* bat bihurtzen da).

III.3.3.2.2 Desbideratze linguistikoak.

Azaleko maila hutsean "itxura" bereko desbideratzeak izan arren (adibidez: *dakiolako*; *oihanøan* → *LEKTE*¹) maila linguistikoan fenomeno linguistiko desberdinak topatzen ditugu (*dakielako* → *esaldi kausala*; *oihanean* → *inesiboaren deklinabidea*)

Honetaz gain, ikuspegi linguistiko batetik badirudi bi maila desberdin behar direla: maila linguistikoa eta maila metalinguistikoa. Maila linguistikoan hizkuntza da aztergaia, maila metalinguistikoan, berriz, linguistika bera lantzen da. Adibidez, *s/z* alternantzia arazo linguistikoa den bitartean, maileguen osaketa, lexiko sorkuntzarako eremuan sartzen denez, maila metalinguistikoan dago. Oraingoan, arazo metalinguistikoak linguistikoaren artean sailkatuko ditugun arren, alde metalinguistikoa identifikatuko dugu, etorkizunean bi maila hauen azterketa sakonago baten atea irekita utzi nahian.

Fenomeno linguistikoaren tratamenduari dagokionez, bi motatako arrazoiak eduki ditugu kontuan: informatikoak eta psikolinguistikoak. Arrazoi informatikoen artean nagusiki gaur egun linguistika konputazionalak ezartzen dizkigun mugen menpe daude eta konkretuki euskararako egin denaren menpe. Momentu honetan euskararako analizatzaile morfologiko

¹LEKTE: E LEtra Kendu hitzaren Tarteko posizio batean.

automatikoak egin daude, 70.000 sarrera inguru landuak gordeta dituen datu-base lexikala ere badago, eta honekin guztiarekin anitz milioi hitz analiza ditzakegu automatikoki. Honek printzipioz hitz mailan edozein desbideratze-mota automatikoki tratatu ahal izateko aukera ematen digu. Esaldi-mailan euskararekin egindako lana oraingoz nahiko teorikoa da, nahiz eta, honez gero, saiakera batzuk egin ditugun zenbait sistema konputazionalekin, Constraint Grammar (C.G.), Lexical Functional Grammar (L.F.G), PATR-II (Gojenola et al., 94) e.a. Honegatik, gure lanean hitz mailan gertatzen diren desbideratze gehienen arrazoiak bilatzen saiatu gara, eta esaldi-mailan, ordea, gutxi batzuk. Esaldi-mailako desbideratzeen aukeraketa egiteko momentuan hizkuntzen ikaskuntzan aipatzen den fosilizazio fenomeno hartuko dugu abiapuntutzat. Beraz, esaldi-mailan aztertuko ditugun egiturak izango dira, hasiera batean behintzat, maila altuetan irakasleen zein euskararen irakaskuntzaren esperientziaz arazoak sortzen dituztenak, agian, fosilizazioarekin lotuta egon daitezkeenak; hauen zerrenda osatu ondoren, bi filtro pasa ditugu: bata, euskararen normalizazio-prozesuak markatua, hau da, hizkuntza estandarrean definitu gabe dagoen zenbait egitura sintaktiko baztertu dugu gure azterketan; eta bestea, Linguistika Konputazionalak markatzen dizkigun mugek jartzen dutena, euskararako tratamendu automatikoan existitzen diren analizatzaile sintaktikoak kontuan izanik.

Sailkapen linguistikoa egiteko hartu ditugun abiapuntuak ondokoak izan dira:

- Aurretik aipatu ditugun hiru azterketa esperimental: IRALE hizkuntz eskolatik 91/92 ikasturtean jasotako 47 idazlanen eskuzko azterketa; ILAZKI hizkuntz eskolatik 92/93 ikasturtean zehar jasotako 20 testuren azterketa erdiautomatikoak; eta hizkuntz eskola beretik 94/95 ikasturtean zehar bildutako 150 bat idazlanen azterketa erdiautomatikoak.
- Ondoko lan bibliografikoak: lehenik, Zubiritik (1991) ateratako zenbait errore tipiko, 1. eta 5. mailen arteko ikasleei egokitu arren eta gure testuetan maiz agertu ez arren, euskararen hizkuntz maila altua duten pertsonen, gure ustez, askotan egiten dutena; eta, bigarrenik, XUXEN euskararako zuzentzaile ortografiko automatikorako egindako errore tipikoen azterketa (Aduriz, 94). Bigarren lan honetan edozein euskaldunek egin ditzakeen errore tipikoak kontuan hartu dira.

Hitz mailako sailkapenean XUXEN zuzentzaile ortografiko automatikorako egindako errore tipikoen azterketa oinarritzat hartuz, azaleko maila hutsezko sailkapenean identifikatutako desbideratzeak erlazionatu ditugu. Zuzentzaileko egindako errore tipikoen azterketan bi multzo nagusi desberdintzen dira: a) erregelen bidez orokortu daitezkeen fenomenoak eta b) morfemak tratatu beharreko fenomeno zehatzak. Fenomeno zehatz batzuk morfemak tratatzeko, sistema lexiko paralelo bat dugu. Sistema honetan morfema desbideratuak sartzen dira sistema lexiko orokorrean dagokion morfema zuzenarekin lotuta.

B eranskinean, bi multzo nagusi hauek zehatzago aztertzen ditugu, eta azaleko mailan egindako sailkapenarekin erlazionatu.

94/95 ikasturtean ILAZKI euskaltegian burututako azterketa esperimentalean euskaltegiko hiru talde desberdinetako ikasleek zein modutan erabiltzen dituzten zenbait fenomeno linguistiko aztertu genuen esaldi-mailan. B eranskinean, aztertutako fenomeno linguistikoak, dagozkien erregela linguistikoak eta identifikatutako testuinguru linguistikoak definitzen ditugu.

III.3.3.2.3 Sakoneko desbideratzeak.

Lehen aipatu den bezala, Selinkerrek (1972) bigarren hizkuntzaren ikaskuntzan bost prozesu psikologiko nagusizat hartzen ditu tartehizkuntzaren osaketan, *gainjeneralizazioa, irakaskuntza transferentzia, ikaste-estrategiak, komunikazio-estrategiak eta interferentziak*. Ez dugu ahaztu behar Selinkerrek tartehizkuntzari buruz ari denean gaitasuna (*competence*) eta ariketa (*performance*) terminoetan ari dela. Testuinguru honetan erroreak ere kontuan hartzen ditu, Beraz, prozesu psikologiko hauek tartehizkuntzaren osaketa-prozesuan garrantzi handikoak direnez, tartehizkuntzaren errepresentazioan desbideratzeak ere kontuan izan ditugu tartehizkuntzaren parte bat direlako. Gure ustez tartehizkuntzaren osaketa ez da prozesu psikologikoen ondorioz bakarrik gertatzen, tartehizkuntzaren osaketan prozesu psikologikoen garrantzi handia, edo agian handiena, izan arren ingurunearen faktoreek ere badute zeresanik. Honegatik ikasleak egiten dituen azaleko desbideratzeen azpian dauden sakoneko arrazoiak bilatzean ikaslearen barruko prozesu psikologikoetan eta ingurunearen faktoreetan bilatuko ditugu.

Gure ikuspuntutik Selinkerrek barruko eta ingurunearen faktoreak nahasten ditu. Irakaskuntza Transferentzia ingurunearen faktorea denez, ez dugu uste prozesu psikologiko bezala defini daitekeenik, prozesu psikologikoa ikaslearen barruan gertatzen baita, eta ez ingurunean. Beraz, Irakaskuntza Transferentzia ingurunearen faktoreen artean sartuko dugu. Beste alde batetik Interferentzia edo transferentzia negatiboaz hitz egiten duenean, transferentzia positiboa ahaztu egiten du. Guk prozesu psikologikoa ez dugu interferentzietara murriztuko eta transferentziaz hitz egingo dugu, hau da bai transferentzia positiboa bai negatiboa prozesu psikologiko bera direla kontsideratuko dugu. Azkenik argitu nahiko genuke Selinkerrek gainjeneralizazioa, interferentzia eta ikaste-estrategiak desberdintzen dituen arren guk bai gainjeneralizazioa bai interferentzia (gure kasuan transferentzia) ikaste-estrategien barruan sartuko ditugula, biak hizkuntza ikasteko estrategiak baitira (ikus O'Malley & Chamoten (1990) ikaste-estrategien sailkapena). Honegatik ikaslearen bost prozesu psikologiko nagusi hauek guretzat bi izango dira: **Ikaste-estrategiak** eta **Komunikazio-estrategiak**. Prozesu psikologiko hauetatik aparte ingurunearen faktoreak kontuan izan beharko ditugu. Ingurunearen faktoreak hiru

azpimultzotan banatuko ditugu: **Irakaskuntza Transferentziak**, **Komunikazio-egoerak** eta **Faktore Soziokulturalak**. Gure kasuan ikasleak normalizazio-prozesuan dagoen hizkuntza bat ikasten ari dira hizkuntza hitz egiten den herrian, beraz inguruneak bigarren hizkuntzaren input-a ematen die egunero eta ikaslearen tartehizkuntzaren osaketa-prozesuan bere eragina du.

Selinkerrek definitutako prozesu psikologien berrantolaketa oinarritzat hartuz eta 92/93 ikasturtean ILAZKI euskaltegiko ikasleek idatzitako testuen azterketa egitean irakasleekin edukitako zenbait elkarrizketa kontuan izanik, sakoneko desbideratzeen taxonomia azaltzen dugu C eranskinean. Taxonomia hau oinarria izan da 94/95 ikasturtean ILAZKI euskaltegian egindako azterketa esperimentalean.

Hau guztia dela eta bi atal nagusi sortu zaizkigu sakoneko desbideratzeen taxonomian: **Barruko Faktoreak** edo **Prozesu Psikologikoak** eta **Ingurunearen Faktoreak**. Barruko faktoreen barruan hiru mota desberdin topatzen ditugu, **Ikaste-estrategiak(SI)**, **Komunikazio-estrategiak(SK)**, eta **Bestelakoak(SB)**. Ingurunearen faktoreen barruan hiru mota desberdin topatzen ditugu, **Irakaskuntza Transferentziak(SR)**, **Komunikazio-egoerak(SE)** eta **Faktore Soziokulturalak(SS)**.

Egindako azterketa bibliografikoaren ondorioz eta 94/95 ikasturtean ILAZKI euskaltegian burututako azterketa esperimentalean irakasleek emandako arrazoen ondorioz, ikasleek egitura linguistikoak modu batean edo bestean erabiltzeko egon daitezkeen arrazoen zerrenda atera dugu. C eranskinean, aipatutako arrazoi-zerrenda azaltzen dugu. Kasu batzuetan adibideak aipatzen ditugu, ikasleak esandakoa letra etzanaz agertuko delarik. Aldi berean, arrazoi bakoitzaren bukaeran parentesi artean honako hau azaltzen da: sakoneko desbideratzeen taxonomiari dagozkion kodeak.

Ikusten denez, **SE** kodea ez da agertuko aipatutako sailkapenean, arrazoi-bilaketan komunikazio-egoerek sortutako hizkuntz erabilpenak testuen azterketan ez dagoelako detektatzerik.

Sailkapenean arrazoen artean honako hauek desberdintzen ditugu: zeintzuk dauden lotuta testu osoarekin eta zeintzuk dauden lotuta aztertzen den egitura linguistiko konkretuarekin. Beraz, bi arrazoi-multzo bereizten ditugu: testu-mailakoak eta esaldi-mailakoak, azken hauek egitura linguistiko konkretuei lotuta — kontuan izan esaldi-mailan gerta daitekeela esaldi berean egitura linguistiko bat baino gehiago aztertu behar izana, ondorioz, bakoitzerako arrazoiak aztertu beharra —.

Hasieran esan genuen tartehizkuntza bere osotasunean tratatzeko sistema informatikoa egin nahi genuen arren, esperimentuak egitura desbideratuetara murriztuko genituela. Honek

beti eramaten gaitu erne egotera diseinuan, desbideratuak ez diren egiturekin lan egitea erabakitzen dugunean. Hala eta guztiz ere, ikaslearen tartehizkuntzaren osaketa-prozesuan eragina eduki dezaketen faktoreak atal honetan aztertu ditugunak izango dira hasiera batean, behintzat tartehizkuntzaren *edozein* egitura linguistikoren osaketa-prozesuan. Dena den, sistema informatiko irekia diseinatuko dugunez, tartehizkuntzaren osaketaren eragigai berriak sortu ahala, sisteman integrazeko aukera izango dugu.

III.4 Esperimentuetan erabilitako metodologia.

Atal honetan landa-lanean psikolinguista, irakasleak eta informatikariak partaide zirela, burutu ditugun esperimentuak laburbilduko ditugu § III.4.1-n. Besteak beste 94/95 ikasturte osoz iraun zuen landa-lana syntaxian zein sakoneko arrazoien azterketan garrantzitsua izan zen. Honetan bi irakaslek ikasturte osoz aztertu zituzten ikasleen corpusak eskuz. Esperientzia honi buruz zehatzago arituko gara § III.4.2-n.

III.4.1 Burututako esperimentuak.

Gure azterketan informazio-iturriak ugari izan dira. Alde batetik, 300 testuz osatutako corpus bat erabili dugu, corpus hau 1991/92 ikasturtean IRALE eskolatik jasotako 47 idazlanek eta 1992/93an, 1994/95ean zein 1996/97an ILAZKI eskolatik jasotako 250 bat idazlanek osatzen dute; beste alde batetik, euskara bigarren hizkuntza moduan ikastean egiten diren errore tipikoei buruzko bibliografia (Zubiri, 91; 94) aztertu dugu. Gainera, ILAZKI hizkuntza eskolako zenbait irakasleekin batera, bai corpusetik, bai bibliografiatik ateratako informazioa komentatu izan dugu zenbait kasutan. Beraz, ikasleek idatzitako testuetan islatzen den tartehizkuntzaren gramatika zenbateraino den konpetentzia eta zenbateraino den performantzia detektatzea zaila denez, informazio-iturri desberdinak hartu ditugu fidagarritasun-maila hobea izango delakoan. Informazio-iturri hauek direla medio, desbideratzeen taxonomiak definitu genituen hasieran. Taxonomia hauek 1994/95 eta 1996/97 ikasturteetan zehar ILAZKI euskaltegian ikasten ari ziren hiru maila desberdinetako ikasleen idazlanen azterketan oinarri izan ziren. Azterketa honetan ikaslearen tartehizkuntza eta honen garatze-prozesuaren atal batzuk aztertu genituen. Azkenik, azterketaren ondorioz, taxonomi finduagoak lortu ditugu. Honetaz gain, aztertu diren mailei egoki liezazkiekeen tartehizkuntz ereduak, edo, beste modu batean esanda, perfilak landu dira.

Ondoren 1994/95 ikasturtean burututako landa-lana zehatzago deskribatzeari ekingo diogu. 1996/97 ikasturtean egindako esperimentuen emaitzak, ordea, VII. kapitulan azalduko ditugu.

III.4.2 Azterketa experimentalak.

III.4.2.1 Helburuak.

Lan experimentalak azterketaren zati handiena bete zuen. Lan experimentalean landu diren hiru arloek beren ekarpenak egin zituzten. Aipatutako arloak ondokoak dira: Linguistika konputazionala, Psikolinguistika eta Adimen Artifiziala. Edozein kasutan, lan experimentalaren helburu nagusia psikolinguistikoa zen, beti ere sistema informatiko baten diseinuak helburu horiek bideratzen zituelarik. Helburu zehatzagoak ondokoak ziren:

1) Hizkuntza baten ikaste-prozesuan ikasleak bereganatzen duen tartehizkuntzaren egitura aztertzekeo tresna lagungarriak garatzea.

Lehenengo helburua bete genuen, ikasleen tartehizkuntza aztertzekeo metodoa finkatu baikenuen. Metodoa ondokoa izan zen:

- Ikasleen aukeraketarako irizpideak zehaztu.
- Aztertu beharreko egitura linguistikoak aukeratzeko irizpideak finkatu.
- Ikasleen idazlanak astero bildu eta transkribatu, corpus bat osatuz.
- Irakasleek idazlanetan aukeratutako egitura linguistikoak identifikatu eta kodetu.
- Irakasleek identifikatutako egitura linguistikoaren sakoneko azterketa egin, ikaslearen ezaugarriak kontuan hartuz.
- Azken bietan lortutako informazioa kodetzeko eta ustiatzeko tresna informatikoak diseinatu eta inplementatu (ezagutza-basea eta tresna batzuk UNIXen PERL eta C lengoia erabiliz)

2) Maila desberdinetako hizkuntz ikasleek egiten dituzten idatzizko azaleko desbideratze eta hauen sakoneko zergatien azterketetatik taxonomiak definitzea.

Bigarren helburuaren emaitza lortutako bi sailkapenak dira, hauek dira: *egitura linguistikoaren sailkapena* eta *sakoneko arrazoen sailkapena*. Hauek dira § III.3-n azaldutakoak. Helburu honi buruz, esan beharra dago, mailak 9., 10., 11. eta 12.-EGA izan zirela aztertu genituenak, hasieran 7., 8. eta 9. mailak aztertzekeo asmoa genuen arren. Mailak aldatzeko arrazoiak, besteak beste, gehienbat egoera errealak markatu zituen. ILAZKI euskaltegian genituen taldeak kontuan izanik, errazagoa zitzaigun maila hauek aztertzea. Gainera, alde psikolinguistikotik euskara-mailan gutxien aztertutako mailak zirenez, interesgarriago ikusi genuen aldaketa hau (HABEren programaketa ikusi besterik ez dago. Programaketa goi mailetan agertzen diren programak oso orokorrak dira eta gutxi zehaztuta).

3) Hizkuntza baten ikaste-prozesuan ikasleek erabiltzen dituzten ikaste-estrategia konkretu batzuen azterketa eta hauen eta desbideratzeen arteko erlazioa zehaztea.

Ikaste-estrategiak desbideratzeak egitearen sakoneko arrazoiak izaten dira. Beraz, sakoneko arrazoiaren sailkapenean koka dezakegu ikaste-estrategien azterketa (ikus C eranskina).

III.4.2.2 Metodologiaren zehaztapena.

Lan esperimentala aurrera eramateko metodologia hasiera batean finkatu genuen arren, urtean zehar, aldaketaren bat egin behar izan genuen. Edozein kasutan hasiera batean definitutako metodoa ia bere osotasunean bete genuen.

Hamabi hilabetez, hamabost eguneroko maiztasunez bilera diziplinartekoa egin genuen. Esperientzia honetan ILAZKI euskaltegiko bi irakaslek, informatikariek, eta zenbait kasutan linguistek ere parte hartu zuten.

Landa-lana hiru fase desberdinetan banatu genuen:

1) Lehenengo fasea

Aztertu beharreko egitura linguistikoak aukeratzeko abiapuntua (Maritxalar et al., 96)-n agertzen den azterketa bibliografikoa izan genuen. Azterketa bibliografiko honetan maila altuko ikasleek egiten dituzten desbideratzeen lehen sailkapena aurkitzen dugu.

Aipatutako sailkapena oinarritzat hartuz, ILAZKIn maila altuetan aritzen diren bi irakaslek ikasleen corpus bat aztertu zuten (§ III.3.3.1) hasiera batean, maila horietan ILAZKIn jarraitzen ziren programak kontuan izanik. Lan honetatik aztertu beharreko egitura linguistikoaren lehen zerrenda atera zen. Zerrenda honi ondoko bi filtroak pasa genizkion: lehenengoa, linguistikoa — UPV/EHUko Informatika Fakultatean lanean aritzen diren linguista konputazionalak aipatutako zerrendatik zenbait fenomeno baztertu zuten, gaur egun euskara-mailan oraindik normalizatu gabe egoteagatik —; bigarrena, informatikoa, honetan konputazionalki arazo konpongaitza edo trataezina duen zenbait fenomeno baztertu genuen; honetaz gain, gaur egun euskara arloan linguistika konputazionalean garaturik dauden tresnen berrerabilgarritasuna ere irizpide izan genuen.

2) Bigarren fasea

Hamabost ikasle desberdinek, 9., 10-11. eta 12. mailetan zeudenek, astero idatzitako idazlanen corpora osatu genuen. Fase honetan ondokoa egin zen:

- Egitura linguistikoaren identifikazioa.

- Egitura linguistikoen erabileraren arrazoen azterketa: irakaslearen zein ikaslearen iritzia.
- Egitura linguistikoen erabileraren arrazoen bi sailkapen desberdinen definizioa, hau da, irakasle bakoitzak lortu zuena.
- Corpusaren kodeketa irakasle bakoitzak bere ikuspuntuari jarraituz.

Hasierako helburua zen bi irakaslek ikuspegi bera jarraitzea sakoneko arrazoen azterketa honetan, hala ere, lana aurreratu ahala aldaketa bat egin genuen irakasle bakoitzak ikuspegi desberdina hartzen zuela ikusita: batek hasieran planteatutako ikuspuntua jarraitu zuen, hau da, ikaslearen izaera, egoera, ezaugarriekin... lotuta zeuden arrazoiak aztertzea; besteak, ordea, testu berean agertzen ziren beste fenomeno linguistiko batzuekin erlazionatzen zituen esaldien desbideratzeak. Azken ikuspuntu hau interesgarria zela ikusita hasierako planak aldatu genituen: irakasle bakoitzak bere ikuspuntua jarraituko zuen *eta hirugarren fasean bigarren buelta eginez, irakasle bakoitzak zenbait testu berriro aztertuko zuen beste irakasleak jarraitutako ikuspegiari jarraituz.*

3) Hirugarren fasea

Azken fase honetan, hasiera batean hiru pausu finkatu arren, azterketa aurrera joan ahala beste bat gehitu behar izan genuen hirugarren fase honen hasieran. Beraz, une honetan egindakoa ondokoa izan zen:

- Corpusaren kodeketa irakasle bakoitzak beste irakasleak bigarren fasean jarraitutako ikuspegia landuz.
- Azterketan zehar egindako lanaren ondorioak.
- Corpusaren kodeketan identifikatutako zenbait arrazoiren egiaztapena burutzeko froga berezia ikasle desberdinekin.
- Bukaerako froga-azterketaren azken ondorioak.

III.4.2.3 Azterketarako aukeratutako ikasleak zehazteko irizpideak.

Ikasleen aukeraketa egitean hiru irizpide nagusi jarraitu genituen:

- A) Maila altuak, 9. mailatik aurrera.
- B) Ikaslearen hasierako konpromezua ikastarora erregularitasunez etorri eta idazlanak egiteko.
- C) Ikaste-prozesu desberdinak, motibazio anitz, hizkuntz ingurugiro diferenteak, ezaugarri pertsonal ezberdinak... hautatu genituen; azken finean perfila bera jarraitzen zutenak baztertzeko saiatu ginen, beti ere maila altuetan zeuden ikasleen artean. Hautaketa honetan konkretuki ondoko ezaugarriak eduki genituen kontutan:
 - a1) Euskara-maila altuko maila desberdinak.

- c2) Maila bera izanik ikaste-prozesuaren abiadura desberdinak (adibidez, 4 urtetan 10. mailara heldu den ikaslea eta 10. mailara heltzeko 12 urte behar izan duena).
- c3) Ingurugiroa, bai euskalduna, bai erdalduna.
- c4) Bai egoera sozialak behartutako ikaste-prozesua, baita bere inizatibaz sortutakoa ere.
- c5) Euskaldun berriak eta euskaldun zaharrak.
- c6) Euskara bigarren hizkuntza eta euskara hirugarren zein laugarren, bosgarren... hizkuntza.
- c7) Ikasten den hizkuntzaren erabilera.
- c8) Gramatikari arreta handia jartzea edo esanahiari lehentasuna ematea.

III.4.2.4 Egitura linguistikoaren aukeraketaren irizpideak.

Aztertu beharreko egitura linguistikoaren aukeraketa egitean hasiera batean egindako zerrenda ondoko lau irizpideei jarraituz burutu genuen:

- A) Maila altuetan, 9. mailatik aurrera, erabiltzen diren egitura linguistikoak.
- B) Ikaste-prozesu desberdinak, motibazio anitz, hizkuntz ingurugiro diferenteak, ezaugarri pertsonal ezberdinak... izaten duten ikasleen tartehizkuntza, azken finean aniztasunaren bidez ikasle konkretu batek bakarrik erabilitako egitura linguistikoak baztertu nahi genituen, orokorrenak aukeratuz.
- C) (Maritxalar et al., 96)-n egindako azterketa bibliografikoa.
- D) Aztertu beharreko mailetan irakasleek dagoeneko duten esperientzia: ikasleek egiten dituzten "akats" tipikoak; maizegi erabiltzen dituztenak eta ekiditen dituztenak.

Lehenengo egitura-zerrenda osatu ondoren orrazketa egin genuen eta egitura batzuk aukeratu genituen. Aukera hori ondoko irizpideen arteko oreka lortuz egin genuen:

- ILAZKI euskaltegiak finkatu zituen ikertu beharreko egitura linguistikoak, maila horietan euskararen irakaskuntzan behin eta berriro *ez-ikaste* arazoak sortzen dituztenak.
- Alde linguistikotik nahiko zehaztuta dauden egiturak, hau da, egitura normalizatuak (kontuan izan euskararen normalizazio-prozesua 1968an martxan jarri zela).
- Alde informatikotik tratamendu posiblea zuten egiturak.

III.5 Ondorioa: sistema erdiautomatiko baten beharra.

Egindako azterketa bibliografikoa eta burututako landa-lanaren ondorio gisa esan dezakegu bigarren hizkuntzaren ikaste-prozesuaren arazo nagusienetako bat ikaslearen tartehizkuntzaren gramatika nolakoa den eta gramatika hori nola jasotzen den argitzean datzala. Honegatik, diseinatuko dugun sistemak:

- Ikaslearen tartehizkuntza errepresentatuta edukiko du.
- Ikasleek erabiltzen duten zenbait ikaste-estrategia ere errepresentatuta edukiko du.
- Zenbait ikasleren desbideratze fosilizatu eta ez-fosilizatuak detektatu eta tratatuko ditu.
- Tartehizkuntzaren osaketa-prozesuan zenbait ikaste-estrategiek duten eragina islatuko du.

Ingurune informatiko honetan bi tresna-mota edo topatuko ditugu: irakaslea eta psikolinguistaren ezagutza eskuratzekoa eta ikaste-prozesuaren azterketarakoa.

- Irakaslea eta Psikolinguistaren Ezagutza Eskuratzeko Sistemak kasuan bai irakasleak bai psikolinguistak ikaslearen tartehizkuntza zein ikaste-prozesuari buruz duen ezagutza jasoko da, beti ere aurretik aipatutako taxonomiak oinarritzat hartuz.
- Ikaste-Prozesuaren Azterketarako Sistemak ikaslearekin lan egingo du, Ezagutza Eskuratzeko Sistemak bidez hartutako hipotesiak oinarritzat hartuz.

Ikaslearen saioei buruzko informazioa bildurik psikolinguistak emaitzak aztertzeko aukera izango du. Behin azterketa eginez gero, emaitzen arabera hipotesi berriak formulatu eta Ezagutza Eskuratzeko Sistemaren bidez gordeko dira. Diseinatuko dugun sistemaren arkitektura, bai hizkuntzaren ezagutzarekiko, bai ikaste-estrategiekiko independentea izango da. Honegatik, psikolinguistak bere hipotesiak aldatu arren, ez du behar informatikariaren laguntzarik, bere aldaketek sistemaren inplementazioan ez baitute eraginik izango.

Ezagutza eskuratzeko sistemak

Ikaste-Prozesuaren Azterketarako Sistema eraikitze behar dugun informazioa eduki ahal izateko, ezagutza eskuratzeko bi sistema erabiliko ditugu. Sistema hauek, irakasle eta psikolinguistarekin egingo dute lan. Hauen helburuak honakoak izango dira:

- 1) Ikasleari buruzko ezaugarriak zein erabiltzen dituen ikaste-estrategiak eskuratzeko.
- 2) Ikaslearen testuetatik abiatuz bere tartehizkuntza osatzea.
- 3) Irakaslearen ustez desbideratzeen arrazoi posibleak zerrendatzea.
- 4) Irakasleak esandakoaren aurrean psikolinguistaren iritzia jasotzea.

Galdera asko etortzen zaizkigu burura helburu hauei buruz, esate baterako ea ikasleak egindako desbideratzeen azpian arrazoi bat baino gehiago egon daitekeen ala ez; eta abar. Honi buruz, hurrengo kapituluaren zehatzago mintzatuko gara.

Ikaste-prozesuaren azterketarako sistema

Ikasteko sistema laguntzailean ikaslearen ikaste-prozesuan lagungarri gertatuko litzatekeen tresnak inplementatuko ditugu. Tresna hauen ezaugarri nagusienak ondokoak izango dira:

- 1) Ez da inolaz ere irakaslearen ordezkari izango.
- 2) Ikaslearen ikaste-prozesurako laguntza izango da.
- 3) Ikasleak egindako desbideratzeetan oinarrituta egongo da.
- 4) Laguntza gidatua emango zaio ikasleari, eta ez laguntza inposatua.

Laguntza ez inposatua esaten dugunean ondokoa esan nahi dugu: irakasle batek bere ikasleekin bi filosofia zeharo desberdinak jarrai ditzake; bata, irakasleak nahi dituen irakaste-estrategiak erabiltzen saiatzea, bestea, ikasleari askatasuna ematea ikasleak nahi dituen ikaste-estrategiak erabil ditzan. Dena den, bigarrenak ez du esan nahi irakasleak prozesua gida ez dezakeenik. Gure filosofia bigarrena izango da.

Sistema informatikoa irakaslearen ordezkari ez dela izango esan dugu, beraz, ez dugu pentsatu behar hizkuntza bat ikasteko metodo autonomoa denik. Neurri batean autonomia izan arren, ikaslea eta ordenagailua izango baitira partehartzaileak, hizkuntzaren ikaste-prozesuan osagarria izatea da gure helburua. Nagusiki taldean gertatzen den ikaste-prozesu bat ez da zeharo osatua izaten. Askotan ikasleak bere ikaskuntza-mailan komunak ez diren desbideratzeak egiten ditu. Irakasleak desbideratze horiek zuzenduko balizkio talde-dinamika aldatuko luke askotan. Hortaz, gertatzen den hutsune hau betetzeari zuzenduta egongo da gure lana. Ordenagailua lagungarria izan daiteke ikasle batek egiten dituen desbideratze horien azpian dauden sakoneko arrazoen tratamendurako, eta, bestalde, oso lagungarria ikaslearengan bere ikaste-prozesuan zehar fosilizatutako egitura ez-zuzenen desfosilizazioa nola burutzen den aztertzeke.

IV. Tartehizkuntza modelizatzeko sortu eta egokitutako tresna linguistiko-konputazionalak.

Hirugarren kapituluaren MUGARRI (bigarren hizkuntzako ikasleen ezagutza eskuratzeko sistema anitzeko ingurunea) garatu aurretiko azterketa bibliografikoa eta burututako landa-lana azaldu dugu. Arloaren beharrak identifikatutakoan MUGARRI ingurunean HITES sistemaren diseinuari ekin genion. Azken honek psikolinguistarekin elkarreagiten du, honek ikasleen hizkuntz ezagutzaren, edo tartehizkuntzaren, eskuratzeko-prozesuan laguntzen baitio sistemari.

Orain dagokigun kapituluaren, HITES sistemaren modulu automatikoaz arituko gara, konkretuki ikasleen testuez osatutako corpusaren azterketa burutzen duenaz. V. kapituluaren ordea, corpusaren azterketatik lortzen den informazioa erabiliz tartehizkuntz ereduak nola osatzen diren azaltzeari ekingo diogu, han psikolinguistarekiko elkarrekintza ere komentatuko dugularik.

Esan bezala, HITES sistemak tartehizkuntzaren errepresentazioa du helburu nagusia, eta honetarako lehen pausua da ikasleen corpusetatik abiatuta tartehizkuntzaren lehen bertsioa osatzea, kanpoko elkarreragileen laguntzarik gabe. Corpusaren azterketan tartehizkuntza modelizatzeko hainbat tresna linguistiko-konputazional egokitu eta sortu behar izan dugu. Aurrera baino lehen, aipatu behar da, ordea, kapitulu honetan azalduko dugun zenbait tresna erabilgarri dela beste zeregin batzuetan ere, hots, ikaslearen ikaste-prozesuan laguntzeko eta bai irakasleari bai psikolinguistari beren interfazeetan tresna lagungarriak eskaintzeko.

Aipa ditzagun, bada, ondorengo ataletan topatuko ditugun gaien ildo nagusiak:

Lehenik, domeinuaren adierazpena deskribatuko dugu, hau baita lortu beharreko azken helburua; hortaz, kapitulu honetan corpusaren azterketa azaldu arren, sistema osoaren bukaerako irteera deskribatuko dugu. Tartehizkuntz ereduaren deskribapenaz gain eredu bera gobernatuko duten printzipio nagusiak azalduko ditugu: dinamismoa, aldakortasuna eta finkotasuna, konpilagarritasuna, ez-monotonia eta ikuspegi aniztasuna. Honela, emaitzak zein behar lukeen ezagututa, hurrengo bi kapituluetan deskribatuko diren sistemen helburuez jabetzeko moduan izango gara.

Bigarrenik, ikaskuntza-maila bereko ikasleek amankomunean duten tartehizkuntz ereduaren osaketa-prozesurako garatu dugun metodologia proposatzen dugu.

Hirugarrenik, garatutako tresna linguistiko-konputazionalak definituko ditugu, hauetako batzuk lehendik sortutako hainbat tresnaren egokitzapenak direla: berrerabilgarritasuna azpimarratu beharreko ezaugarria dugu gure lanean. Bereziki, egokitutako analizatzailea azaltzeko atala izango dugu bukaeran.

Eta, azkenik, inplementazio-egoera zertan den komentatuko dugu: inplementazioaz aritzean batez ere hedapenak eta tresnen zabalkuntzak komentatuko dira.

IV.1 Tartehizkuntzaren modelizazioa corpusetatik.

Lehen esan bezala hizkuntza bat ikasten duten ikasleen tartehizkuntza osatzen duten egitura linguistikoen detekzioa da gure nahia. Horretarako, ezinbestekoa da ikasle, irakasle zein psikolinguistaren elkarrekintza sistemarekin. Hala ere, elkarrekintzarik izan gabe zenbait tartehizkuntz egituraren detekzioa egin daiteke ikasleek idatzitako testuetatik soilik. Jakina, kasu askotan hautagai gisa besterik ezingo ditugu proposatu corpusetatik detektatutako egiturak.

Hizkuntza bat ikasten dutenen tartehizkuntza modelizatzeko egitura linguistikoen detekzioaz gain hauen kontestualizazioa beharrezkoa dugu (Selinker, 92). Honetarako, aurreneko pausu gisa, corpusen azterketaren bidez lortuko dugu ikasle bat edo ikasle-multzo baten idatzietatik ondoriozta daitezkeen egitura kontestualizatu posibleak (Díaz et al., 98a). Lengoia naturalaren prozesamenduan egun dauden teknikak erabiliz beteko dugu gure helburua, hau da, tartehizkuntzaren aurreneko bertsioa automatikoa izatea.

LNPko teknikak, hala ere, ez dute detektatzerik gizaki batek ikus ditzakeen egitura guztiak; hauen maiztasunari buruzko informazioa, ordea, zehatzagoa izango dute gizakiek baino. Gainera, gaur egun semantika arloan hainbat aurrerapen suertatu den arren, oraindik ere, ezagutza eremua beti mugatua izaten da. Guk edozein gairi buruzko idazlanak aztergai ditugunez, semantika baztertu eta morfologia eta sintaxiaren mailan kokatu dugu gure lana. Baina, mugek ez dute hor bukaera; esperimenterako erabilitako hizkuntza euskara izanik, bi murriztapen oso garrantzitsu topatzen ditugu: bat, euskararen normalakuntza-prozesua garatzen den bitartean zenbait kasutan ez da jakiten zein den hobesten den egitura linguistikoa, eta, bi, LNPan egun dauden ikerguneetan oso urriak dira euskara bera lantzen dutenak.

Dena den, aurretik aipatutako mugak izango ez bagenu ere, testu hutsetatik egindako azterketetan testuei buruzko zenbait datu ere beharrezkoa litzateke emaitza fidagarriak nahi izanez gero, adibidez, idazlana idatzi duen ikaslearen hainbat ezaugarri (ikaskuntza-maila, hizkuntza ikasteko erabiltzen dituen baliabideak, e.a.). Esan behar da, datu hauen beharra

aurreko kapituluaren azalduetako landa-lanean egin genituen elkarrizketetan agertu zela, kasu batzuetan irakasleekin, eta beste batzuetan ikasleekin aritzean.

Beraz, ez dugu ahaztu behar corpusei buruzko informazio nahikoa era egokienean prozesatuta ere, emaitzak findu egin beharko direla sistema elkarreragileen bidez, non bai psikolinguista, bai irakaslea, baita ikaslea ere arituko baitira tartehizkuntzaren osaketa-prozesua fintzen.

IV.2 Domeinu-ereduaren adierazpidea.

IV.2.1 Tartehizkuntz ereduak. Printzipioak.

HITES sisteman domeinua eskuratzea dugu helburu, eta horregatik Ezagutza Oinarritutako Sistemak (KBS edo Knowledge-Based Systems) gure sistemaren abiapuntua izango dira arkitekturari dagokionez. Has gaitzen, bada, domeinua definitzen bere deskribapenaren azalpenak eman aurretik.

HITES sistemaren domeinua bi modulutan banatzen da: ikaslearen tartehizkuntza eta ikaskuntza-maila bereko ikasleek amankomunean duten hizkuntz ezagutza. Aldi berean, azken honetan bi modulu bereizten dira:

- a) **Tartehizkuntza komuna:** ikaslearen tartehizkuntza gordetzen duen modulua izatera berekoa, hau da, ikaslearen ikuspegitik adierazitako ezagutza.
- b) **Tresnen hizkuntz ezagutza:** aurrekoaren eskuratze-prozesuan erabiltzen diren tresna linguistiko-konputazionalak erabiltzen dena, hau da, ikuspegi linguistikotik adierazitako ezagutza.

Modu honetan ikaskuntza-maila bakoitzari dagokion ezagutza gordeko duten bina modulu egongo dira. Honek esan nahi du ikaskuntza-maila bakoitzaren ezagutza egokitzen dugula, maila guztien azterketarako tresna bera erabiltzen dugularik. Tresnen hizkuntz ezagutza gordetzen duen modulua errepresentazioa tresnetan erabilitako formalismo linguistiko-konputazionalak estuki lotuta dago, nahiz eta hauen ikuspegia hizkuntza baten ikaste-prozesua aztertzen duten teoretatik oso urruti dago. Gure helburua tresnen hizkuntz ezagutzaren errepresentazioa lantzea ez denez, tartehizkuntzaren deskribapenari buruzko adierazpidea soilik aztertuko dugu. § IV.4 eta § IV.5-n tresnen ezagutzaren eguneraketa komentatuko dugu sistemak duen hainbat tresna azaltzean.

Ondoren, tartehizkuntz ereduaren ezaugarri orokorrei buruz arituko gara. Dena den, gauden kapituluaren gaia ezagutzaren adierazpena baldin bada ere, hizkuntza bat ikasten ari den ikaslearen jokabidea komentatuko dugu hainbatetan. Woolf et al.ek (1993:406) dioten bezala

ezagutzaren errepresentazioa estuki lotua dago ikaslearen ikaste-prozesuari buruzko diagnosiarekin:

Two closely related design issues are knowledge representation and diagnostic inference; that is, a system cannot diagnose what it can't represent, and should not represent what it cannot make inferences about.

Ikaslearen ezagutza modelizatzeko aurreko kapitulan azaldutako esperimenduak izan ditugu abiapuntu, eta hori dela medio definituko dugun ikasle-ereduaren hizkuntz ezagutza behar errealekara egokituko da, ahal den neurrian. Ikaste-prozesuan zehar ikasleak eskuratzen duen ezagutzaren dinamismoa landa-lanean argi geratu zen. Woolf et al.ek (1993:403) oso ondo adierazten dute ideia ondoko paragrafoan:

Observations about human learning constrain the design in a student model. We have shown that the model should reflect a student's changing knowledge, spontaneous reorganization of that knowledge and changes in its own beliefs about that knowledge. It should also generate non-persistent inferences which can later be overturned. It should continually reassess student knowledge and expect dynamic movement through stages.

Hizkuntzaren ikaste-prozesuan hizkuntz ezagutza osatzen joan ahala, ikasleak baliabide berriak eskuratzen ditu egitura berriak sortzeko ahalmena garatuz. Ahalmen hori, ordea, ez da modu berean bilakatzen ikasle guztiengan. Ikasle ausartei egitura berriak sortzea gustatzen zaie, eta sortutakoa idazlanetan, esaterako, beren erabilpena gehitzen dute ikaste-estrategia moduan. Bada, berriz, arrisku hartzerik atsegin ez duen ikaslea, honek irakatsitakora murrizten du bere eskuratze-prozesua, sorkuntza baztertuz estrategia bezala. Aipatu bi ikasle-mota horien kasuan garbi dago ikasgai bera jaso arren eskuratzen duten ezagutza desberdina dela, edo era desberdinean gordetzen dutela. Bigarrenak irakatsitakoa era konpilatuan ikasiko duen bitartean, lehenengoak unitate txikiagotan banaturik eskuratuko du ezagutza "bera". Adibidez, eman dezagun ikasleek lehen aldiz ikusten dutela kontsonantez bukatutako hitz baten deklinabidea ariketa berri batzuetan, eta adibide guztiak *non* kasuan daudela. Ikasle ziurrak ondoko egitura eskura zezakeen: *lema kontsonantez + non kasua => -ean atzizkia gehitu*. Ikasle ausartak, berriz, ondokoa egingo luke seguraski: *lema kontsonantez + kasua => -e + bokalez bukatuko balitz gehituko genukeen atzizkia*. Aztertzen baditugu bi jokaerak, sorkuntza-prozesua lantzen duen ikasleak *-ean* itema bitan banatu du, hau da, *-e + -an*. Honek esan nahi du sorkuntza garatzen denean eskuratze-prozesuan gordetzen diren egiturak ugariagoak direla, garatzen ez denean baino. Horregatik, esan dezakegu *ikasleek erabiltzen dituzten ikaste-estrategiek ezagutzaren antolakuntza baldintzatzen dutela*. Brent-en (1996:26) *Advances in the computational study of language acquisition* izeneko artikuluan sorkuntza fenomenoaren lantzen da ama-hizkuntzaren eskuratze-prozesuan. Gure kasua bigarren hizkuntzaren ikaste-prozesua izanik ere, baliagarria da *autonomous bootstrapping* izeneko fenomenoari buruz han egiten den azterketa. Bukaeran

hainbat hurbilpenen konbinaketa proposatzen du: ikaste automatikoa, inferentzia estatistikoa, LNPko teknikak eta psikologia esperimentalak. *Autonomous bootstrapping* izeneko fenomeno honela definitzen du:

Autonomous bootstrapping strategies work by representing partial and/or uncertain linguistic knowledge and using it to analyze the input. In order to develop an autonomous bootstrapping analysis, one must answer the following questions:

1. What bit of tentative, partial knowledge can be gleaned from unanalyzed inputs?
2. How can that knowledge be represented?
3. How can such tentative, partial knowledge be used to analyze other utterances more fully, and hence extract more knowledge from them?

Jakina dugunez, ama-hizkuntza eskuratzean bilakatzen den *autonomous bootstrapping* izeneko garatuagoa da beste hizkuntza bat ikaste baino. Hala ere, ezagutzaren antolakuntza partziala eta ez-ziurra bigarren hizkuntzaren ikaste-prozesuan ere agertzen dira, modu desberdinetan ikasle-motaren arabera, ezagutzaren adierazpena baldintzatuz. Partzialtasuna bigarren hizkuntza ikaste ezagutza eskuratzeko moduarekin lotua dago: ikasleak buruz ikasten duenean egitura linguistiko bat modu konpilatuan barneratzen du, egitura osatzen duten erregela linguistikoek kontziente denean, berriz, era partzialean eskuratzen du ezagutza, guk honi modu ez-konpilatua deituko diogu.

Esandakoa oinarri izanik, tartehizkuntz ereduaren gobernatuko duten printzipio nagusiak honela definituko ditugu:

a) **Dinamikoa.** Ikaste-prozesuaren izaerak definitua. Ikaskuntza-maila bereko ikasleentz tartehizkuntz ereduaren kasuan, denboran zehar alda daitekeela esan dezakegu, hau da, maila batentzat une batean definitzen den tartehizkuntz komuna desberdina izan daiteke hamar urte beranduago, hizkuntzak ikasteko programetan batzuetan ikasgaien ordena berritzen delako, eredu berriak agertuz. Beraz, honelako dinamismoak kontsideratu beharrekoak dira, tartehizkuntzaren azterketa helburutzat hartzen duen sistema batean.

b) **Aldakorra eta finkoa.** Ezagutzaren garapenak behartua. Ezagutzaren eskuratzeko prozesua ez da bat-batekoa, pixkanaka osatzen dena baizik. Hortaz, ikaste-prozesuan zehar ezagutza linguistikoaren forma batzuk aldatuko dira eta beste batzuk desagertuko dira. Horiez gain, ezagutza barneratzen den heinean forma finkoak ere gauzatuko dira. Hizkuntza baten ikaste-prozesuetan maiz gertatzen da, ikasleak modu desberdinetan erabiltzen duela fenomeno bera, esaterako, ergatibitatean *-k* markatzea batzuetan eta ez markatzea beste batzuetan. Honelako egoeretan ezagutza linguistikoa aldakorra dela esango dugu, hortaz, aldakortasunaren forma bakoitzari, aldaera bakoitzari hain zuzen, egonkortasuna adierazteko balio kualitatiboa asignatuko diogu.

c) **Konpilatua eta ez-konpilatua.** Sorkuntza-prozesuak bideratua. Aurretik komentatu ditugun bi ikasle-moten ezagutzen arteko desberdintasuna adierazteko ezagutza konpilatua eta ez-konpilatua bereiztu behar dira. Ezagutza konpilatuaren kasuan ikaslea ez da jabetzen ezagutzaren barneko egituraz, ez-konpilatuaren kasuan, ostera, ezagutzaren antolakuntza partziala ere eskuratzen du.

d) **Ez-monotonoa.** Sistemaren elkarreragile desberdinek eskatua. Arrazonamendu ez-monotonoa beharrezkoa da sisteman hasiera batean psikolinguistak ikasleen ezagutzari buruz onartutakoa geroago baliogabetua gera baitaiteke ikaslearekiko elkarrekintzan.

e) **Ikuspegi anitzekoa.** Ezagutza beraren ikuspegi desberdinak eduki ditzakegu, hots, linguistikoa eta azalekoa. Adibide gisa, ikasle batek **dakilako* desbideratzea erabiliz gero linguistikoki *-elako* atzizkiaren ordez *-lako* atzizkia erabili duela esango genuke, azaleko ikuspegitik, berriz, *-e* letra ahaztu zaiola.

IV.2.2 KADS metodologia.

Tartehizkuntz domeinuaren printzipioak aztertu ondoren eta tartehizkuntzaren eredu kontzeptuala azaldu aurretik azken honen deskribapena adierazteko metodologia komentatuko dugu labur-labur. Ereduaren deskribapen kontzeptuala sistemaren diseinutik independentea behar genuenez honi egokitzen zitzaion metodologia aukeratu genuen, KADS ("Knowledge Acquisition and Documentation Structuring") metodologia hain zuzen.

KADS metodologian (Schreiber et al., 93) ezagutza eskuratzea, funtsean, modelizazio-saioa da. Zentzu horretan, Ezagutzan Oinarritutako Sistema, ezagutza gordailu baino gehiago da, hau da, mundu errealearen fenomeno bat adierazten duen eredu operazionala dugu. Metodologia honetan bi eredu-mota agertzen zaizkigu ezagutzaren eskuratze-prozesuan: eredu kontzeptuala, sistemaren portaera-espezifikazioa, eta diseinu-eredua, portaera aurrera eramateko errepresentazio- eta konputazio-teknikak zehazten dituen. Guk, eredu kontzeptuala erabili dugu sistemaren diseinu kontzeptuala egitean, eta honi buruzko ideia orokorrak aipatuko ditugu hemen.

Adituen ezagutza modelizatzean ondo bereiztu behar da domeinu-ezagutza eta kontrol-ezagutza. Kontrolean inferentziak, atazak eta estrategiak lantzen dira; domeinuan, berriz, aplikazio partikular baten ereduaren kontzeptualizazioa errepresentatzen da. Domeinuko kontzeptuak eta horien atributuak, egitateak, egiturak eta erlazioak errepresentatzen dira. Schreiberrek (1993) domeinuaren deskribapenerako hainbat formalismo aztertzen du. KADS metodologian egoki izan daitezkeen hiru hurbilpen aztertzen ditu autoreak: entitate-erlazio eredu, entitate-erlazio eredu hedatua eta KL-ONE eredu. Alde batetik, entitate-erlazio ereduak *entitate* kontzeptua definitzen du objektuen errepresentaziorako, *erlazioa* entitateen arteko erlazioak adierazteko eta, azkenik, *atributua* eta *balioa*. Atributua entitate bat balio

batekin erlazionatzen duen funtzioa da. Beste alde batetik, entitate-erlazio eredu hedatuan datu-base semantikoen modelizazioan *is-a* erlazioa beste erlazio motetatik bereizten da. Erlazio-mota honen bidez entitateen arteko herentzia erreprezentatzen da. Eredu hedatueta, *is-a* erlazioaz gain multzoak eta agregazioak topatzen ditugu. Bestalde, atributuak moten arteko funtzioak dira. Azkenik, KL-ONE hurbilpenean ere *is-a* erlazioa beste erlazio motetatik bereizten da, herentziarako aukera emanez. Bestelako erlazioak entitatearen rolak bezala definitzen dira, eta azpikontzeptuen bidez heredatutako roletan murriztapenak ezar daitezke. Gainera, KL-ONEn kontzeptuak eta hauen instantziak desberdintzen dira. Eta egiturazko deskribapenen bidez rolen arteko erlazioak adierazten dira. Baina, bi arazo nagusi daude KL-ONE lengoian: bat, erlazioak, *is-a* erlazioa izan ezik, zeharka daude adierazita rolen bidez, eta, bi, ez dira bereizten entitateen arteko erlazioak erreprezentatzen dituzten rolak eta atributu-rolak.

DDL (Domain Description Language) izeneko lengoaiak entitate-erlazio eredu du oinarri, eta KL-ONE zein datu-base semantikoen modelizaziorako formalismoen hainbat ezaugarriren bidez aberastu da. DDLk objektu intentsionalak (intensional objects), estentsionalak (extensional objects) eta eraikuntza lagungarriak (auxiliary constructs) erabiltzen ditu modelizaziorako. Objektu intentsionalen egiturak ondokoak dira: kontzeptuak, multzoak, erlazioak eta egiturak. Objektu estentsionalak instantziak eta tuplak dira, eta hauen mota objektu intentsionalek definitzen dute. Azkenik, eraikuntza lagungarriak dira objektu intentsionalen definizioetan erabiltzen direnak: adierazpenak, azpimotak, ezaugarriak, balioak eta balio-multzoak.

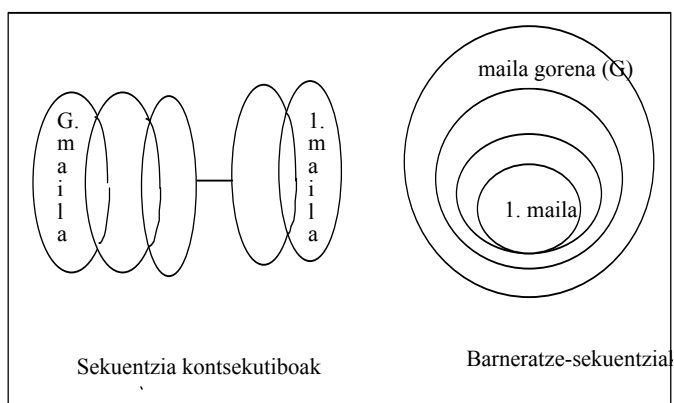
HITES sistemaren eredu kontzeptualaren definiziorako DDL lengoia hobetsi dugu. Arrazoi garrantzitsuenak, batez ere, ondokoak dira: sistemaren diseinuarekiko independentea izatea, objektuei zuzendutako ikuspegia eskaintzea eta agregazioaren aukera ematea. Sistemaren diseinuan, ordea, GBC sistema (Elorriaga et al., 95) erabili dugu. Honek CLIPS programazio-lengoaiaren objektuei zuzendutako programazio-ingurunea eskaintzen digu, besteak beste, objektuen definizioa eta objektuen arteko herentzia-erlazioak ingurune grafiko baten bidez lantzeko aukera. Baina honi buruz ez gara hemen arituko diseinu-fasean erabilitako tresna baitugu.

IV.2.3 Tartehizkuntz ereduaren deskribapena.

Atal honetan ikaskuntza-maila bati dagokion Tartehizkuntzaren Eredu Kontzeptuala (IMTEK) deskribatuko dugu. Hau eta Ikaslearen Tartehizkuntz Eredua (ITE) antzekoak izanagatik desberdintasunen bat badute, eta azalpenean zehar azpimarratuko dugu.

Ikasle bakoitzak bere hizkuntza propioa eskuratzen du ikaste-prozesuan zehar. Honetaz gain, ikaskuntza-maila bereko ikasleek mailari dagokion hizkuntz ezagutza komuna izaten dute ezagutzaren parte. Hortaz, hizkuntzaren egitura komunak izango dira IMTEKen errepresentatuko ditugunak.

Ikaskuntza-maila bakoitzean bi ezagutza-mota desberdin adierazten dira: *ezagutza aldakorra*, ikaslea une horretan ikasten ari dena, eta, *ezagutza finkoa*, dagoeneko ikasleak barneratu duena. Ikaskuntza-maila konkretu batera heltzean finkatutako hizkuntz ezagutzaren barne aurkitzen dugu beheko mailaren batean dauden ikasleek barneratutako ezagutza. Ikaste-prozesuan dagoen ezagutza, berriz, era desberdinean dago antolatuta: ikaskuntza-maila kontsekutiboan ezagutza aldakorrak elkarrekiko independenteak dira, elkarren arteko ebakidurak izaten badira ere. Hau dela eta, hizkuntz mailetan zehar hizkuntz egitura aldakorren multzoak *sekuentzia kontsekutiboak* dira, eta hizkuntz egitura finkoen multzoak, berriz, *barneratze-sekuentziak* (ikus IV.1 irudia).



IV.1 irudia. Ezagutza aldakorra eta ezagutza finkoa.

Ereduaren deskribapenari ekin aurretik gogoraziko dugu III. kapituluan tartehizkuntza kontzeptua definitzean azpimarratutako bi ezaugarri, garrantzitsuak direnak erreduaren deskribapenean.

- Lehenik, hizkuntz egitura aldakorra edo finkoa diogunean ez dugu inolako erreferentziarik egiten ezagutza zuzena edo "okerra"ri buruz. Tartehizkuntz egiturak definitzean ezagutza zuzena eta ezagutza "desbideratua" era berean daude errepresentatuta, biek osatzen baitute ikaslearen ezagutza. Ikaskuntza-maila altuko ikasleek erabilitako egitura konkretu bat desbideratutzat ematen dugun bitartean, beheko maila batean dagoen ikasle batek egitura bera erabiliz gero zuzena dela esango genuke hainbat kasutan. Adibidez, 12. mailako ikasle batek *nor ote da?* esan beharrean *nor da?* galdetzen duenean egitura desbideratua erabili duela esaten dugu, eta lehenengo mailako ikasle batek *nor da?* erabiliko balu, ordea, edozein egoeratan egitura zuzena dela esango genuke.

- Bigarrenik, tartehizkuntzaren hizkuntz egiturak testuinguruarekiko dependenteak dira (Selinker, 94):

Language learners create discourse domains as contexts for interlanguage development. Such domains are constructed in connection with life experiences that have importance for the learner, containing prototypical interlanguage forms associated with the content area by the learner. Interlanguage may be developing in one domain while at the same time it may be stabilised, or possibly fossilised, in another.

Honegatik, gure eremuan tartehizkuntz egituren adierazpenean testuinguruari buruzko informazioa topatuko dugu. Dena den, badira egiturak edozein testuingurutan gertatzen direnak, eta horiek testuinguruarekiko egitura independenteak izango dira eremuan.

Ezaugarriak ezaugarri, IMTEKren lehenengo hurbilpena ondokoa dugu (DDL erabiliz):

```
set Hizkuntz_ezagutza_mailarena;  
  element-type: instance(Hizkuntz_ezagutza_komuna);  
  cardinality: min 0 max infinite;  
  properties:  
    identifikadorea: string;  
    ikaskuntza_maila: integer-range(1, MAILA_GORENA);  
  
structure Hizkuntz_ezagutza_komuna;  
  parts:  
    ezagutza_finkoa : set(instance(Tartehizkuntz_eredu_finko_komuna));  
      cardinality: min 0 max infinite;  
    ezagutza_aldakorra: set(instance(Tartehizkuntz_eredu_aldakor_komuna));  
      cardinality: min 0 max infinite;  
  
set Tartehizkuntz_eredu_finko_komuna;  
  element-type: instance(Tartehizkuntz_eredu_komuna);  
  cardinality: min 0 max infinite;  
  
set Tartehizkuntz_eredu_aldakor_komuna;  
  element-type: instance(Tartehizkuntz_eredu_komuna);  
  cardinality: min 0 max infinite;  
  
structure Tartehizkuntz_eredu_komuna;  
  parts:  
    testuinguruan_oinarrituta: set(instance(Egitura_kontestualizatu_komuna));  
    testuinguruarekiko_independentea: set(instance(Egitura_independente_komuna));  
  
structure Egitura_kontestualizatu_komuna;  
  subtype-of: Tartehizkuntz_egitura_komuna,  
  parts:  
    baldintzak: set(instance(Baldintza));  
  
structure Egitura_independente_komuna;  
  subtype-of: Tartehizkuntz_egitura_komuna,
```

```
structure Tartehizkuntz_egitura_komuna;  
subtypes:      Egitura_kontestualizatua_komuna,  
                Egitura_independentea_komuna;  
parts:  
    fenomenoak: set(instance(Fenomeno_linguistikoa));  
properties:  
    identifikadorea: string;  
    deskribapena: string;  
    desbideratzea: boolean;  
    ikaskuntza_maila_gorena: integer;  
    mota: list(string-value); value={testua, esaldia, hitza}  
    maila_linguistikoa: list(string-value); value={morfo, sintak}  
    kontestualizatua: boolean;  
    egonkortasuna: string; function(fenomeno linguistikoen egonkortasuna);  
                    value={gutxitan, batzuetan, askotan, beti}
```

IV.2 irudia. *Hizkuntz_ezagutza_mailarena* eta *Tartehizkuntz_eredu_komuna* definizioak.

IV.2 irudian agertzen denez baldintza konkritu batzuetan gertatzen diren fenomeno linguistikoek definitzen dute tartehizkuntz egitura.

Ondoren, aipatutako bi kontzeptuak sakonago aztertuko ditugu:

- Tartehizkuntz egiturak.
- Fenomeno linguistikoak.

Beraz, fenomeno linguistikoen aurkezpenari ekin aurretik tartehizkuntz egiturak definituko ditugu.

Tartehizkuntz egiturak

Esan bezala, tartehizkuntz egitura bat fenomeno linguistikoez osatua dago. Hauetaz gain, hainbat ezaugarri topatzen dugu tartehizkuntz egitura baten barruan: dagokion *identifikadorea* eta *deskribapena*; aztertzen ari garen ikaskuntza-mailan *desbideratzea* kontsideratzen den; *ikaskuntza_maila_gorena*, hau da, ezagutza finkoari badagokio zein ikaskuntza-mailatan finkatzen den egitura, eta ezagutza aldakorari egokituz gero zein ikaskuntza-mailatik aurrera mantendu den ezagutza aldakorrean; *mota*: testu, esaldi edo hitz-mailako egitura den; *maila_linguistikoa*: morfologikoa, morfosintaktikoa edo sintaktikoa; *kontestualizatua* testuinguruarekiko dependentea den ala ez adierazteko; eta, azkenik, *egonkortasuna*.

Aipatutako ezaugarrien artean *desbideratzea* ez da agertuko ikasleen hizkuntza propioa osatzen duten tartehizkuntz egituretan, hau da, ikaskuntza-mailarekiko komuna ez den ezagutza estandarra dena bakarrik adieraziko dugu. Honen arrazoi nagusia ondokoa da: MUGARRIren arkitekturan ikaskuntza-maila bakoitzeko diagnosi-modulua proposatzen

dugu, eta ez ikasle bakoitzekoa. Gainera, lehen ikusitako adibidean (*nor ote da?* esan ordez *nor da?* erabili) argi ikusi da desbideratze kontzeptua ikaskuntza-mailarekin lotu dugula.

Mota eta *maila_linguistikoa* ezaugarriak kontuan izanik identifika dezakegu ikaslearen aberastasuna maila desberdinetan, hau da morfologian, morfosintaxian... Honetaz gain, ez dugu ahaztu behar proposatzen dugun eredua ikaste-sistema batean ere erabiliko dugula. Kasu honetan egoera bakoitzean aztertzen den tartehizkuntz egiturak *mota* eta *maila_linguistikoa* eremuetan dituen balioen arabera tresna desberdinak aktibatuko dira ikasleari ikaste-prozesuan laguntzeko. Definizioan ikusten denez, tartehizkuntz egitura fenomeno desberdinek osatzen dute, hitz-mailakoak eta esaldi-mailakoak nahastuta egon daitezkeelarik. Adibidez, ikasle batek ondoko esaldi-mota maiz erabiltzen duenean: *Esan du aterako dela eta zinemara joango da*, koordinazioa eta subordinazioa fenomenoak batera biltzen ditu tartehizkuntz egitura berean (tartehizkuntz egitura: *eta* juntagailuaren bidez bi esaldi elkartzean *-ela* atzizkia lehenengo esaldiaren aditzari bakarrik gehitzen dio). Kasu honetan esaldi-mailako analizatzailea eta hitz-mailako segmentatzailea, konkretuki aditza analizatzekoa, aktibatu beharko lirateke egitura hau aztertzeko.

Tartehizkuntz egitura fenomeno-multzoak osatzen duen bezala, fenomeno linguistikoa erregela-multzoak osatzen du; adibidez, aditzen azpikategorizazioa fenomenoan ikasleak aditzaren erregela sintaktikoak eta semantikoak (adib. aditzaren osagaiak bizidun izatea) batera biltzen ditu. Gure azterketan oraingoz erregela semantikoak baztertu ditugun arren, etorkizunean garatu beharreko arloa dela iruditzen zaigu. Edozein kasutan, fenomenoaren definizioari aurrerago ekingo diogunez, jarrai dezagun tartehizkuntz egituraren definizioa azaltzen.

Tartehizkuntz egituren egonkortasuna

Egonkortasuna ikaskuntza-maila konkretu batean tartehizkuntz egituraren barneratze-gradua neurtzen duen balio kualitatiboa dugu. Egonkortasuna ezaugarriaren bidez hizkuntz egitura ezagutza finkoa edo aldakorrari dagokion zehazten da. Ezaugarri honen balio posibleak ondokoak dira: *gutxitan*, *batzuetan*, *askotan* eta *beti*. Esate baterako, ezaugarriaren balioa *beti* denean ikasleak egitura barneratu duela esan nahi du, honelako kasuetan egitura garatzea ia ezinezkotzat ematen da. Literaturan egitura desbideratua denean eta zeharo barneratua dagoenean fosilizazio fenomeno dela esan ohi da. Balioa *gutxitan* denean hizkuntz egituraren aldakortasuna handia izaten da, kasu hauetan irakasleak (edo OLHIA sistema batek) ikaslea bidera dezake egitura barneratzeko edo aldatzeko xede-hizkuntzaren egituretatik urruti balego.

Tartehizkuntz egituren baldintzak

Orain arte emandako deskribapenean ikusi dugu tartehizkuntz egiturak hainbat baldintzatan gertatzen diren fenomeno linguistikoek osatzen dituztela. Beraz, alde batetik, zein motatako baldintzak aurkituko ditugun, azaltzeari ekingo diogu.

structure Baldintza;

parts:

testuinguru_linguistikoa: **set(instance(Baldintza_linguistikoa));**

testuinguru_ez_linguistikoa: **set(instance(Baldintza_ez_linguistikoa));**

IV.3 irudia. Baldintza egituraren definizioa.

Zer baldintza-mota izan behar dugu kontuan tartehizkuntz egituren definizioan? Arestian aipatu dugu tartehizkuntz egitura batzuk testuinguruarekiko dependente eta beste batzuk independente direla (Selinker, 92). Testuinguruarekiko dependenteak diren kasuetan bi baldintza-mota bereizten ditugu: linguistikoak eta ez-linguistikoak (IV.3 irudia).

Baldintza ez-linguistikoek (IV.4 irudia) diskurtsoaren domeinuari egiten diote erreferentzia. Tartehizkuntzak diskurtsoaren arlo bakoitzean garapen propioa du. Ikasleak istorio berria asmatzen ari denean aktibatzen dituen hizkuntz egiturak ez dira testu zientifiko-teknikoa idaztean erabiltzen dituen berberak (*baldintza_tematikoa*). Ariketa-mota ere, esaterako, galdera-erantzunak, eskutitz bat idatzi, hutsuneak bete, itzulpena, e.a. baldintza ez-linguistikoa da. Tartehizkuntz formak askotan ariketa-mota batetik bestera aldatzen dira, diskurtsoaren gaia aldatzen ez bada ere. Horretaz gain, baldintzak testu osoari egokitzen zaizkio batzuetan eta egitura txikiago bati beste batzuetan, konkretuki esaldiari edo hitzari (*egiturazko_baldintza*). Adibide gisa, esaldiaren luzerak eragina izaten du aditza eta osagaien arteko komunztadura markatzean: ikasleen testuetan esaldi motzetan komunztadura egokia topatzen dugu maizago luzeetan baino. Mota honetako baldintzak corpusaren azterketan landu ditugu. Testu osoari dagozkien baldintzak, ordea, ikasle eta irakasleekiko elkarrizketetan detektatu dira.

```

concept Baldintza_linguistikoa;
properties:
  hitz_maila: list(string-value); value={kategoria_gramatikala, deklinabide_kasua...}
  esaldi_maila: list(string-value); value={subordinatuen_erabilera, posposizioak...}

concept Baldintza_ez_linguistikoa;
subtypes: Egiturazko_baldintza, Baldintza_tematikoa, Ariketa_baldintza;

concept Egiturazko_baldintza;
subtype-of: Baldintza_ez_linguistikoa;
properties:
  mota: list(string-value); value={testua, esaldia, hitza}
  balioa: list(string-value1) | list(string-value2) | list(string-value3);
  value1={laburpena, eskutitz formala, itzulpena...}
  value2={esaldi_luzea ...}
  value3={hitz_luzea, aplikazioa_leman, aplikazioa_morfeman...}

concept Baldintza_tematikoa;
subtype-of: Baldintza_ez_linguistikoa;
properties:
  mota: string; value={orokorra, tekniko, ...}

concept Ariketa_baldintza;
subtype-of: Baldintza_ez_linguistikoa;
properties:
  balioa: string; value={...}

```

IV.4 irudia. Baldintza-mota kontzeptuen definizioa.

Baldintza linguistikoak corpuseko azterketa bidez landu dira. Analizatutakoen artean interesgarri iruditzen zaigu esaldiaren osagaien pluraltasunaren arazoa aipatzea. Zenbaitetan esaldiaren osagaiak pluralean agertzeak aditza ere pluralean jartzea ekartzen du, nahiz eta osagaiak ez izan inolako komunztadurarik aditzarekin. Fenomeno bera frantsesa ikasten dutenengan ere detektatu dute Lessard et al.ek (1994:457).

In the case of subject-verb number agreement, we showed that there was again a major strategy at work which tended to produce the unmarked singular in cases where other non-number-marked elements intervened between subject and verb, and the marked plural form where such intervening elements were plural.

Fenomeno linguistikoak

Behin tartehizkuntz egiturak azalduta fenomeno linguistikoaren deskribapenari ekingo diogu. Fenomenoaren adierazpenean nagusiki ondoko ekarpenak nabarmentzen dira: fenomeno beraren ikuspegi desberdinak, fenomenoaren konpilagarritasuna (erregela konpilatu eta ez-konpilatuak) eta fenomenoaren aldakortasuna. Hauei buruz arituko gara ondoko lerrootan:

Fenomenoen ikuspegi linguistikoa eta azalekoa

§ IV.2.1-n tartehizkuntza ikuspegi anitzekoa dela definitu dugu. Hau dela eta, fenomeno linguistikoak eta hauek osatzen dituzten fenomeno-aldaerak definitu dira domeinuaren deskribapena egiterakoan (IV.5 irudia). Fenomeno linguistikoa, adibidez, **dakilako* desbideratzea erabiltzea izan daiteke. Desbideratzea, ordea, bi ikuspegi desberdinetatik azal dezakegu: bat, -e letra gehitzea ahaztu zaio ikasleari, azaleko ikuspegia, bi, *-elako* atzizki kausala erabili beharrean *-lako* atzizkia erabili du, ikuspegi linguistikoa.

```
structure Fenomeno_linguistikoa;  
  parts:  
    azaleko_ikuspegia: set(instance(Fenomeno_aldaera));  
    ikuspegi_linguistikoa: set(instance(Fenomeno_aldaera));  
  properties:  
    identifikadorea: string;  
    fenomeno_mota: string; value={...}
```

IV.5 irudia. Fenomeno_linguistikoa egituraren definizioa.

Domeinuan, beraz, bi fenomeno ikuspegiak agertuko zaizkigu. Dena den, fenomeno guztietan ez ditugu bi ikuspegiak topatzen, eta, bestalde, kasu batzuetan, ikuspegi berean interpretazio bat baino gehiago izan dezakegu.

<u>Adibidez:</u> <i>*institutuako</i>	(=> <i>institutuko</i>) (lokatiboa)
<u>Bi interpretazio azalekoak:</u>	<u>Bi interpretazio linguistikoak:</u>
1) *LEGBA (A LEtra Gehitu lemaren Bukaeran)	1) <i>institutoa</i> lema + -ko morfema
2) *MOG-A (A MORfema Gehitu)	2) <i>instituto</i> lema +-a morfema + -ko morfema

IV.6 irudia. *institutoko* hitzaren ikuspegi azalekoak eta linguistikoak.

Tartehizkuntzaren sistema linguistikoaren aldakortasuna kontuan izanik, hizkuntz irakasle adituen ustetan ikasle berak bi lema, hots *instituto* eta *institutoa*, eduki ditzake aldi berean bere hiztegian (IV.6 irudia).

Fenomenoen konpilagarritasuna

Fenomenoa definitzean ikasleak eskuratutako erregelak agertzen dira, bai azaleko erregelak, baita linguistikoak ere. Hala ere, ikasleak ez ditu beti fenomenoaren osatzen duten erregela guztiak eskuratzen. Esaterako, aditza eta esaldiaren hainbat osagaien arteko komunztadura fenomenoaren jabe izan daiteke, fenomenoaren existentzia besterik ez jakinik. Aurrerago, bere ikaste-prozesuan zehar fenomenoaren barruko erregelak ikasiz joango da. Honegatik fenomenoaren erregelak esplizituki adieraztea beharrezkoa deritzogu. Aipatutako arrazoia ez da alabaina bakarrik erregelak esplizituki azaltzeko. Tartehizkuntz ereduaren printzipioen artean konpilagarritasuna ezaugarria aipatu dugu, hau da, fenomenoaren eskuratzean ikasleak batzuetan fenomenoaren barneko egitura aztertzen du, fenomenoaren osatzen duten erregelak ere eskuratuz, beste batzuetan, ordea, ikasleak era konpilatuan ikasten du fenomenoaren: ez du lantzen barneko egitura eta ez du erregularik inferitzen. Honen ondorioz, erregelak esplizituki behar izateaz gain *konpilatua* ezaugarria ere agertzen zaigu definizioan. Dena den, konpilatua ezaugarria fenomeno-aldarekin lotua dago, eta ez fenomeno linguistikoarekin. Hau dela eta, fenomeno-aldara definitzeari ekingo diogu.

Fenomenoen aldakortasuna

§ IV.2.1-n esan bezala, tartehizkuntz egitura bat berez fenomeno-aldarez osaturik dago, hauek fenomeno linguistikoaren bai azaleko ikuspegiak bai ikuspegi linguistikoak izan daitezkeelarik. Adibidez, komunztadura fenomenoaren izan dezakegu aztergai, baina fenomeno horren aldara asko topa ditzakegu ikasleengan. Horregatik, IV.7 irudian fenomeno-aldara definitzeari ekin diogu.

```
structure Fenomeno_aldara;
  parts:
    erregelak: set(instance(Erregela_linguistikoa)) | set(tuple(Ordezkapen_erregela);
    baldintzak: set(instance(Baldintza));
  properties:
    identifikadorea: string;
    deskribapena: string;
    konpilatua: boolean;
    mota: list(string-value); value={testua, esaldia, hitza}
    maila_linguistikoa: list(string-value); function(erregelen maila linguistikoa);
    value={morfo,sintak}
    adibidea: list(string-value);
  axioms:
    baldintzak= Ebakidura(erregelen baldintzak)
```

IV.7 irudia. Fenomeno_aldara egituraren definizioa.

Irudian agertzen denez aldaera baten erregelak baldintza morfologiko edo sintaktiko batzuetan gertatzen dira. Horretaz gain, aldaerak hiru motatakoak izan daitezke, burututako esperimientuetan irakasleek hiru aldaera-mota nagusi identifikatu dituztelako erregelei erreparatzean. Aldaera-motak honako hauek dira:

- Erregela linguistikoez osatutakoak.
- Erregela konkretuak ebitatzen direnekoak (adib. *nor da?* erabili *nor ote da?* erabili beharrean).
- Erregela linguistiko batzuen erabileraren maiztasun handikoak beste erregela egokiago batzuk saihesten baitira (adib. *ordea, hala ere, hortaz...* esan ordez *eta* juntagailua erabili).

Sailkapen honek adierazten du tartehizkuntzaren erregela linguistikoen artean erlazioak sortzen direla ezagutzaren eskuratze-prozesuan. Honelako erlazioak *Ordezkapen_erregela* egituraren bidez definitu ditugu (ikus IV.8 irudia).

```
structure Erregela
  subtypes: Erregela_linguistikoa, Ordezkapen_erregela;
  parts:
    baldintzak: set(instance(Baldintza));
  properties:
    identifikadorea: string;
    deskribapena: string;
    mota: string; value={testua, esaldia, hitza}
    maila_linguistikoa: list(string-value); value={morfo, sintak}

structure Erregela_linguistikoa;
  subtype-of: Erregela
  subtypes: Erregela_morfologikoa, Erregela_sintaktikoa;
  parts:
    implementatuta: list(string-value); value={erregela identifikadoreen zerrenda}

structure Ordezkapen_erregela;
  subtype-of: Erregela
  subtypes: Erregela_morfologikoa, Erregela_sintaktikoa;
  parts:
    zein_erregela: instance(Erregela_linguistikoa);
    zein_erregelaren_ordez: instance(Erregela_linguistikoa);
  axioms:
    zein_erregelaren_ordez = hutsa erregela bat ebitatzean osagaien bat kendu besterik egiten ez duenean, hau da, ez du ordezkorik erabiltzen.
```

IV.8 irudia. Erregela egituraren definizioa.

Fenomeno_linguistikoa motako egituran (ikus IV.5 irudia) bi identifikadore daude: dagokion *identifikadorea* eta *fenomeno_mota*, lehenengoak fenomenoaren portaera partikularra identifikatzen du eta bigarrenak fenomeno bera. Alde batetik, identifikadoreak fenomenoaren aldaera bat identifikatzen du. Identifikadoreak ikaste-denboran zeharreko

posizioa ere adierazten du, hau da, *FEN8* fenomenoa ikaste-prozesuan zehar *FEN15* fenomenoa baino lehen agertu da -. Ordea, informazio eza dela eta, denboran kokaturik ez dauden fenomenoak, detektu ahala zenbakitzen dira: *F1*, *F2*, *F3* ... -. Fenomeno_motak, beste alde batetik, esanahidun kodea du, fenomenoa bera identifikatzen baitu, adibidez, KOMSA (Subjektuaren eta Aditzaren arteko KOMunztadura). Dena den, identifikadore hauek ez dira nahiko. Tartehizkuntzaren printzipioak azaltzean finkoa izateaz gain aldakorra ere badela esan dugu; beste modu batean esanda, fenomeno bereko (*fenomeno_mota*) hainbat portaera desberdin topa dezakegu ikaslearen tartehizkuntzan; portaera horiek ezaugarri berak (properties) eduki ditzakete kasu bakoitzean fenomeno osatzen duten erregelak desberdinak izanik. Hauei fenomenoaren aldaerak deituko diegu. Hortaz, aldaera desberdinen multzoak fenomeno beraren aldakortasuna adieraziko du. Honegatik, aldakortasuna esplizituki errepresentatzeko *Aldakortasuna* izeneko multzoa (set) definitzen dugu (IV.9 irudia).

```
set Aldakortasuna ;
  element-type: instance(Fenomeno_linguistikoa);
  cardinality: min 0 max infinite;
  properties:
    fenomeno_mota: string; value={...}
```

IV.9 irudia. *Aldakortasuna* multzoaren definizioa.

Atal honi bukaera eman aurretik hizkuntzaren ikaste-prozesuan hain garrantzitsu diren ikaste-estrategiei hitz batzuk eskaintzea beharrezkotzat jotzen dugu, hauek domeinuaren antolakuntzan eragina baitute. § IV.2.1-n aipatu dugu hizkuntzaren ikaste-prozesuan hizkuntz ezagutza osatzen joan ahala ikasleak baliabide berriak eskuratzen dituela egitura berriak sortzeko ahalmena garatuz. Ahalmen hori, ordea, ez da modu berean bilakatzen ikasle guztiengan. Ikasle ausartak egitura berriak sortzen saiatzen diren bitartean, ikaste-estrategia moduan, arrisku hartzea gustuko ez duen ikasleak era desberdinean eskuratzen du hizkuntz ezagutza, input-a bera izan arren. Hori dela medio, domeinuan ez dugu nahikotzat jo ikaskuntza-mailaren arabera soilik ezagutza antolatzea; maila bakoitzean ikasteko strategi moten araberako eskuratze-prozesuaren emaitza ere adierazi dugu (ikus IV.10 irudia).

```
set Ikasle_multzoa;
  element-type: instance(Ikasle_eredua);
  cardinality: min 0 max infinite;
  properties:
    identifikadorea: string;
    ikaskuntza_maila: integer-range(1, MAILA_GORENA);
    ikaste_estrategiak: set(instance(Ikaste_estrategia));
    hizkuntz_ezagutza_mailarena: instance-name(Hizkuntz_ezagutza_mailarena);
```

IV.10 irudia. *Ikasle_multzoa* multzoaren definizioa.

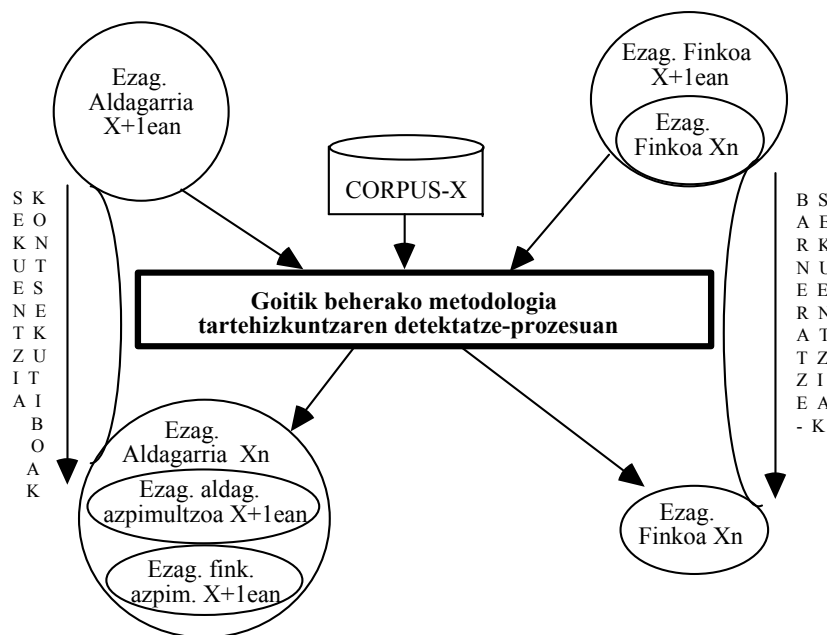
IV.3 Ikaskuntza-mailako ereduaren eraikuntzarako erabilitako metodologia.

Atal honetan ikaskuntza-maila baten tartehizkuntza detektatzeko proposatzen dugun goitik beherako metodologia azalduko dugu. Honetaz gain, prozesuan zehar ezagutza finkoa eta aldakorra desberdintzeko irizpide nagusiak aipatuko ditugu.

IV.3.1 Goitik beherako metodologia.

Ikaskuntza-mailako ereduaren eraikuntzarako goitik beherako metodologia aukeratu dugu, hau da, ikaskuntza-maila gorenaren modelizaziotik hasi eta hurrengo hizkuntz mailen ereduak atera goikoaren emaitzetan oinarrituz, honela ikaskuntza-maila baxueneraino helduz (ikus IV.11 irudia). Metodologia hau jarraitzeko arrazoiak ondokoak izan dira: egun euskararako ditugun tresna linguistiko-konputazionalak (lematizatzailea, zuzentzaile ortografikoa, desanbiguatzaile morfologikoa, ...) ikaskuntza-maila altuetara hobeto egokitzen dira maila baxuetara baino; gainera, testu idatziak analizatzen dituzten tresna konputazionalak osatuagoak izaten dira ikaskuntza-maila altuetan eta, azkenik, maila altuetan maila baxuetan baino material idatzi gehiago topatzen dugu hizkuntz eskoletan.

Tartehizkuntzaren gaitasun gramatikalean ezagutza orokorra eta espezifikoa izaten da, hau da, testuinguruarekiko independentea eta dependentea hurrenez hurren. Desberdintasun hau, halere, ez dugu irudian adierazi. Bertan ezagutza finkoa eta aldakorraren garapena hizkuntz mailetan zehar islatzen da. IV.1 irudian azaldu genuenez ikaskuntza-maila baten ezagutza finkoa hurrengo mailaren azpimultzoa da; ezagutza aldakorraren kasuan, berriz, ikaskuntza-maila kontsekutiboen ezagutza aldakorrak independenteak dira, elkarren arteko ebakidurak izaten badira ere. IV.11 irudian adierazten da nola X ikaskuntza-maila emanda, bere ezagutzaren modelizazioak X+1 mailako ezagutza osatuta erabiltzen duen, X mailako corpusa (CORPUS-X) aztertzean.



IV.11 irudia. X ikaskuntza-mailaren tartehizkuntz ereduaren modelizazioa.

Egindako azterketan maila ertainak eta altuetan idatzitako testuek osatutako corpusaren azpimultzoak aztertu ditugu. Aztertze-unitatea testua izanik, hainbat testuk osatutako multzoak aztertu ditugu automatikoki eta emaitzak sakonki landu. Ezaugarri desberdinak betetzen zituen zenbait ikaslek idatzitako testu-mota desberdinek osatzen zuten multzo bakoitza.

Ikaskuntza-maila gorenaren (G maila) eredia eraikitze dagoeneko euskararako garatuta dauden tresna linguistiko-konputazionalak hartu ditugu abiapuntutzat. Emandako pausuak ondokoak izan dira:

- 1) Euskararako Datu-Base Lexikaleko sarreren orrazketa. Oinarrizko tresnen ezagutzan hizkuntzaren ezaguera zabala duen pertsonaren hiztegia garatu zen, hizkuntza menperatzen duen ikasle trebeenak ere ia eskuratu gabea. Honegatik, hitzen analisisan batzuetan agertzen diren interpretazioen lema ez da ezagun izaten ikaslearentzat. Hori dela eta, honelako lemen interpretazioa baztertzen duen *hiztegi_eguneraketa* izeneko tresna sortu dugu. Oraingoz, tresnak lemen eta hauen hainbat ezaugarri islatzen duten erregelen ezagutza-basea sortzen du. Erregela hauek erabiliz etorkizunean egoki ikusten dugu ikaskuntza-maila bakoitzerako Datu-Base Lexikal bana egokitua sortzea.
- 2) Corpusaren (CORPUS-G) eta honen azpimultzoen aukeraketa. Testuak aukeratzeko irizpideak § III.3.3.1-n azaldu ditugu.
- 3) Morfologia eta morfosintaxiaren definizioa CORPUS-Gren azpimultzo batean oinarrituz. Helburu honetarako erabilitako tresnen deskribapena hurrengo atalean egingo dugu.

- 4) Ezagutza finkoa eta aldakorraren identifikazioa. Domeinuaren adierazpenean azaldutako testuinguru-motak kontuan izan ditugu.
- 5) Emaitzen ebaluazioa:
 - a) Ereduaren zehaztasunaren azterketa, CORPUS-Gren azpimultzo berriak erabiliz.
 - b) G ikaskuntza-mailako irakasle batek emaitzak ebaluatzea.

Aditu batek emaitzak ebalua ditzan beharrezkoa da ereduaren eraikuntzan, corpusaren azterketa ez baita nahiko izaten. Corpusaren azterketaren emaitzak ikaskuntza-maila bereko beste corpus batzuen azpimultzoak erabiliz ebaluatu genituen, eta, gero, ebaluazioa irakasleen azterketarekin konparatu. Adibide gisa, ikaskuntza-maila altuko 16 testutako (5310 hitz) emaitzak konparatutakoan bi ebaluaketa nahiko antzeko izan arren, kasu batzuetan corpusaren emaitzak eta irakaslearen iritzia ez zetozen bat. Esaterako, hitz-mailako azterketa egitean irakaslearen ustetan ikasleek ohitura handiagoa zuten desbideratzeetan *h* letra ezabatzeko gehitzeko baino. Corpusaren emaitzetan, ordea, ez genuen detektatu diferentzia nabarmenik bi fenomenoaren artean.

IV.3.2 Ezagutza finkoa eta aldakorraren identifikaziorako irizpideak.

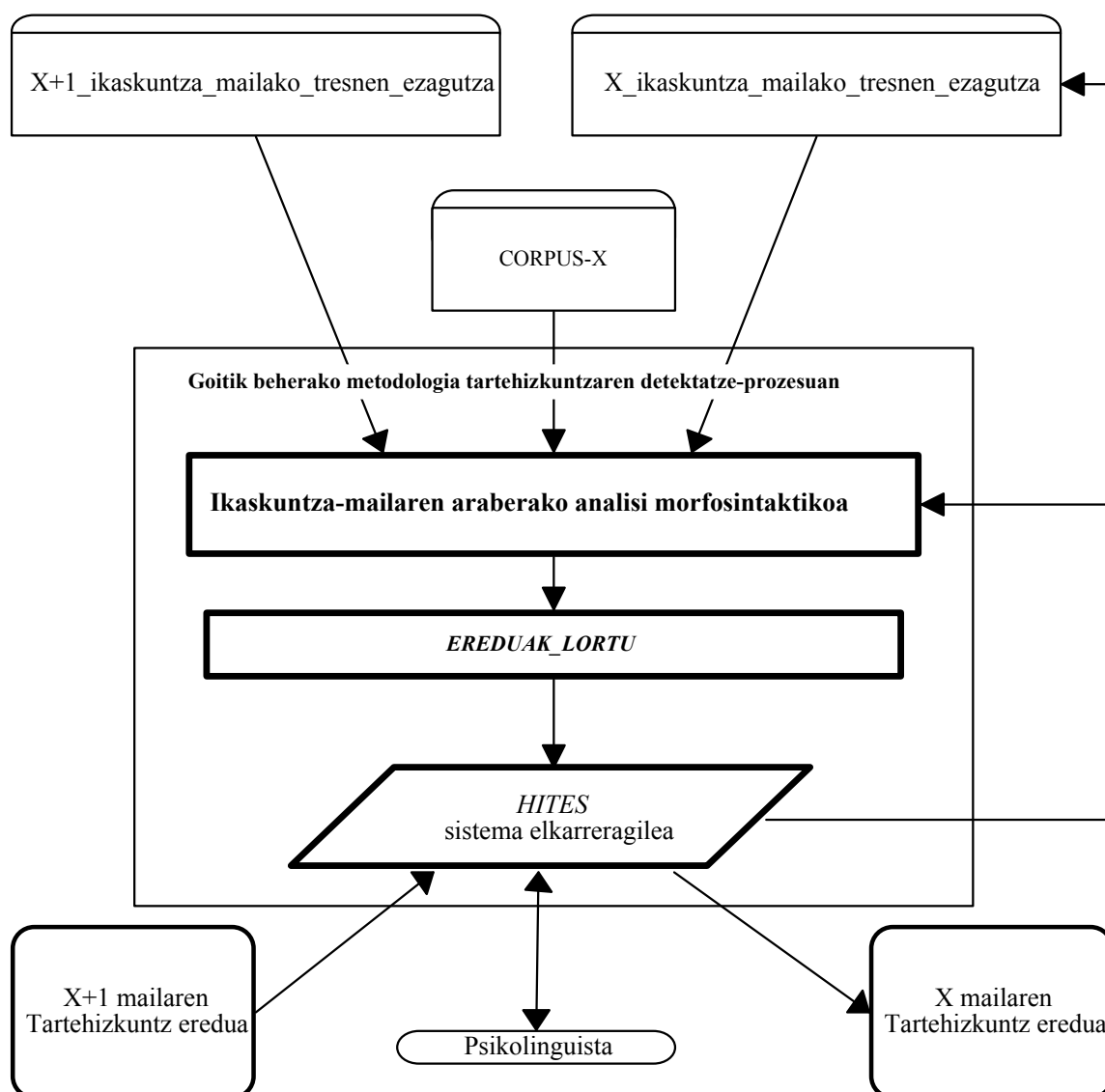
Tartehizkuntz egiturak detektatutakoan ezagutza finkoa edo aldakorra den zehaztea ez da zeregin erraza. Landa-lanean ondorioztatutako hainbat irizpide izan da erabakiaren oinarria. Ondoren irizpideak komentatuko ditugu:

- Lehenengoz, erregela guztiek ez dute erabilpen-maiztasun bera, ez estandarrek, ez desbideratzeek. Beraz, erregela zehatz bat finkatu dela esateko probabilitatea erregela espezifikoaren menpe eta sortu den ingurunearen menpe dago. Adibidez, eman dezagun irakasle batek idazlan bat idaztea eskatzen diela ikasleei egitura linguistiko batzuk erabil ditzaten. Kasu honetan egitura linguistikoaren bertsio estandarra erabili arren behin edo bitan erabiliz gero irakaslearen egitura barneratu dela esaterik ez dago.
- Bigarren, hainbat erregela ikaskuntza-maila batean detektatu eta maila altuagoetan desagertzea fenomeno arrunta da. Honen ondorioz, maila baxuetan detektatutako erregela bat *X* mailan desagertuz gero, *X* maila baino altuagoetan ere ez bada azaltzen, erregela beti ezagutza aldakorrean egongo da *X* baino baxuago diren mailetan.
- Hirugarren, analizatutako corpusean agertzen diren lema desberdinen kopuruak eragina du ezagutza finkoa eta aldakorraren bereizketan; adibidez, lema desberdin gutxiako corpus batean erregela morfologiko bat behin agertzeak ez du esan nahi finkaturik ez dagoenik, edo ez agertzeak aztertzen ari den ikaskuntza-mailari ez dagokionik.

- Azkenik, esperimentuetan baieztatu dugu ariketa-motak ere eragina izaten duela hainbat kasutan.

IV.4 Ikasle eta ikaskuntza-mailako ereduen hizkuntz ezagutza detektatzeko tresnen deskribapena.

Esan bezala, HITES sistemaren domeinua ondoko bi elementuk osatzen dute: ikaslearen tartehizkuntzak eta ikaskuntza-maila bereko ikasleek amankomunean duten hizkuntz ezagutzak. Azken honetan bi ezagutza-mota topatzen dugu HITES sisteman: alde batetik, ikasleen tartehizkuntza eta, beste alde batetik, tresna linguistiko-konputazionalen ezagutza linguistikoa (ikus IV.12 irudia). Lehenak bigarren hizkuntza eskuratzen ari den ikaslearen ikuspegiari darraio, bigarrenak, berriz, tresna linguistiko-konputazionalen filosofia du oinarri. HITES sistemak alde aurretik garatutako tresna linguistiko-konputazionalen berrerabilpena eta egokitzapena lantzen du ikasleen tartehizkuntzaren fenomeno linguistikoak detektatzea helburu izanik. Prozesu honetan tartehizkuntza osatzeaz gain tresnen hizkuntz ezagutza ere joango da eguneratzen kasuan kasu.



IV.12 irudia. Goitik beherako metodologia.

Aurreko atalean azaldutako goitik beherako metodologia jarraitzen duen *ereduak_lortu* izeneko tresnak hiru emaitza desberdin ematen ditu: bat, bai goiko mailan bai behekoan suertatzen diren analisiak; bi, goiko mailan bakarrik agertzen direnak, eta, hiru, beheko mailan bakarrik azaltzen diren analisiak, hurrengo mailako ikaste-prozesuan aurrerago desagertuko direnak. Tresnak ikaskuntza-maila baten corpusa (CORPUS-X) erabiliz bi ikaskuntza-maila desberdinen arabera (X eta X+1) egiten du analisia behin eta berriro. Iterazio bakoitzean HITES sistemak X ikaskuntza-mailaren tartehizkuntz eredua eta tresnen ezagutza eguneratzen du psikolinguistaren laguntzaz: sistemak, beraz, ikasi egiten du iterazioen ondoren. Honela, hurrengo iterazioan ikasitako ezagutza erabiliz *ereduak_lortu* programak emaitza zehatzagoak lortuko ditu.

Ikaskuntza-mailaren araberako analisi morfosintaktikoa burutzeko tresnen deskribapena eta *ereduak_lortu* programaren azalpen zabalagoa emango dugu § IV.4.1-n, eta IV.20 irudian ikaskuntza-mailaren araberako analisi morfosintaktikoa zehatzago adieraziko dugu. Ondoren, § IV.4.2-n hain zuzen, tresna konputazionalen ezagutzaren eguneraketarako tresnak komentatuko ditugu. Kontuan izan, hemen ez dugula azalduko tartehizkuntz eredua eraikitzekeo prozesu osoa. Gauden honetan corpusetik abiatuak tartehizkuntzan hautagai izan daitezkeen fenomeno linguistikoak detekzioa bakarrik ari gara lantzen. Psikolinguistaren laguntzaz (HITES sistema elkarreragilea) tartehizkuntz eredu osatua lortuko duen sistemari buruz, berriz, hurrengo kapituluan arituko gara.

IV.4.1 Tartehizkuntzaren erregela eta fenomeno linguistikoak detektatzeko tresnen deskribapena.

Lan honen hasieran genienez, euskararako tratamendu automatikorako tresnen garapena helburu orokortzat hartu zuen IXA ikertaldean sortu zen aurrean dugun tesia. Une horretarako garatutako hainbat tresna berrerabili dugu HITES sisteman, zenbait egokitzapen burutu behar izan delarik, analizatzaile morfologikoa (Agirre et al., 92) eta lematizatzailea (Aldeazabal et al., 94) esaterako. Egokitzapenez gain tresna berriak ere sortu behar izan ditugu; ikaskuntza-mailaren araberako desanbiguatzailea eta *Estatistikak* programa elkarreragilea aipagarrienak izanik. Ondoren, sortutako eta egokitutako tresna linguistiko-konputazionalak deskribatuko ditugu:

- Analizatzaile morfologikoa.
- Post-prozesu morfologikoa.
- Ikaskuntza-mailan oinarritutako desanbiguatzailea.
- Testuinguruetan oinarritutako desanbiguatzailea.
- Desanbiguatzaile morfosintaktikoa.
- Statistikak.
- Laginak.
- *Ereduak_lortu*.

Hauen artean bai *Estatistikak* bai *Laginak* ereduaren sorkuntzarako prozesu lagungarriak dira, baina ez beharrezkoak. Tresnen arteko loturak adierazten dira IV.20 irudian. Irudi hau IV.12 irudiaren zabalpena da.

• **Analizatzaile morfologikoa:** egokitutako analizatzaileak hitz baten analisisetan aplikatu erregela morfofonologikoak detektatzen ditu, bai erregela estandarrek, baita desbideratzeak ere. Analizatzaileak Bi Mailatako Morfologia erabiltzen du (Koskenniemi, 83) eta Agirre et al.ek (1992) inplementatutako analizatzailearen egokitzapena da. Emaitzan informazio morfologiko eta morfosintaktikoa lortzen da.

Analizatzaileari buruzko azalpen zehatzagoa hurrengo azpiatalean emango dugunez, beste tresna batzuk komentatuko ditugu oraingoa.

- **Post-prozesu morfologikoa:** detektatutako erregelak, estandarrak zein desbideratzeak (adib. err4, err39, ...), identifikatu eta hitz barruan non aktibatu diren (leman/morfeman) erakusten du. Honetaz gain, bestelako ezaugarriak ere detektatzen ditu: hitzaren luzera eta lemaren azken letra-mota (bokala/kontsonantea). Honelako ezaugarrien garrantzia nabarmen geratu zen landa-lanean irakasleekin izandako elkarrizketetan. Beraz, erregelen aplikazio-eremuen azterketa egin eta hainbat arazo topatu genuen tratamendu automatikoan. Adibidez, honelako galderei erantzun behar izan genien: *zer da hitza?*. Guk honela definitu genuen: zenbaki edo hizki batez hasten den unitatea, beraz, "a" eta "3" hitzak izango dira, baina ";" ikurra ez da hitza izango.

```
/<hark>//  
  ((forma "hark")  
...  
  %T::0  
    ((anal 3)  
    ((lema "haQ")((SAR hura)(KAT DET)(AZP ERK)(LUZ LUZ4)(NOLA KTS)))  
    ((morf "Ek")((SAR k)(KAT DEK)(KAS ERG)(NUM S)(MUG M)(FS1 @SUBJ))))  
  %T::0  
    ((anal 4)  
    ((lema "haQ")((SAR hura)(KAT IOR)(AZP PER)(LUZ LUZ4)(NOLA KTS)))  
    ((morf "Ek")((SAR k)(KAT DEK)(KAS ERG)(NUM S)(MUG M)(FS1 @SUBJ))))
```

IV.13 irudia. *hark* hitzaren bi analisi.

IV.13 irudian agertzen den analisisan ikusten denez analizatzaile morfologikoak hainbat diakritiko (Q) erabiltzen du aldaketa morfofonologikoak adieraztearren. Post-prozesuan diakritikoen korrespondentzi taula gehitu da lemen bukaera nolakoa den jakiteko. Adibidez, R,Q diakritikoak r letra bihurtzen dira; beraz, *hark* hitzaren "lema" kontsonantez bukatzen dela ondorioztatzen da.

IV.14 eta IV.15 irudietan post-prozesuaren emaitza nolakoa den azalduko dugu **oihanan* eta *oihanean* hitzen analisisen bidez.

```
/<oihanan>//  
  ((forma "oihanan")  
  %T:err39:1  
    ((anal $ 1)  
    ((lema "oihan")((SAR oihan)(KAT IZE)(AZP ARR)(LUZ LUZ7)(NOLA KTS)))  
    ((morf "0")((SAR 0)(KAT DEK)(NUM S)(MUG M)))  
    ((morf $ "Ean") (ald "Ean")((SAR an)(KAT DEK)(KAS INE)  
    (FS1 @ADLG))(39)(EMOT MOR_LEKTE3)  
  )
```

IV.14 irudia. *MOR_LEKTE3*: -e letra kendu du morfeman.

```

/<oihanean>/
((forma "oihanean")
 %T:err4:l
 ((anal $ 1)
 ((lema "oihan")((SAR oihan)(KAT IZE)(AZP ARR)(LUZ LUZ8)(NOLA KTS)))
 ((morf "0")((SAR 0)(KAT DEK)(NUM S)(MUG M)))
 ((morf $ "Ean") (ald "Ean")((SAR an)(KAT DEK)(KAS INE)
 (FS1 @ADLG))(4)(EMOT MOR_LEGBE)
 )

```

IV.15 irudia. MOR_LEGBE: -e letra gehitu du morfeman.

- **Ikaskuntza-mailan oinarritutako desanbiguatzailea:** desanbiguatzaile honen helburua da aztertzen den ikaskuntza-mailan posible ez den hainbat analisi ezabatzea. Modulu honetako erregelak esperimentuetan oinarriturik aldatuko dira maila bakoitzean. Atal honen bukaeran, hiztegia eta kategoria gramatikalen azterketa azaltzean, desanbiguatzaile honetan aplikatutako heuristikoak komentatuko ditugu. Heuristiko hauen arabera, esaterako, maila altuetan hitz baten interpretazio berean ez da agertuko bidesbideratze-erregela baino gehiago. Adibidez, *arriskurik* hitzaren analisisetan (ikus IV.16 irudia) 4 interpretazio baztertzen ditugu (ikus IV.17 irudia) — 3 eta 4 desbideratze aktibatu dira interpretazio horietan —.

Hitzaren luzera eta aktibatutako erregela-kopuruaren arteko erlazioa ere landu da hainbat kasutan.

```

/<arriskurik>/
("arrisku" IZE ARR LUZ10 BOK + DEK PAR MG @OBJ @SUBJ)
("hari" /hari/ IZE ARR LEM_LEGBS LEM_LEGTR LEM_LEKHH LUZ10 BOK + DEK NUMS
 MUGM + DEK GEL @IZLG> @<IZLG @ADLG MOR_LEOEOU + DEK PAR MG
 @OBJ @SUBJ)
("hari" /hari/ IZE ARR LEM_LEGTR LEM_LEKHH LUZ10 BOK + DEK INS MG @ADLG
 MOR_LEOEZS + DEK GEL @IZLG> @<IZLG @ADLG MOR_LEOEOU + DEK PAR
 MG @OBJ @SUBJ)
("harri" /harri/ IZE ARR LEM_LEGBS LEM_LEKHH LUZ10 BOK + DEK NUMS MUGM + DEK
 GEL @IZLG> @<IZLG @ADLG MOR_LEOEOU + DEK PAR MG @OBJ @SUBJ)
("harri" /harri/ IZE ARR LEM_LEKHH LUZ10 BOK + DEK INS MG @ADLG MOR_LEOEZS +
 DEK GEL @IZLG> @<IZLG @ADLG MOR_LEOEOU + DEK PAR MG @OBJ @SUBJ)

```

IV.16 irudia. *arriskurik* hitzaren analisiak desanbiguatzailea pasa aurretik (formato laburtua).

```

/<arriskurik>/
("arrisku" IZE ARR LUZ10 BOK + DEK PAR MG @OBJ @SUBJ)

```

IV.17 irudia. *arriskurik* hitzaren analisiak desanbiguatzailea pasa ondoren (formato laburtua).

- **Testuinguruetan oinarritutako desanbiguatzailea:** honek erregela morfofonologikoak aktibatutako testuinguruaren arabera (adib. leman/morfeman) hainbat interpretazio baztertzen du hitz baten analisisetatik. Aplikatutako erregela eta lemaren

kategoria gramatikala bat ez badatoz ere interpretazioak baztertzen dira (adib. *LEOEUO¹ aditzetan ez da gertatzen). Desanbiguatzaileen erregelak definitzeko *Estatistikak* izeneko programa elkarreragilea diseinatu dugu testuen azterketa partzialak egin ahal izateko. Honi buruz beranduago hitz egingo dugu.

```
MS_DESB_12:      (REMOVE MOR_LEOEZS);
"<analysis>" MS_DESB_1
  "analisi" "analisi" IZE ARR LEM_LEGBS LUZ8 BOK + DEK ABS MG @OBJ @SUBJ
  "analisi" "analisi" IZE ARR LEM_LEGBS LUZ8 BOK ZERO
-> "analisi" "analisi" IZE ARR LUZ8 BOK + DEK INS MG MOR_LEOEZS @ADLG
```

IV.18 irudia. Testuinguruetan oinarritutako desanbiguatzailearen adibidea.

- **Desanbiguatzaile morfosintaktikoa:** hitz-mailako eta esaldi-mailako informazio morfosintaktikoa erabiliz hitzen interpretazioak aukeratu, etiketatu edo ezabatzen ditu Murriztapen Gramatika erabiliz (Karlsson et al., 95; Alegria et al., 96). IV.19 irudian azaltzen den MS_DEK_GABE_111 erregela morfosintaktikoaren bidez **analisi* hitzaren interpretazio batekin geratzen da. Konparatzen baditugu IV.18 irudian eta IV.19 irudian dauden analisiak lehenengoan *ZERO* ezaugarria duen interpretazioa da bigarren irudian aukeratu dena MS_DEK_GABE_111 erregelaren aplikazioaren ondorioz. Erregela honek *ZERO* ezaugarria duten interpretazioak aukeratzen ditu izen arruntak direnean, ezker aldeko hitza ez bada determinatzailea edo mailakatzailea, eta eskubikoa deklinatua ez badago (postposizioak kenduta)³.

```
MS_DEK_GABE_111: SELECT (ZERO) IF (0C IZE + (ARR))
                  (NOT 0 AORG OR GEN/GEL OR POSPOZ OR ADIKONP)
                  (NOT -1 DET OR MAILAKATZAILEAK) (NOT 1 POSPOZ)(1C DEK);

"<analysis>" MS_DESB_1, MS_DEK_GABE_111
  "analisi" "analisi" IZE ARR LEM_LEGBS LUZ8 BOK ZERO MS_SINT_53 @KM>
```

IV.19 irudia. *analisi* hitzaren analisi morfosintaktikoa.

- **Estatistikak:** corpusaren azpimultzoak aukeratzen dituen programa elkarreragile honek ondoko aukerak eskaintzen ditu: ikaskuntza-maila, ariketa-mota, ikasle konkretua, testu konkretua eta testu-zerrenda. Behin corpusa aukeratu ondoren hainbat azterketaren emaitza eskaintzen du, besteak beste:

1) erregela bat aktibatu duten hitz-kopurua.

¹*LEOEUO erregela: *u* letraren ordez *o* letra erabiltzea.

²MS_DESB_1 erregelak aztergai dugun ikaskuntza-mailan gertatzen ez den DESBideratze bat tratatzen du:
(REMOVE MOR_LEOEZS);

³Adibidean bertan ikusten da @KM> funtzio sintaktikoa esleitu dion MS_SINT_53 erregela morfosintaktikoa aplikatu dela, MS_DEK_GABE_111 erregela desanbiguatzaileaz gain.

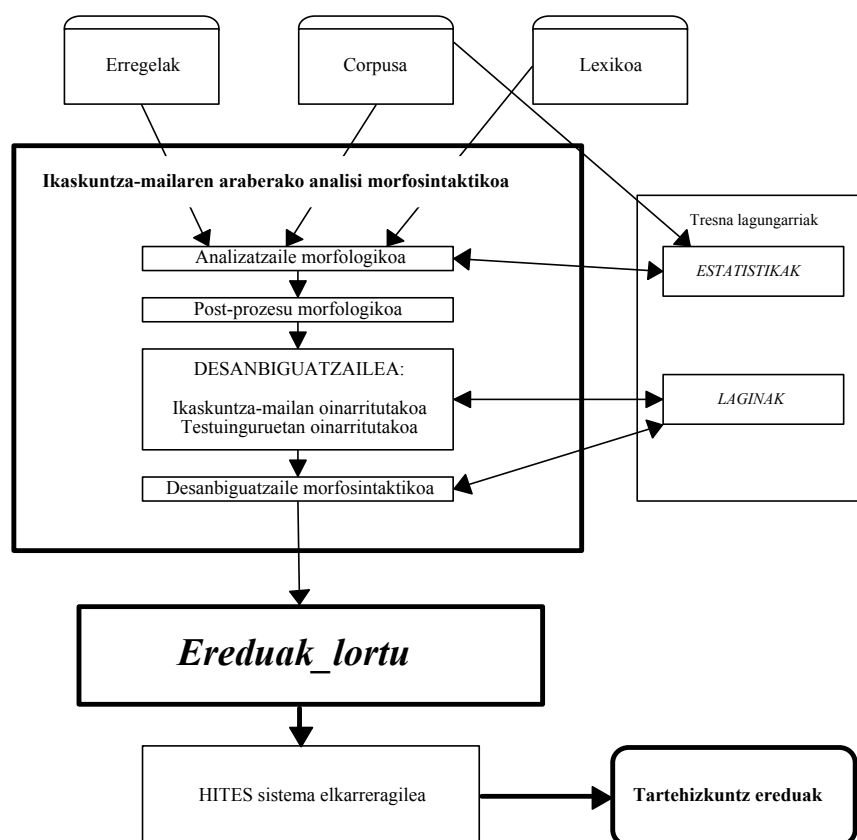
- 2) morfemetan bakarrik agertzen diren erregelen zerrenda.
- 3) erregela baten aktibazio-kopurua corpusaren hitz-luzeren batezbestekoaren arabera.
- 4) kategoria gramatikal konkretu batean aktibatzen diren erregelak.

Estatistikak programak analizatzailearen irteera erabiltzen du sarrera gisa, eta desanbiguitzaileen erregelak definitzeko laguntza eskaintzen digu. Adibidez, detektatzen badugu erregela bat, *LEOEZS¹ esaterako, lemetan bakarrik aplikatzen dena, erabaki dezakegu erregela berri bat desanbiguitzailean gehitzea, zeinek baztertzen baitu *LEOEZS erregela morfeman aplikatzen denean. Honela, ikasle batek **analisi* idatziko balu hitzaren hirugarren interpretazioa baztertuko genuke ondoko erregelaren bidez:

- **Laginak:** tresna honek, *Estatistikak* izeneko programak bezala, desanbiguitzaileen erregelak fintzeko laguntza ematen digu, eta bai erregela morfosintaktiko berriak zehazteko ere. Desanbiguitzaileen irteera sarrera gisa erabiliz, ezaugarri morfosintaktikoen formula logiko konkretua beren interpretazioetan betetzen duten hitzak lortzen dira, testuaren ondorengo hitza ere agertzen delarik.

- **Ereduak_lortu:** tresna honek (ikus IV.20 irudia) tartehizkuntz ereduaren tartehizkuntz egiturak sortzeko hautagai izan daitezkeen fenomeno linguistikoak eta hauen erregelen ereduak ematen ditu. Testuinguruaren informazioa itzultzen du, hau da, esperimentuetan garrantzitsutzat jo den informazioa: kategoria gramatikala, ariketa-mota, lema-azken letra-mota (kontsonantea/bokala), eta abar. Esana dugu, atal honen hasieran, *ereduak_lortu* programak goitik beherako metodologia aplikatzen duela, hortaz, fenomenoak lortzeaz gain goiko maila eta beheko mailaren arteko konparaketa ere egiten du. Konparaketaren emaitza *mailakatzaileria* izeneko eremuan gordetzen da (ikus IV.21 irudia). Gainera, fenomenoak eta erregelak identifikatzean hitz edo esaldi-mailakoak diren eta maila morfologikoan edo morfosintaktikoan kokatzen diren esplizituki adierazten du.

¹*LEOEZS erregela: z letraren ordez s letra erabiltzea.



IV.20 irudia. Ikaskuntza-mailaren araberako analisi morfosintaktikoa.

Esan bezala, erregelen erdez gain fenomeno linguistikoaren ereduak ere ateratzen ditu *ereduak_lortu* izeneko tresnak, hain zuzen ere, hautagai izan daitezkeen fenomenoak, psikolinguistak geroago aztertuko dituenak. Fenomeno-multzoa osatzean hiru konbinaketa lantzen ditu:

1. Erregela morfofonologikoen arteko konbinaketak.
2. Erregela morfofonologiko eta morfosintaktikoen artean interpretazio-mailan aplikatzen direnak.
3. Erregela morfosintaktikoen konbinaketak, bai interpretazio-mailakoak, bai hitz-mailakoak.

Adibide gisa, erregela morfofonologikoen artean ondoko konbinaketa topatzen dugu: "LEOEAE_1"¹ "LEGBE_1". Bi erregela hauek "*da + -En + -Ean -> denean*" fenomenoak adierazten dute.

¹LEOEAE erregela: *a* letraren ordez *e* letra erabiltzea.

LEGBE erregela: *e* letra gehitzea.

Ikus dezagun 10. mailako hainbat ikaslek idatzitako corpusaren lagin txiki batetik (29 esaldi eta 524 hitz) lortutako emaitzaren erregela-adibide batzuen irteera nolakoa den:

- Adibide1 : LEOEAE erregela morfologikoa: aditzen bukaerako a-ren aldaketa.
Testu zatia: ... *etorri **direla** esan zuen* ...
- Adibide2 : LEKTE1 erregela morfologikoa: e-ren atzetiko e-ren galera.
Testu zatia: ... *hamabost **urtekin** ezagutu zuen* ...
- Adibide3 : MS_EZ erregela morfosintaktikoa: ezezko esaldien detekzioa.
Testu zatia: ... *mutilak **ez** daki hori* ...

Adibide1

LEOEAE¹ erregela estandarra aditz laguntzaile (ADL) zein aditz trinkoen (ADT) lemetan (LEM) azaltzen da beti lema bokalez (BOK) bukatzen dela; hitzen luzeraren batezbestekoak ez du eraginik (LUZ>= zein LUZ< duten hitzetan agertzen da) eta k motako ariketetan topatu da. *Baldintzak* eremuan hitza analizatzean erabilitako Bi Mailatako Morfologiak ematen digun testuinguru-informazioa agertzen da. Emaitzan ikusten denez bai deskribapen linguistikoa, *deskribapena* eremua, bai azalekoa, *azaleko_deskribapena*, lortzen dira (ikus IV.6 irudia). Gainera *erregela_agindu*-mota ere lortzen da, eta, hala, ikaslearen jokaerak aztertzeke aukera dugu.

```
(erregela_linguistikoa LEOEAE_1
 (identifikadorea "LEOEAE")
 (deskribapena "aditzen bukaerako a-ren aldaketa")
 (azaleko_deskribapena "A-ren ordez E jarri du hitzaren lemaren bukaeran ")
 (agindugaiak "a" "e")
 (erregela_agindua "ORDEZKATU")
 (adib "(den izan)" "(diren izan)" "(garela izan)" "(direla izan)" "(dela izan)" "(dena izan)"
 "(denean izan)" "(dena izan)" "(direnak izan)")
 (baldintzak "@:e <=> () _ ((+:=)+[a:a|e:e|o:o])"
 "@:e <=> () _ ((+:=)+E:0[k:k|an:an|tik:zik|z:z])"
 "@:e <=> () _ ((+:=)+E:0[n:n|l:l])")
 (tes_ling "(ADL LEM)" "(ADT LEM)")
 (tes_ez_ling "BOK" "LUZ>=" "LUZ<" k)
 (mota_linguistikoa "morfo")
 (mota "hitza")
 (mailakatzailea "BI_IKASKUNTZA_MAILATAN"))
```

IV.21 irudia. *ereduak_lortu* tresnaren irteera. LEOEAE erregelaren eredia.

¹LEOEAE erregela: a letraren ordez e letra erabiltzea.

Adibide2

Ikus dezagun beste adibide bat (IV.22 irudia), non erregela morfeman aplikatu baita. Kasu honetan testuinguru linguistikoan erregela atzizkian (MOR*¹) aplikatu dela azaltzeaz gain atzizkiaren kategoria gramatikala ere agertzen da (DEK).

```
(erregela_linguistikoa LEKTE1_1
 (identifikadorea "LEKTE1")
 (deskribapena " e-ren atzetiko e-ren galera")
 (azaleko_deskribapena " E letra kendu du hitzaren lemaren barrualdean, hau da tartean")
 (agindugaiak "e")
 (erregela_agindua "EZABATU")
 (adib "(urtekin urte)")
 (baldintzak "e:0 => (e:e(+:0)+)_(n|M(+:0|$: $))")
 (tes_ling "( IZE DEK MOR*")
 (tes_ez_ling "BOK" "LUZ>=" "k" )
 (mota_linguistikoa "morfo")
 (mota "hitza")
 (mailakatzailea "BI_IKASKUNTZA_MAILATAN"))
```

IV.22 irudia. *ereduak_lortu* tresnaren irteera. LEKTE1 erregelaren eredua.

Adibide3

Erregela morfosintaktikoei dagokienez bi erregela-mota desberdintzen dira, eta kasu bakoitzean ateratzen den informazioa desberdina da. Alde batetik, interpretazio-mailan aplikatzen diren erregelak ditugu, hauek erregela morfofonologikoen kasuan bezala, hitz-mailako testuinguru linguistiko eta testuala izango dute; beste alde batetik, hitz-mailan aplikatzen diren erregelak, hauek hitzaren interpretazio guztiei egiten diete erreferentzia, hortaz, ez dago esaterik, adibidez, lema kontsonantez edo bokalez bukatzen den, interpretazio guztiek ez baitute zertan lema bera izan. Azken hauetan, beraz, testuingururik ez da agertuko eremuan.

¹MOR* ezaugarriak erregela atzizki batean aktibatu dela adierazten du. Aurrizkian balitz *MOR ezaugarria topatuko genuke.

```
(erregela_linguistikoa MS_EZ_1
 (identifikadorea "MS_EZ")
 (deskribapena "ezezko esaldien detekzioa")
 (azaleko_deskribapena " ")
 (baldintzak " "<ez>" REMOVE (IZE) IF (1C ADL/ADT) ; ")
 (agindugaiak "ez " "IZE")
 (erregela_agindua "REMOVE")
 (mota_linguistikoa "morfo_sintak")
 (mota "esaldia")
 (mailakatzaila "BI_HIZKUNTZ_MAILETAN"))
```

IV.23 irudia. *ereduak_lortu* tresnaren irteera. MS_EZ erregelaren erdua.

IV.23 adibidean agertzen den erregelak ondokoa adierazten du: "ez" hitzaren IZE interpretazioak ezabatu baldin eta eskubiko hitza ADL edo ADT bada. Erregela hitzaren interpretazio guztiei buruz ari da, ondorioz hitz-mailako testuinguru linguistikoak (hots, *tes_ling*) eta ez-linguistikoak (hots, *tes_ez_ling*) ezin dute ereduan agertu, hauek interpretazio konkretuei buruz azalpenak ematen dituztelako. Oraingoz esaldi-mailako testuinguruak ez dira ereduan agertzen; hau aztertze-prozesuan dago oraingoz. Identifikadoreari dagokionez, zenbakidun kodea dugu. Erregela morfosintaktikoei zenbaki kontsekutiboak asignatzen zaizkie ezagutza-basean, oraingoz ez direlako kodetu erregela morfosintaktikoak.

Adibide honek, bestalde, argi uzten du psikolinguistaren beharra ereduaren eraikuntzan, erregelaren ikuspegi linguistiko-konputazionala eta ikaslearen tartehizkuntzaren interpretazioa ez baitatoz bat. Hizkuntza ikasten ari denaren ikuspegitik *ezezko esaldia* erabili duela esango genuke *ez* hitzaren ezezko interpretazio ez direnak ezabatu baitira. Eta honek ez du esan nahi ikaslearen jokaera ezabaketa (REMOVE) denik. Hemen psikolinguistaren partehartzea ezinbestekotzat hartzen da sisteman. Baina, honi buruz zabalago arituko gara hurrengo kapituluan.

Hiztegia eta kategoria gramatikalen azterketa

Orain arte deskribatutako tresnen helburua zen hautagai izan daitezkeen fenomeno linguistikoak automatikoki detektatzea. Esana dugu fenomeno semantikoaren azterketa baztertu dugula gure lanean. Hala ere ikaslearen hiztegia aztertzeko lehen pausuak eman ditugu. Lemen azterketa honek hiru helburu ditu:

- Ikaslearen hiztegia aztertea.
- Ezagutza finkoa eta aldakorraren azterketan corpuseko lema desberdinen kopuruaren eragina zehaztea.
- Kategoria gramatikalen erabilpena lantzea ikaskuntza-maila bakoitzean.

Lemen azterketa dela eta, diseinu mailan beharrezkotzat jotzen dugu Euskararako Datu-Base Lexikalaren (EDBL) sarrerei egonkortasun balioa asignatzea tartehizkuntzaren definizioan fenomenoei gehitu zaien moduan. Honetaz gain sarrera bakoitzak dagozkion hizkuntz mailen zerrenda beharko luke, honela ikaskuntza-maila bakoitzeko tresnek, hots, analizatzaileak, desanbiguatzaileak eta lematizatzaileak, mailari dagokion hiztegia erabiliko lukete. Gainera, sarrera berriak ere beharko genituzke, adibidez, *-a* morfema bikoiztu beharko genuke ergatiborako sarrera gisa, komuntadurako desbideratzeak errepresentatu ahal izateko. Hiztegi-mailako tartehizkuntzaren azterketa aurreragoko utzitako ikergeia bada ere saio batzuk ere egin ditugu, eta, honetarako, euskararako garatuta dagoen lematizatzailean hainbat aldaketa egin dugu.

Lematizatzailea: IV.20 irudian agertu ez arren, euskararako lematizatzailean (Aldeazabal et al., 94) egindako egokitzapenei ere interesgarri deritzegu ikasleen corpusaren lema desberdinen azterketarako. Jatorrizko lematizatzaileak hitz bakoitzeko lema bakarra lortzea zuen helburu. Honetan testuak euskaraz adituak diren euskaldunek idatziak suposatzen ziren, hortaz, desbideratze gutxiko testuak. Hau dela eta, jatorrizko lematizatzaileak, esaterako, hitz baten analisi posibleen artean estandarren bat bada, desbideratze diren analisiak automatikoki baztertzen ditu. Honelako filtroak, berriz, ezin ditugu aplikatu hizkuntza ikasten ari direnen testuen azterketan. Horregatik, lematizatzailearen hainbat filtro kendu behar izan dugu, eta lema baztertzeko unetan heuristikoak aldatu.

Hortaz, hitz baten analisisian ondoko heuristikoak hartu dira kontuan:

- a) Interpretazio guztiak estandarrak badira, denak mantentzen dira.
- b) Interpretazio guztiak desbideratzeak badira, denak mantentzen dira.
- c) Batzuk estandarrak eta beste batzuk desbideratzeak direnean, estandar guztiak mantentzen dira eta desbideratzeak dituztenen artean batzuk aukeratu dira ikaskuntza-mailarentzat definituta dagoen filtro baten arabera.

Aipatutako filtroa aplikatu ondoren ondoko heuristikoak jarraitu ditugu hitzaren lema bakarrarekin gelditzeko:

1. Interpretazio estandarrez gain desbideratze-kopuru (kopurua ≥ 1) txikiena duten interpretazioak mantendu eta besteak baztertu.
2. Geratzen direnen artean lema eta kategoria bera dutenen artean bat aukeratu (analisiaren lehenengoa).
3. Lortzen direnetatik lema luzeena dutenak aukeratu.
4. Lema luzeeneko bat baino gehiago balira, kategoria gramatikala desberdina izanez gero aukeraketa testu-motaren arabera (adib. deskribapena \rightarrow adjektiboa) egin, posible balitz.

5. Gelditzen diren etik lehenengoa hautatu.

Azkenik aipatu nahi dugu ikasleen corpusetik abiatutako hiztegi-azterketan oso garrantzitsua dela hitz ezezagunen analisia. Ikaste-prozesuan ikasi eta sortzen diren hitz berriak analizatzen duen formalismoa izateak helburu behar luke honelako azterketa batean. Gure kasuan, lematizatzailea lema ezezagunak detektatzeko gai da, analisi morfologikoan lexikorik gabeko hitzak analizatzen baitira (Aldeazabal. et al., 94; Black et al., 91).

IV.4.2 Ikaskuntza-mailari dagozkion tresnen ezagutzaren eguneraketa sistemaren autoikasketa dela medio.

IV.12 irudian adierazten denez modelizazio-prozesuan tartehizkuntz eredua sortzeaz gain prozesuan parte hartzen duten tresnen ezagutza ere eguneratzen da. Ondorioz, sistemak berak autoikasketa automatikoa burutzen duela esan dezakegu. Ikaskuntza-mailaren tartehizkuntz eredua corpus desberdinen azterketaren ondoren osatzen da, hasierako abiapuntua goiko mailaren eredua izanik. Eredua hau behin eta berriro finduko du sistemak, aztertzen ari garen mailari dagokion eredua lortu arte, aldi berean honen hizkuntz tresnen ezagutza linguistikoa eguneratzen dela.

Sistemaren autoikasketa hiztegi-mailan eta maila morfosintaktikoan gertatzen da. Eguneraketa HITES sistemaren emaitzatzat har dezakegu. Psikolinguistak emandako iritzien arabera hainbat agindu sortzen du HITES sistemak berak. Agindu horiek tresna linguistiko-konputazionalen ezagutza eguneratzen dute. Lau ezagutza-mota eguneratzen da:

- Erregelen eguneraketa.
- Desanbiguatzaileen eguneraketa.
- Erregela morfosintaktikoen eguneraketa.
- Hiztegiaren eguneraketa.

• **Erregelen eguneraketa:** goiko mailaren analizatzaile morfologikoan inplementatutako automaten artean beheko mailan mantentzen direnak automatikoki gehitzen dira beheko mailaren tresnen ezagutzan. Mantentzen ez direnak ezabatu egiten dira eta aldaketa behar dutenak psikolinguistarentzako sistema elkarreragiletik jasotako informazioaz osatzen dira. Psikolinguistaren partehartze honetan sistemak goiko mailaren eta aztertzen ari denaren arteko automaten baliokidetzataula sortzen du. Taulako automaten arteko erlazioak hiru motatakoak izan daitezke: automata berdinak, automata batetik edo batzuetatik abiatutako egokitapena/k eta automata berria, hau da, goiko mailan baliokiderik ez duena. Aurreko azpiatalean azaldutako *ereduak_lortu* izeneko programak baliokidetzataula erabiltzen du bi hizkuntz mailen arteko konparaketak egiteko.

• **Desanbiguatzaileen eguneraketa:** hitzen interpretazioak filtratzen dituzten desanbiguatzaileek irizpide desberdinak jarraitzen dituzte ikaskuntza-maila

bakoitzean. Irizpide horiek, besteak beste, gehienezko desbideratze-kopurua interpretazio batean, psikolinguistarekiko elkarrizketaren arabera eguneratzen dira.

- **Erregela morfosintaktikoen eguneraketa:** badira erregelak kategoria gramatikala konkretuekin gertatzen direnak. Une batean ikasleak kontsonantez bukatutako hitzen deklinabidea izenekin bakarrik ikasten du, eta aurrerago izen eta adjektiboekin lantzen du. Honen ondorioz, hasierako egoeran adjektibo den interpretazio batean *-e* epentetikoa detektatzen duen erregela morfofonologikoa agertuko balitz, interpretazio hori ezabatu beharko genuke. Hau egiten duen erregela morfosintaktikoa sistemak sortzen du automatikoki eta ikaskuntza-mailaren tresnen ezagutzan gehitzen du. Erregela honelakoa litzateke:

REMOVE (ADJ) IF LEGTE;

Batzuetan, goiko mailan zegoen erregela baten egokitzapena izaten da gehitu beharreko erregela, psikolinguistak esandakoaren arabera noski. Hauek detektatzen dira askotan analisi ez-egokia egiten duelako sistemak eta psikolinguistari, ez-egokizat markatzean, aktibatutako erregelak egunera ditzan eskatzen zaio.

- **Hiztegiaren eguneraketa:** lematiztzailea azaltzean EDBLn hainbat ezaugarri berriren beharra aipatu dugu. Sistemak, ikaskuntza-mailaren arabera, ezaugarrien balioak eguneratzeko aginduak automatikoki sortzen ditu. Hala eta guztiz ere, datu-base lexikalaren diseinu-eguneraketak diseinu eta azterketa sakonagoa eskatzen du tartehizkuntzaren hiztegia aztertzeari dagokionez —kontuan izan 70.000 sarrera inguru dituela datu-baseak eta, zeregin desberdinetarako erabiltzen ari den tresna dugunez, ezin dela diseinua aldatu laboratoriko esperimenterako—. Hori dela eta, HITES sistemak oraingoz Murriztapen Gramatikak erabiltzen dituen erregela-motak sortzen ditu, ikaskuntza-mailari ez dagozkion lemak baztertu ahal izateko hitz baten analisisien artean. Adibidez, *arraio* hitzaren interpretazioetako lehenengoak (ikus IV.24 irudia) ez luke agertu behar ikaskuntza-maila ertainetan, esaterako. Hortaz, *jarraiki* lema *ADT* kategoriadunak ez agertzeko, hiztegiaren sarrera gisa honelako erregela sortzen da:

```
MS_HIZT_1: REMOVE <"jarraiki">;  
  
"<arraio>"  
  "arraio" IZE ARR ... @OBJ @SUBJ  
  "arraio" IZE ARR ... ZERO  
  "jarraiki" "jarraiki" ADT A1 ... LEM_LEKHH ... (agertuko ez den interpretazioa)
```

IV.24 irudia. *arraio* hitzaren analisia.

Kasu honetan *jarraiki* aukera ezabatu da aztertzen ari den ikaskuntza-mailan ezezagun delako. Adibide honen tratamendua psikolinguistarekiko elkarrekintzan azalduko da zehatzago § V.7.2-n.

IV.5 Analizatzaile morfologikoaren egokitzapena.

Esan bezala, euskararen tratamendurako hainbat tresna linguistiko-konputazional egokitu dugu tartehizkuntzaren modelizazio-prozesuan. Besteak beste analizatzaile morfologikoa dugu berrerabilitako tresnetako bat. Bi Mailatako Morfologian (Koskenniemi, 83; Agirre et al., 92; Alegria et al., 96) oinarritutako analizatzailean burutu ditugun berrikuntzak azalduko ditugu ondoren.

Nagusiki, hiru arrazoi dira aldaketen beharra sortu dutenak:

1. Euskara bigarren hizkuntza bezala ikasten duenaren hainbat desbideratze-egitura eta euskara hizkuntzan aditua denarenak desberdinak izaten dira. Aplikatutako testuinguruak, esaterako, hitz-mailako egiturazko baldintzak, aldatzen dira kasuak kasu.
2. Jatorrizko hiztun adituarentzako tresnaren helburua zuzentzaile ortografikoa genuenez, hainbat desbideratze erregela berean konpilatzen zen, eta, hala, hiztunaren benetako antolamendua baztertzen zen; ikuspegi linguistiko-konputazionalak agintzen zuen ezagutzaren antolamenduan. Hizkuntza ikasten ari direnen kasuan, ordea, hainbat erregelak berrantolaketa behar du erregela gehiago sortuz, fenomeno zehatzagoak tratatu ahal izateko.
3. Hasierako inplementazioan erregelen aktibazio-testuingurua anbigua gelditzen zen. Testuingurua detektatzea tartehizkuntzaren modelizazioan garrantzitsua da, eta horregatik testuinguru zehatzak detektatzeko aldaketak ere egin behar izan dira, IV.25, IV.26 eta IV.27 irudietan ikusiko denez.

Hortaz, modelizazio-prozesuan ikasleak erabiltzen dituen egiturak, estandarrak eta desbideratzeak, eta hauen testuinguruak detektatzea dugu helburu (Díaz et al., 97). Eta honen arabera egokitu dugu analizatzailea. Egindako aldaketak bi motatakoak dira:

- Analizatzailearen ezagutza linguistikoa eguneratzea, hau da, fenomenoak inplementatzen dituzten erregelak, gure kasuan automatik, berritzea.
- Aplikatutako erregelak eta testuinguruak detektatzeko aukera gehitzea analizatzailean bertan.

Ezagutzaren egokitzapenean kontuan izan behar dugu ikaskuntza-maila bakoitzak dagokion automata behar duela izan. Goitik beherako metodologia aplikatzean tresnen ezagutza eguneratzen denean, besteak beste, sortzen da automata (ikus IV.12 irudia). Guk, orain, hiztun adituarentzako analizatzailea (Agirre et al., 92) eta ikaste-prozesuaren

ikaskuntza-maila gorenean lortutako analizatzailearen (Maritxalar et al., 96) arteko konparaketa egingo dugu jatorrizko analizatzailea eta egokitutakoaren arteko diferentziak azaltzeko.

Erregela morfofonologikoak dira automaten bidez inplementatzen direnak dagokigun formalismoan (Koskeniemi, 83). Hasierako analizatzailean 30 automata genituen: 11 erregela estandarrak analizatzeko, hauen aktibazioa ez zen detektatzen, eta 19 desbideratzeak deskribatzeko, hauek detektatu arren ez zeuden tipifikatuak analizatzailearen irteeran. Ikaskuntza-maila gorenerako analizatzaile egokituan 61 automata ditugu, 61 fenomeno desberdin identifikatuz, bai estandarrak, bai desbideratzeak. Aplikaturako erregelak eta testuinguruak detektatzen dira. Automata-kopurua handitzeko bi arrazoi nagusi daude: desbideratze-mota berrien agerpena ikaskuntza-maila gorenean, eta hasierako automaten banaketa era espezifikoagoan hainbat fenomeno tratatatu ahal izateko. Hau ilustratzeko adibide bat azalduko dugu ondoren.

h:0 => R:=(+:=)+ _ (Fenomeno estandarrerako erregela)	<i>Adib. beR+hasi ->berrasi</i>
h:0 => [V:V (0:*) \$:\$] _ V:V (Desbideratze-fenomenorako erregela)	<i>Adib. hau ->*au / mehe ->*mee</i>

IV.25 irudia. h letraren ezabatzea tratatzeko erregelak hasierako analizatzailean.

Fenomeno estandarraren erregela ez da detektatzen, desbideratze-fenomenoa, ordea, identifikatu egiten da testuingurua anbigua geratzen delarik.

h:0 => R:=(+:=)+ _ (LEKBH: H LEtra Kendu lemaren Bukaeran)	<i>Adib. beR+hasi ->berrasi</i>
h:0 => V:V _ V:V (*LEKTH: H LEtra Kendu hitzaren Tarteko edozein tokitan)	<i>Adib. mehe ->*mee</i>
h:0 => (0:*) \$:\$ _ V:V (*LEKHH: H LEtra Kendu lemaren Hasieran)	<i>Adib. hau ->*au</i>

IV.26 irudia. Erregelak egokitutako analizatzailean.

Tartehizkuntzaren hiru erregelak eta hauen testuingurua detektatzen dira aktibatzen direnean. Baina nola jakin erregela aplikatu den ala ez erregelaren lexiko eta azaleko maila bera denean?. Automata finitua zena irteeradun automata bihurtu dugu. Automata bera hirukoizten dugu eta automata bakoitzean erregelari dagokion aktibazio-testuingurua definitzen duten trantsizio-egoerak markatzen ditugu. Automata egoera horretatik pasatzean informazioa gordetzen dugu, ondoren automataren aktibazioa kanporatzeko.

Ondoren aurreko bi erregela desbideratzedunei dagozkien automatak azaltzen dira:

"22.LEKHH." 6 7	"23.LEKTH." 6 7
h h R + V 0 =	h h R + V 0 =
0 === V * =	0 === V * =
1: -5 4 2 4 6 1 4	1: 5 4 2 4 6 1 4
2: 0 4 2 3 6 4 4	2: 0 4 2 3 6 4 4

3: 4 4 2 3 6 4 4	3: 4 4 2 3 6 4 4
4: 0 4 2 4 6 4 4	4: 0 4 2 4 6 4 4
5: 0 0 0 0 1 0 0	5: 0 0 0 0 1 0 0
6: 5 4 2 4 6 1 4	6: -5 4 2 4 6 1 4

IV.27 irudia. LEKHH eta LEKEH erregelen automatak.

Adibidean ikusten denez h letraren galera gertatzen den egoera markatu dugu (minusa jarritz) kasu bakoitzean testuinguru desberdinak identifikatuz. Dena den, soluzioa ez da egokia fenomeno guztietarako, kasu batzuetan gertatzen baita bi erregela desberdinek testuinguru bera betetzen dutela (ikus IV.28 irudia). Adibidez, har dezagun *Azpeitia* hitza. Kasu honetan *Azpeitira* eta *Azpeitian* forma estandarrek diren bitartean **Azpeitiara* eta **Azpeitin* desbideratzeak kontsideratzen dira ikaskuntza-maila gorenetan. Ikus ditzagun *Azpeitia* (*Azpeiti&*) bezalako hitzei dagozkien erregela morfofonologikoak, hauek dira &¹ diakritikoa tratatzen dutenak.

LEKBA31 Erregela estandarra ->	&:0 <=>	_ (+:=)+ [Era:0ra Et:0t Eko:0ko ko:ko]
LEGBA31 Desbideratze erregela ->	&:a =>	_ (+:=)+ [Era:0ra Et:0t Eko:0ko ko:ko]
LEKBA32 Desbideratze erregela ->	&:0 =>	_ [+:= \$:\$ (+:=)+ EM:0n]

IV.28 irudia. & diakritikoa duten hitzei dagozkien erregela morfofonologikoak.

Erregeletan ikusten denez, alde batetik, estandarra eta lehenengo desbideratzearen testuinguruak berdinak dira, eta, beste alde batetik, estandarra eta hirugarren desbideratzearen osagaiak berdinak dira (&:0) eta testuinguru batzuen hasiera berdina da ((+:=)+ E:0). Honelako fenomenoak ezin dira tratatu aurreko adibideko automatetan azaldutakoa aplikatuz. Hauetan osagaiak eta testuinguruaren egoerak markatuko ditugu egoera berezi hauetatik automata zenbat alditan pasako den kontuan izanik. Era honetan zein erregela ari den aplikatzen detektatzeko gai izango gara. Ondoren aurreko hiru erregelak inplementatzen dituen automata azalduko dugu.

¹& diakritikoaren esanahia: lekuzko izen propioetan galtzen den -a mugatasun-adierazlea.

"61.LEKBA31, LEGBA31, LEKBA32." 8 11
& & + E r a t k o = M
0 = = 0 r a t k o = n
1: -2 7 1 1 1 1 1 1 1 1 1
2: 0 0 3 0 0 0 0 0 0 0
3: 0 0 3 4 0 0 0 -6 0 0
4: 0 0 0 0 -5 0 -1 -6 0 0 1
5: 0 0 0 0 0 1 0 0 0 0
6: 0 0 0 0 0 0 0 0 1 0 0
7: -2 7 7 8 1 1 1 -1 1 1 1
8: -2 7 1 1 -1 1 -1 -1 1 1 1

IV.29 irudia. LEKBA3 erregelaren automata.

Automatan (IV.29 irudia) ikusten denez ikasleak **Azpeitin* dioenean, hitzaren lema bukaerako azken karakterea tratatzean (**Azpeitin = Azpeiti + n*) 1 egoeratik (&:0) 2ra pasatzen da eta hemendik 3ra. Gero, 4 egoeratik pasa eta handik 1era doa. Honek esan nahi du bi desbideratzeetako bat detektatu dugula. Lemaren bukaera aztertuz gero LEKBA32 dela jakin dezakegu. **Azpeitiara* hitza erabiliz gero (**Azpeitiara = Azpeitia + ra*) 1 egoeratik 7ra doa eta handik 8ra. Kasu honetan lema -a letraz bukatzen denez LEGBA31 erregela detektatzen da. Azkenik, *Azpeitira* esango balu ikasleak, LEKBA32 erregela identifikatu luke automatak; 2, 3, 4 eta 5 egoeratik pasa ondoren.

Jakina, aurreko bi adibideetan azaldutako fenomenoak programak detektatu ahal izan dezan, hainbat aldaketa egin da analizatzaile morfologiko egokituari.

IV.6 Inplementazio-egoera eta landu beharrak.

Linguistika konputazionalari dagokionez, nagusiki bi formalismo erabili ditugu azaldutako tresnetan: Bi Mailatako Morfologia (Koskenniemi, 83) eta Murriztapen Gramatika (MG) (Karlsson et al., 95). Programazio-tresnak, ordea, honakoak izan dira: C, Perl (Randal, 91), eta UNIX sistema eragileak corpusen azterketarako, besteak beste, eskaintzen duen zenbait aukera. Euskararako Datu-Base Lexikala dela eta, diseinu-proposamen batzuk aipatu ditugu. Guk uste, bide honi jarraituz, azterketa sakonagoa interesgarria litzateke. § IV.4-n esan bezala, inplementaziorako, oraingoz MG formalismoa erabili dugu hiztegia lantzeko, EDBLn aldaketak egin beharrean. Dena den, ikaskuntza-maila bakoitzaren hiztegi azterketa sakona egiteko datu-base lexikala garatzea beharrezkotzat jotzen dugu. Sarrerren ezaugarrien artean testuinguru-aplikazioa, hots, egiturazko testuingurua eta gai-testuingurua, eta ikaskuntza-maila gehitzea proposatzen dugu. Honetaz gain, sarrerren definizioak datu-base lexikalean gehitzea eta azterketa semantikoak lantzea etorkizuneko ildoak ikusten ditugu.

Tartehizkuntz ereduen lehen azterketa honetan bi ikaskuntza-maila tratatu ditugu nagusiki: hiztun adituarena eta ikaskuntza-maila goreneko ikaslea. Honetarako morfologi mailan automata bana landu ditugu. Hiztun adituarentzako automata dagoeneko garatuta zegoenaren (Agirre et al., 92) egokitzapena da. Bertsio berrian edozein automataren aktibazioa detektatzen da, erregela estandarrak barne. Egokitzapena egin behar izan da bi hizkuntz mailen arteko konparazioak automatikoki egin ahal izateko. Hasiera batean ikaskuntza-maila goreneko automata eskuz egin behar izan genuen ez zegoelako oraindik sistema garatua. Une honetan ikaskuntza-maila berrien automatikoki erdiautomatikoki sor genitzake. Oraingoz, ordea, § IV.4.2-n azaldutako lema eguneraketan aipatu baliokidetzataularen sorrera erdiautomatikoki egiten dugu. Hau hobetzeaz gain, erregelen testuinguruen definiziotik automatikoki sortzen duen Xeroxen tresna (Karttunen et al., 92) integratzeari egoki deritzogu. Honela, psikolinguistaren interakziotik zuzenean tresnen ezagutza linguistikoa automatikoki egunera genezake. Automaten kasuan ez bezala, desanbiguatzaileetarako erregela berriak landu beharko lirateke oraindik maila gorenean, hainbat esperimendu egin dugun arren.

Erregela morfofonologikoen detekzioan kodea eta erregelaren azalpena lortzen da automatikoki. Erregela morfosintaktikoen kasuan, ordea, zenbaki kontsekutiboak asignatzen zaizkio ezagutza-basean, oraingoz erregela hauek ez baitira kodetu. Garatze-prozesuan dagoen modulu honetan kodeketa modularizatu baten beharra ikusten dugu. Dena den, une honetan euskararen tratamendu orokorrerako modulua abiapuntu hartzen dugu, horren gainean erregelen aldaketak eginez. Horregatik, ez dira erregela morfosintaktikoen kodeak eta azalpenak agertzen.

Azaldutako tresnak, hau da, analizatzailea, desanbiguatzaileak, *Estatistikak* sistema elkarreragilea, laginak, eta beste guztiak, bukatutzat eman daitezke, batzuetan ezagutza linguistikoa landu behar dela ahaztu gabe. Corpusaren bidezko ereduen azterketa aipatu tresnen bidez egin dugu, VII. kapituluan ikusiko dugunez.

Hala eta guztiz ere, tartehizkuntza kontzeptuaren definizioa dela eta, badago landu beharreko puntua, bereziki, garatutako tresnetan. Tartehizkuntzan desbideratzeak ikaslearen ezagutzaren partetzat hartzen dira. Horregatik, hitzen analisisian kopurua nabarmenki biderkatzen da. Ondorioz, desanbiguatzaileei lan handia uzten zaie. Honek guztiak eranginkortasunean ondorioak dakartza. Hortaz, dimentsio errealeko aplikazioei begira tresnen azkartasuna landu beharko litzateke. Honetarako aztertu beharreko bi tresna ikusten dugu: Xeroxen *Twolc* eta *XFST (Finite State Tools)* tresnak. Hauetan aurkitzen dugun arazo handiena erregelen aplikazioaren detekzioan datza. Automata guztiak itzultzaile bakarrean bihurtzen direnez ez dago hain garbi egin daitekeenik. Oraingoz, eta lehenengo hurbilpen gisa, bi aukera egon daitezke:

- *Twolc* tresnaren (Karttunen et al., 97) kasuan lexiko-mailan beharko genuke erregelei buruzko informazioa. Honek eskatuko luke, analisi morfologikoa egin aurretik, lexikoan hitz guztiak aldeztu aurretik sortuak izatea, erregelei buruzko informazioa hitzetan dagoela; esaterako **suhaitz* desbideratzea hiztegian izango bagenu aldeztu aurretik ondokoa topatuko genuke: **LEOEZSuhaitz*. Kontuan izan behar dugu hiztegiaren sorkuntza automatikoki egin litekeela, beti ere, muga batzuk jarritz. Edozein kasutan hitz guztiak sortuak gordetzeak diskoen edukieran arazoak ekar litezake, baina bakarrik konpilazio garaian.
- *Finite State Tools* (Karttunen et al., 97) izeneko tresna erabiliz gero bideragarri ikusten dugu ordezkapen erregeletan erregela esplizituki jartzea lexiko-mailan. Honek eskatuko luke erregela guztien testuinguruan erregela guztiak jartzea (erregela bera ezik) lexiko-mailan ordurarte agertu diren beste erregelak kontuan ez izateko azalduz. Etorkizunean aztertu beharreko aukera iruditzen zaigu honako hau.

Azkenik, esan beharra dago testuinguruen detekzioa hitz-mailan automatikoa dela, baina esaldi-mailan erdiautomatikoki lantzen dela, psikolinguistaren laguntzaz, horregatik IV.23 irudian ez da testuinguruei buruzko informazioa agertzen. Sintaxi maila ere era apalean tratatu da, eta etorkizunean euskararako garatzen ari den analizatzaile sintaktikoa (Gojenola, 99) integratzeko asmoa dugu ereduaz azterketan.

V. HITES: bigarren Hizkuntzaren Ikaste-prozesuan bereganatzen den Tartehizkuntz Eredua modelizatzeko Sistema adimenduna.

Kapitulu honetan HITES sistema azalduko dugu: bigarren hizkuntza ikasten ari diren ikasleek duten Tartehizkuntzaren konpetentzia gramatikalaren modelizaziorako sistema adimenduna da. Tartehizkuntza bigarren hizkuntza ikasten ari den ikasleak bere ikaste-prozesuan zehar eskuratzen duen egitura linguistikoaren multzoa genuke. Prozesuan zehar aipatu egiturak aldatuz doaz eta gainera gauzatzen diren testuinguruaren arabera ere aldatuz doazenak dira, (Selinker, 92).

HITESek helburu hirukoitza du: lehena, ikaskuntza-maila berean dauden ikasleek jaso duten ezagutza komunaren errepresentazioa lortzea du xede (ezagutza kolektiboa); bigarrena, ikasle jakin baten tartehizkuntzaren errepresentazioa ere lortu nahiko luke (ezagutza berezitua) eta, hirugarrena, ikaskuntza-maila bati dagozkion tresna linguistikoetan erabilitako ezagutza linguistikoaren finketa.

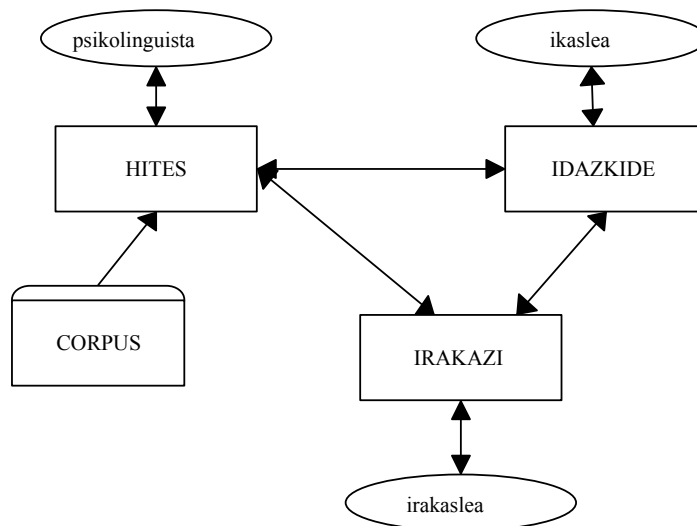
Kapitulu honen xedea HITES deskribatzea bada ere, § V.2-n HITES garatu deneko MUGARRI ingurunean kokatu asmotan IDAZKIDE eta IRAKAZIri buruz aski orokorra den hainbat ideia laburki aditzera emango dugu. Ondoren, HITES domeinuaren errepresentazioa nolakoa den azaldu eta gero, sistemaren arkitekturaren deskribapen zehatzagoari helduko diogu sistemaren ikaskuntza-moduluekin osatuko dugularik. Jarraian aurkeztera goazen lan honetan agente ikuspegiaren erabilerari buruzko hausnarketa egingo dugu. Eta amaiera gisa, HITESen inplementazioaz jardungo dugu, objektuei zuzendutako programazio-paradigmaren baitan burutuko dugularik.

V.1 Sarrera.

HITES autonomia den sistematizat hartzen dugu eta bere jokatzeko modua tartehizkuntza-ereduen eraikitze-prozesuan jasotzen den esperientziaren arabera izan ohi da, horretarako, batetik maila berean dauden ikasleen idatzizko testuen analisia hartzen da abiapuntu gisa, eta bestetik, baita psikolinguistak ikaste-prozesuaren maila desberdinetan tartehizkuntza-ereduei buruz jasotako informazioa bera ere. Sistemak goren-gorenko ikaskuntza-mailarako eredua eraikitzen du, eta ondoren goitik hasi eta mailaz-maila beherantz joaz behekoetan tartehizkuntza-ereduak eratuz jarraitu ohi du. (ikus § IV.3). Kolektiboa eta berezitua den ezagutzen eskuratze-sistema baten aurrean geundeke: lehen

kasuan arrazoibidea maila desberdinetako ikasleek idatzitako testuetan eta irakasleak nahiz psikolinguistak emandako informazioan oinarritua da; bigarren kasuan, berezitua den ezagutzaren jabetza ikasle jakin batek idatzitako testuetan, ikaste-prozesuan zehar nabarmentzen den hainbat ezaugarritan, eta irakasle nahiz psikolinguistak ikaslearen tartehizkuntz ereduari buruz emandako datuetan oinarritzen da.

HITES sistema IDAZKIDEREkin harremanetan dago, euskarazko idazketaren ikaste-prozesuan ikaslearentzat lagungarri gertatu ohi den laguntzaile elkarreragilearekin, alegia, eta bestalde, irakaslearentzat ikaslearen ikaskuntzan gertatu ohi diren prozesuez diagnostikoa egin dezan lagungarri den IRAKAZI sistema ere. Bi sistema hauek HITESekin komunikaturik daude (ikus V.1 irudia). Bai IDAZKIDE (Díaz et al., 98b) eta baita IRAKAZI ere VI. kapituluan landuko ditugu.



V.1 irudia. Hiru sistemen arteko komunikazioa.

V.2 HITES, IDAZKIDE eta IRAKAZI: elkarrekin harremanetan dauden sistema adimendunak.

V.2.1 Sistemen definizioa.

HITES, IDAZKIDE eta IRAKAZI sistemak, ikasle-talde batek, irakasleak eta ikasleen ikaste-prozesua aztertzen duen psikolinguistak osatzen duten MUGARRI ingurunean erabiltzen dira.

HITES bigarren hizkuntza ikasten ari diren ikasleek duten tartehizkuntzaren konpetentzia gramatikalaren modelizaziorako sistema adimenduna da. Sistemak ikasle jakin

baten tartehizkuntza aztertzeaz gain, ikaskuntza-maila berean dagoen ikasle-talde baten tartehizkuntza ere aztertzen du. Esan liteke beraz, HITESen saio bat hauxe dela:

- ikasle bakar bati dagozkion testu-multzoen azterketa, hain zuzen helburua ikasle jakinaren tartehizkuntz eredua eguneratzea baitu.

edo/eta

- ikaskuntza-maila berean dagoen zenbait ikasleren testu-multzoa aztertzea ikaskuntza-maila jakinari dagokion tartehizkuntz eredua eguneratzeko asmotan.

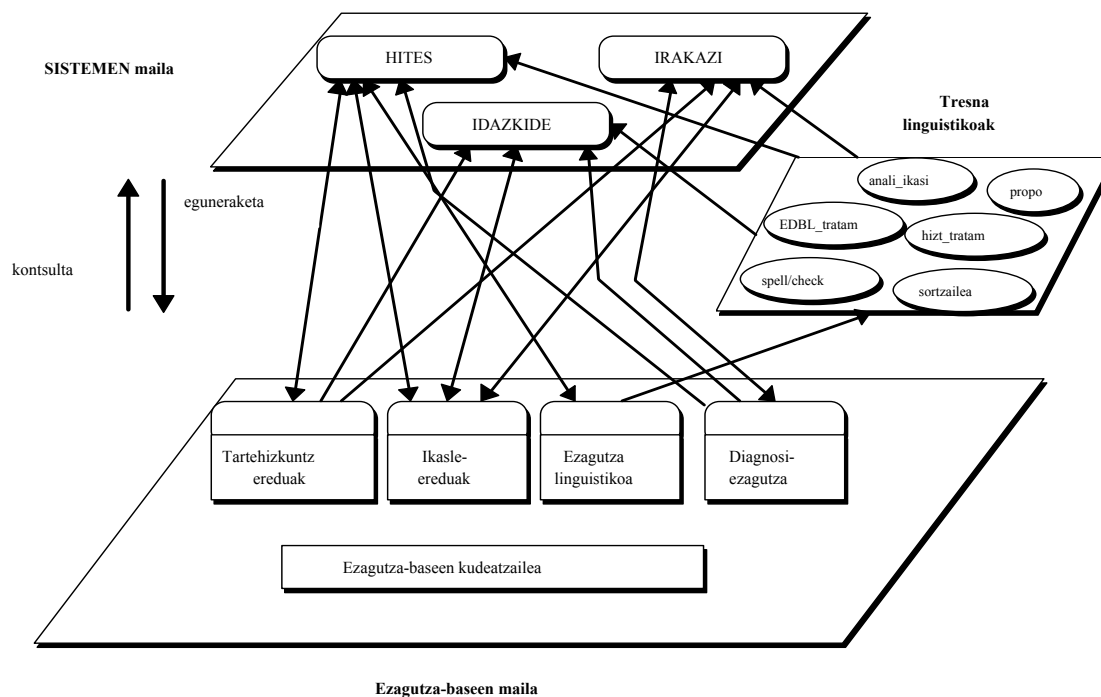
IRAKAZI ikaslearen ikaskuntzan gertatzen diren prozesuez diagnostikoa egiteko irakasleari lagungarri zaion sistema bat da. Horretarako HITESek sortutako tartehizkuntz ereduak erabiltzen ditu.

IDAZKIDE idazte-prozesuan zehar ikasleari laguntza ematen dion laguntzaile elkarreragilea da, eta HITESek sortutako ikaskuntza-maila baterako tartehizkuntz ereduatan nahiz IRAKAZIk emandako informazioan oinarriturik dago.

V.2.2 Ingurunearen arkitektura.

Esan bezala, azaltzekotan garen MUGARRI ingurunea hiru sistemek osatzen dute: HITES, IDAZKIDE, eta IRAKAZIk hain zuzen. Sistema hauek, alde batetik, elkarrekiko zuzeneko komunikazioa dute eta, beste aldetik, zeharkako komunikazioa burutzen dute ezagutza-baseak direla medio. Hiru sistemek konpartitzen dituzten ezagutza-baseak ondokoak dira: ikaskuntza-maila bakoitzari dagokion ezagutza linguistikoa, tartehizkuntz eredua eta diagnosi-ezagutza, eta ikasle bakoitzari dagokion ikasle-eredua. V.1 irudian, dagoeneko sistemen arteko zuzeneko komunikazioa errepresentatu dugunez, V.2 irudian zeharkako komunikazioa irudikatu dugu gesien bidez. Bertan hiru sistemek bere ataza egoki burutzeko erabiltzen dituzten tresna linguistikoak ere agerian jartzen zaizkigu. Aipatu tresnak sistemen maila eta ezagutza-baseen mailaren artean irudikatu ditugu.

Irudian agertzen denez, sistemak eta ezagutza-baseen arteko komunikazioaz arduratzen den ezagutza-baseen kudeatzailea definitu dugu. Ondoren, ezagutza-baseen maila azalduko dugu labur. Esan bezala, arkitekturaren maila honetan lau ezagutza-base definitu ditugu: ezagutza linguistikoa, tartehizkuntz ereduak, diagnosi-ezagutza eta ikasle-ereduak.



V.2 irudia. MUGARRI ingurunearen arkitektura.

Ezagutza linguistikoa ikaskuntza-mailaren baitan antolatuta dago. Ezagutza kolektiboaren zeharkako errepresentazioa da: tresna linguistiko-konputazionalentzat beharrezkoa den ezagutza linguistikoa definitzen du.

Tartehizkuntz-ereduak ikaskuntza-mailaren baitan eraten dira eta maila bakoitzean dauden ikasleek duten ezagutza komuna adierazten dute.

Diagnosi-ezagutza errepresentatzen duen moduluak ikaskuntza-maila baten ikasleek bere tartehizkuntz komunaren egitura linguistiko desbideratuak erabiltzeko eduki ditzaketen arazoak gordetzen ditu.

Ikaslearen ereduak ikaslearen tartehizkuntz-ereduen berri eta beren ikaskuntzaren ezaugarrien berri ematen duten eredu generikoak genituzke.

V.2 irudiak dagerzkigun tresna linguistikoak dira hiru sistemek gehien erabiltzen dituztenak. IV. kapituluan zehazkiago garatu ditugu tresna hauek. Atal honetan beraz, V.2 irudian agertzen zaizkigunak baizik ez ditugu gogora ekarriko:

- anali_ikasi: analizatzaile morfosintaktikoa.
- propo: proposatzaile morfologikoa.
- EDBL_tratam: lexikoaren informazioari buruzko kontsulta-interfazea.
- hizt_tratam: lexikoaren definizioari buruzko kontsulta-interfazea.
- spell/check: hitz-mailan akatsen detektatzaile/zuzentzailea.
- sortzailea: deklinabideen sortzailea.

V.2.3 Hiru sistemen arteko komunikazioa.

HITES, IRAKAZI eta IDAZKIDE elkarren osagarri diren helburuak dituzten hiru sistema dira, eta zenbait ezagutza-base ere komunak dituzte. Hala nola ezagutza-base hauek direla medio lortzen da zeharkako komunikazioa, V.2 irudian adierazi dugun bezala. Zeharkako komunikazioaz gain, zuzeneko komunikazioa ere topatzen dugu HITES, IDAZKIDE zein IRAKAZIren artean.

Ondoren, hiru sistema hauen arteko komunikazioaren helburuak aipatuko ditugu. Kasu guztietan bi zentzuko komunikazioa egon daiteke, hau dela eta kasu bakoitzeko bina helburu nagusi aipatuko ditugu.

HITES eta IDAZKIDEn arteko komunikazioa:

- HITESek IDAZKIDEn informazioa eskatzean helburua ondokoa da:

Ikasleari idazketan aholku ematerakoan IDAZKIDEn erabaki-hartzeak hobetzea.

IDAZKIDEk ikaslearekiko elkarrekintzan egitura linguistikoak lantzeko ariketak proposatuko dizkio HITESek eskatuta, azken honek corpusetan detektatutako zenbait egituraren ulermena lortzeko, hau da, tartehizkuntz egiturak osatu ahal izateko.

- IDAZKIDEk HITESi informazioaren berri ematean helburua ondokoa da:

HITESen tartehizkuntz ereduaren eraketa osatzea.

HITESek IDAZKIDetik jakingo du ikasleak bere testuen zuzenketan zein aukera egin duen, eta honek bide emango digu aztertzen ari garen ikaslearen ikaskuntza-mailaren tartehizkuntz ereduaren IDAZKIDEk detektatu dituen egitura linguistiko berri horiek gehitzeko.

HITES eta IRAKAZIren arteko komunikazioa:

- HITESek IRAKAZIri informazioa eskatzean helburua ondokoa da:

HITESen tartehizkuntz ereduaren eraketa osatzea.

HITESek IRAKAZItik jakingo du ikasleak ikasgelan erabiltzen dituen egitura linguistikoak corpusetan agertzen ez direnak, eta honek bide emango digu aztertzen ari garen ikaslearen ikaskuntza-mailaren tartehizkuntz ereduaren irakaslearen bidez detektatu diren egitura linguistiko berri horiek gehitzeko.

- IRAKAZIk HITESi informazioaren berri ematean helburua ondokoa da:

HITESek tartehizkuntzaren ezagutzari buruzko informazioa xehetasun handiagoz hornitzea.

IRAKAZIk egitura horiek gauzatzen diren testuinguruari buruzko informazioa are xehetasun handiagoz emango dio, gero HITESek tartehizkuntzari buruzko ezagutza egunera dezan.

IDAZKIDE eta IRAKAZIren arteko komunikazioa:

- IRAKAZIk IDAZKIDERI informazioa eskatzean helburua ondokoa da:

IRAKAZIk erabiltzen duen diagnosi-ezagutza fintzea.

IDAZKIDEk ikaslearekiko elkarrekintzan egitura linguistikoak lantzeko ariketak proposatuko dizkio IRAKAZIk eskatuta, azken honek irakasleak proposatutako zenbait egituraren ulermena lortzeko, hau da, irakaslearen diagnostikoa fintzeko.

- IDAZKIDEk IRAKAZIri informazioaren berri ematean helburua ondokoa da:

IRAKAZIk erabiltzen duen diagnosi-ezagutza fintzea.

IDAZKIDEk bere saioetan ikaslearekin izandako prozesuaren trazek egitura desbideratuen diagnosirako gakoak eman ditzakete. Hortaz, IDAZKIDEk informazioa hori bidaliko dio IRAKAZIri honen bidez irakasleak diagnostiko zehatza egin dezan.

V.3 HITESen domeinuaren definizioa.

HITESen domeinua tartehizkuntz ereduaz eta ikaslearen ereduaz osaturik dago. Kapitulu honetan tartehizkuntz ereduaz arituko gara gehienbat, ikaslearen ereduari buruz, berriz, ideia orokor batzuk bakarrik emango ditugu, honi buruz VI. kapituluan zabalago arituko garelako. Han diagnosi ezagutza ere azalduko dugu, IDAZKIDE sistema lagungarriarekin batera.

V.3.1 Tartehizkuntz eredu.

IV. kapituluan xeheki asko garatua dugu domeinuaren eredu kontzeptuala. Atal honetan beraz, domeinuaren definizioa eta berau gidatzen duten printzipio nagusien berri baizik ez dugu aditzera emango.

Tartehizkuntz eredu testuinguru jakin bat duen egitura-multzo bat da, hain zuzen fenomeno linguistikoez osaturik dago, eta era berean fenomeno linguistiko hauek erregela linguistikoek osatzen dituzte. Garatzen ari garen tartehizkuntzaren modelizazioa idazketa-mailara mugatu dugu eta hauek izan dira kontuan hartu ditugun testuinguru edo baldintza-motak (ikus § IV.2.2):

- Baldintza linguistikoak (kategoria gramatikala, menpekoen erabilera...)
- Baldintza tematikoak (orokorra, tekniko, ...)
- Egiturazko baldintzak (testu-mota, esaldien luzera, ...)
- Ariketa-baldintzak (testu baten sorkuntza, testu baten laburpena, ...)

Tartehizkuntzaren errepresentazioaz azaltzen ari garen gure proposamenean finkatua eta aldakorra den ezagutzaren arteko desberdintasuna kontuan hartu badugu ere, fenomeno

linguistiko bakar baten aurrean ikasle berak izan ditzakeen aldaera desberdinak ere adieraziko ditugu. Aldaera hauek fenomeno bakar bat adierazteko modu desberdinak izan ohi dira, eta litekeena da inoiz aurkari izatera heltzea ere. Adibide bat aipatzearen, ergatiboaren kasua dugu. Ikasleak testu berean ergatiboari -k ikurra jarri izana suerta liteke eta hurrengo batean haatik deus jarri ez izana.

Ezagutza pilatzeko moduari dagokionean konpilatua edo ez-konpilatua izan daitekeela esaten dugu eta bereizketa hau ezagutza hau jaso dugun moduaren baitakoa izan ohi da.

HITESen tartehizkuntz eredia diseinatzerakoan gidari bezala proposatzen ditugun printzipiorik nagusienak hauexek ditugu: dinamismoa, aldakortasuna, konpilagarritasuna, monotoniarik eza eta ikuspegi anitza (ikus § IV.2.1).

V.3.2 Ikaslearen eredia.

MUGARRIren ikasle-eredua ikaslearenak berarenak diren ezaugarriek eta bere tartehizkuntz ereduak osatzen dute. Gaian are gehiago barneratuz esango genuke ikasle-eredua batetik estatikoak diren ezaugarriek osatzen dutela (adib. ama-hizkuntza, bigarren hizkuntzaren ikasketari ekindako data, hasiera-hasieratik ikaste-prozesuak ezagututako historia, hizkuntzaren ikasketara bultzatutako arrazoiak, ...) baina bestetik ezaugarri dinamikoak ere baditu (adib. ikasketa-maila, ikaste-estrategiak, ...). HITES, IRAKAZI, eta IDAZKIDE sistemek ikasle-eredu bera konpartitzen dutela jo badugu ere, ezaugarri bakoitzak sistema desberdinekin bere-berea duen dinamismoarekin jokatzen du. Adibide bati helduz, HITESen ezagutza kolektiboa dinamikoa da, eta ezagutza hau bera IRAKAZI eta IDAZKIDEn estatikoa da aldiz.

Ikasleak berak garatzen dituen tartehizkuntz ereduak eta maila bereko gainontzeko ikasleengan komunak ez direnak, ikaslearen ereduaren barruan kokatuko genituzke. Ikasle-talde bati komunak zaizkion ezagutza kolektiboaren egiturak alabaina, ikaslearen eredutik at dagoen tartehizkuntz ereduaren moduluan gordetzen dira. Dena den, ezagutza kolektiboa adierazten duten ereduak ikaslea berarena den eredutik kanpo egon arren ikaskuntza-mailaren tartehizkuntz eredia ikasle jakin honen ezagutzaren zati ere izango da.

Ikaste-prozesua luzea denean tartehizkuntz egituren jabetze-mailan erreparatzerakoan ikaslea mailaz-maila aurreratuz doan heinean egituren progresioa ere aldatuz doala konturatzen gara. Egitura-multzo hauek dira hain zuzen ere ezagutza finkoa eta ezagutza aldakorra osatzen dutenak. Maila bakoitzean finkoa den ezagutzaren azpitalde bat ikaskuntza-maila bat beheragokoa aurretik finkatua izaten da, baina gainontzeko guztia lantzen ari garen ikaskuntza-maila horretantxe bertan finkatutako ezagutza izango da (ikus IV.11 irudia). Tartehizkuntz eredia erabakitzeko beraz, eredia eraikitzerakoan identifikatu beharko dugu azpitalde hauxe. Aldakorra den ezagutzak halere, ikaskuntza-maila batetik

besterako urratsa ematerakoan oso bestelako bilakaera jarraitzen du. Maila jakin bateko ezagutza aldakorraren egiturak gerta liteke gainera dagoen hurrengo ikaskuntza-mailako ezagutza finkokoak izatea edota baita ezagutza aldakorretik eratorriak izatea ere. Azkenik, bukatzeko gainera dezagun, ikaskuntza-maila bateko ezagutza aldakorra maila berean sortu eta ondoren desagertzea ere litekeena dela.

Orain artean ikaslearen ereduari topatzen dugun ezagutza-motaz jardun dugu solasean baina beti ere ezagutza horren jabetze-mailan erreparatu garelarik. Azpimarratzekoa den beste alderdi bat ezagutzaz jabetzeko modua genuke: konpilatua edota ez-konpilatua izan bailiteke. Ikasleak konpilatua den modua erabiltzen duenean ikaslea ez da ohartu ere egiten egitura linguistikoak zein erregelez osaturik dauden. Ez-konpilatua den moduan ordea, ikasleak hainbat erregela kontzienteki segidan aplikatuz lortzen ditu egitura linguistikoak. Gure ustean bigarren hizkuntza ikasten ari den ikasleak ezagutza linguistikoa nola bereganatzen duen, hala garatuko du tartehizkuntz egitura-mota ere. Adibide bat aipatuz, goiko maila batean dagoen ikasleak honako hau esango balu: "*Etorri den aitona atsegina zen*" erlatibozko lehenaldiko esaldien eraikuntza seguraski, jadanik automatizaturik edo konpilaturik izango du, baina esaten duena egitura hau ikaste-bidean bada oraindik ere, ziurrenik erregela hauek izango dira erabiliko dituenak:

- erlatibozko esaldia -> aditza + -n
- -n-z bukatutako aditza -> -n + -n = -n

Beharagoko mailan den ikasleak egitura ez du konpilaturik.

Zenbait kasutan prozesua alderantzizkoa da, egitura konpilatua ikasten baita. Aditza alokutiboan (famili-erregistroan) ikastea esaterako adibide garbi askoa dugu, eskuarki ikasleak behin eta berriro errepikatzearen poderioz burutzen du eta beranduago heltzen zaio egiturari lotua dena deskonposatzeko joera, eta hau egiten duenean ezagutza linguistiko berri bat sortzeko ahalmena izango du. Gramatikaren barru-barruko erroetara jotzeko joera duen ikasle batek egiturak deskonpilatzera joko du eta bera bakarrik saiaturik da gainera ezagutza berri bat sortzen, aitzitik, errepikatzearen poderioz ikasten duen ikasleak ezagutzaren bertsio konpilatu hura besterik ez du ohikoan ikasiko.

Ondorio gisa, baieztatu dezagun ezagutza linguistikoaren jabetze-modua ikaskuntza estrategiekin zuzen-zuzenean dagoela erlazionaturik. Hau dela eta, beraz, gure proposamena zera da: domeinuaren ereduari ikaskuntza-maila bakoitzeko tartehizkuntz egiturak ikaste-estrategia bateratu batzuk dituzten ikasle-taldeetan arabera era daitezela.

V.4 HITESen arkitektura.

Atal honetan HITES sistemaren arkitekturaz arituko gara. Arkitektura osatzen duten osagaiak adieraztera eman ondoren, sistemaren ezagutza-baseak azaltzeari ekingo diogu. Horretaz gain, sistemaren arrazonatze- eta erabakitze- moduluaren funtzionamendua deskribatuko dugu, eta, azkenik, psikolinguistarekiko interfazearen ezaugarri nagusiak aipatuko ditugu.

V.4.1 HITESen arkitekturaren osagaiak.

Sistema oro baliabide eta ekintzen bidez ingurunearekin elkarrekintzan aritzen da. Gure kasuan, HITESek bigarren hizkuntza ikasten ari diren ikasleen output-a jasotzen du (*ikaskuntza-maila jakin batean diren ikasleen testu-corpusa*). Corpus hau abiapuntutzat hartuz HITESek erdiautomatikoki sortuko du dagokion tartehizkuntz eredua, izan ere elkarrekintzaren bidez psikolinguistak horretan lagunduko baitu. Elkarrekintza hauei esker sistemari tartehizkuntz egiturei buruzko informazio zehatzagoa eskaintzen zaio. Prozesu honetan zehar psikolinguistak badu kontsultak burutzeko aukera, adibidez egitura bakar baten aldaera desberdinak zeintzu diren jakiteko, goragoko ikaskuntza-maila batean tartehizkuntzaren ezagutza finko eta aldakorra zein den jakiteko, etab.

HITES sistemaren arkitekturaren errepresentazioaren mailarik abstraktuenean (ikus V.3 irudia) osagairik garrantzitsuenak honako hauek dira:

HITESen arkitekturaren osagairik nagusienak honako hauek ditugu:

- Ezagutza-baseak.
- Arrazonatze- eta erabakitze-modulua.
- Interfaze-kontrolatzailea.
- Komunikazio-modulua.
- Ikaskuntza-modulua.

Ezagutza-baseak tartehizkuntz egituren detekzio-prozesua irudikatzen dute. Alde batetik, helburuak eta atazak ezagutza-basean helburuen egitura eta hauek burutzeko atazen egitura gordetzen da. Beste alde batetik, saio batean zehar sortzen diren helburu konketuak eta ataza konketuak gordetzen dira lan-memoriako agendan.

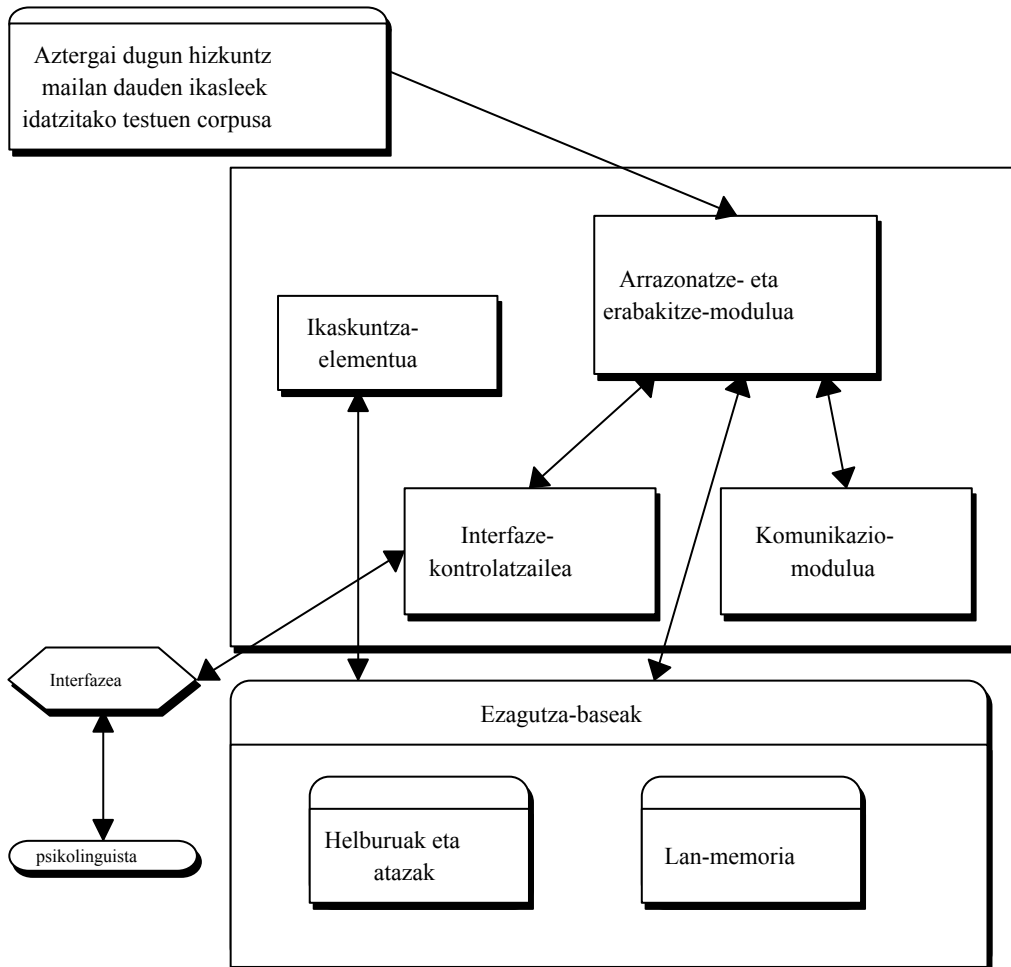
Arrazonatze- eta erabakitze-moduluak psikolinguistarekin duen elkarrekintzaren baitan eta corpusaren bidez jasotako informazioaren baitan, egoera konketua kontuan harturik planteatuko du zein helburu hautatu eta nola bete. Hau dela eta ikuspegi logiko batetik begiraturaz hiru maila bereizten dira: estrategikoa (zer egin), taktikoa (nola egin), eta exekutiboa (noiz egin). Garrantzitsua da gero kontuan hartzea helburu guztiak ez direla izaera berekoak, izan ere adibide bat aipatuz, hainbatetan sistema psikolinguistarekin duen

harremanaz arduratzen bada ere nagusiki, inoiz IDAZKIDetik informazioa lortzen ere saiatzen da eta beste batzuetan ezagutza linguistikoen kontsulta egitea izaten du xede, horretarako MUGARRI ingurunean diren tresna linguistiko desberdinetara jotzen duelarik (ikus V.2 irudia). Helburu-mota desberdinek, nola ez, ataza-mota desberdinak eskatzen dituzte, eta hauek HITESek bere ezagutza-basean aurretik gordeak izaten ditu. Bere baitan daramatzen ezagutza-baseen artean beraz, ataza-egitura eta lan-memoria genituzke. Azken honek sistemaren barne-egoeraren berri ematen du, eta bertan, beste hainbat gauzaren artean, bere ekintzak ingurunearen gain zer nolako eragina duten jakiteko informazioa gordetzen du.

Interfazearen kontrolzailea arrazonatze- eta erabakitze-moduluarekin dago komunikaturik, eta kanpoan gertatzen ari denari buruzko informazioa ematen dio. HITESek ingurunea aldatzeko hauta ditzakeen ekintzek psikolinguistaren erabakietan bakarrik eragin lezakete. Ekintza hauek interfazearen kontrolzailearen bidez ere gauzatuko dira.

Komunikazio-modulua aurrerantzean garatuko dugun modulu bat da, HITES, IRAKAZI eta IDAZKIDEekin komunikatzea bilatzen baitu. § V.2.3-n hiru sistemen arteko komunikazioaren helburuak jadanik aipatuak ditugu.

Ikaskuntza-elementuak baditu desberdinak diren bi helburu: ezagutza kolektiboari zein banako ezagutzari dagokien tartehizkuntz egiturak eguneratzea; eta konputazionalak diren tresna linguistikoek erabilitako ezagutza linguistikoa eguneratzeko erregelak sortzea. Modulu honen argibide zehatzagoa § V.5-n ikusgarri izango dugu.



V.3 irudia. HITESen arkitektura.

Beranduago, hain zuzen § V.6-n, gure arkitektura COSY (Cooperating Systems) (Haddadi, 96) agente-arkitekturarekin konparatuko dugu, izan ere gure sistemari moldagarri zaion arkitekturaz osatua baita azken hau, eta beraz, bere egitura HITESen arkitekturari ezin hobeki egokitzen zaio. COSY asmo orokorra duen arkitektura bat da, eta batez ere kontrol-estrategiak eta agente anitzekoak diren sistemen komunikazioa izan dira gehien garatu dituenak.

V.4.2 Ezagutza-baseak.

V.4.2.1 Lan-memoria.

Lan-memoriak sistemaren barne-prozesua jasotzen du, bere baitan (agendan) zera gordetzen du: hiru lan-mota, hau da, lan aktiboak, une bakoitzean zai dauden lanak eta unean martxan direnak; lantzen ari den testua bera (testua) eta saioaren egoera, hau da, ikaskuntza-maila jakin baterako edota ikasle jakin batentzat (saioa) tartehizkuntz ereduaren eraketa-

prozesuaren aldi bateko egoera. V.4 irudian aipatutako osagai guztien definizioa aurkitzen dugu.

<p>structure Lan_memoria; sub-type-of: Datu_modulua; parts: <i>saioa</i>: instance(Saioa); <i>agenda</i>: instance(Agenda); <i>testua</i>: instance(Testu_eredua);</p>	<p>structure Agenda; sub-type-of: Lan_memoria_osagaia; parts: <i>uneko_lana</i>: instance(Lana); <i>lan_aktibatuak</i>: set(instance(Lana)); <i>lanak_zai</i>: set(instance(Lana));</p>
<p>structure Saioa; sub-type-of: Lan_memoria_osagaia; parts: <i>ezagutza_iraunkorra</i>: instance(Ezagutza_iraunkorra); <i>ezagutza_dinamikoa</i>: instance(Ezagutza_dinamikoa); <i>ikasleak</i>: set(instance-name(Ikasle)); properties: <i>ikaskuntza_maila</i>: integer;</p>	<p>structure Lana; sub-type-of: Saioa_osagaia; parts: <i>helburua</i>: instance-name(Helburua); <i>atazak</i>: set(instance-name(Ataza));</p>
<p>structure Ezagutza_dinamikoa; sub-type-of: Saioa_osagaia; parts: <i>erregela_multzoak</i>: set(instance(Erregela_multzoa)); <i>fen_zer_uz_sailk</i>: instance(Fenomeno_multzoa); <i>fen_zer_jabetze_mailaz</i>: set(instance(Fenomeno_multzoa)); <i>fen_zer_ik_eredu</i>: instance-name(Fenomeno_multzoa); <i>fen_zer_hizk_eredu</i>: instance-name(Fenomeno_multzoa);</p>	
<p>structure Ezagutza_iraunkorra; sub-type-of: Saioa_osagaia; parts: <i>fenomenoen_aldakortasuna</i>: set(instance(Aldakortasuna)); <i>fen_zerrendak</i>: set(instance(Fenomeno_multzoa));</p>	

V.4 irudia. Lan-memoriaren errepresentazioa.

Saioa bi zatiz osaturiko objektu gisa adierazten da: ezagutza iraunkorra (*ezagutza_iraunkorra*) eta ezagutza dinamikoa (*ezagutza_dinamikoa*). Ezagutza dinamikoan saioan zehar detektatzen diren tartehizkuntz egiturak genituzke; hauetarik batzuk ikaslearen edota ikaskuntza-mailaren tartehizkuntz ereduen barruan kokatuko dira, eta beste batzuk ezagutza iraunkorrerako urratsa emango dute eta hala ondorengo saioetan kontuan hartuak izango dira. Azkenik, aipa dezagun *saioa* izeneko objektuen *ikasleak* multzo-atributuaren kardinalitatea 1 denean, ikasle jakin baten tartehizkuntz eredia eguneratuko genukeela; beste zernahi kasutan, ikaskuntza-maila jakin baten tartehizkuntza izango da eguneratzen ari garen hura.

Agendak lanak pilatzen ditu, bai uneko lanak, bai aktibatuak daudenak eta baita zai daudenak ere.

Azkenik, gainera dezagun lan bat definitzen dela ondoko bi osagaien bidez: helburua eta helburu hori lortzeko beharrezkoak diren atazak. Lanen garapenaz § V.4.3-n jardungo dugu.

Tartehizkuntz eredia eguneratzeko irizpideak

HITES sistemaren exekuzioaren lehen aldian, ikasleek idatzitako testu corpora abiapuntutzat hartuz, fenomeno linguistikoak erabat automatikoki sortzen dira. Fenomeno hauek jakina, hizkuntza-erregela multzo batek eta multzo hori detektatua izan deneko testuinguruak definitzen dituzte. Bigarren aldian, psikolinguistak fenomeno hauei buruzko informazioa osatu ondoren sailkapen bat egiten laguntzen du, bakoitza zenbateko maiztasunarekin agertu den kontuan hartuz. Egiten den sailkapenaren arabera sistemak lan-memoriaren ezagutza iraunkorrean edota dinamikoan kokatuko du. Tartehizkuntzaren eguneratzea honelaxe burutzen da.

Ezagutza iraunkorrean saioetan detektaturiko fenomeno linguistikoei buruzko informazioa pilatuko dugu, bere agerpen-maiztasuna dela eta, ondorengo saioetan tartehizkuntz ereduren batean sartzeko hautagai bezala izendatuak izango direnak.

- Fenomeno linguistikoez gain, fenomeno hauen aldakortasuna ere gordetzen dugu (ikus V.5 irudia) , hau da, fenomeno bakar bati dagozkion gertaera desberdinen multzoak (fenomeno-aldaera) jasotzen dira. Adibide bat jartzearren ergatiboak aditzarekin duen komuztaduran, esaterako, komuztadura fenomeno honen bi aldaera baditugu: lehenean ikasleak ergatibitatea -k batez markatzen du eta bigarrean ergatibitate hori markatu gabekoa da.

```

set Aldakortasuna;
  sub-type-of: Saioa_osagaia;
  element-type: instance-name(Fenomenoa);
  cardinality: min 0 max infinite;
  properties:
    fenomeno_identifikadorea: string;

set Fenomeno_multzoa;
  sub-type-of: Lista_cjto_fen;
  element-type: instance(Fenomenoa);
  cardinality: min 0 max infinite;
  properties:
    identifikadorea: string;
    jabetze_maila: string; value={barneratua, erdi_barneratua, gutxi_barneratua, ezezaguna}

```

V.5 irudia. Tartehizkuntzaren aldakortasuna.

Ezagutza dinamikoan sistemak detektatzen dituen fenomenoak arian-arian pilatzen ditu eta testu bakoitzeko corpusaren lanketaren ondoren detektatutako erregela-multzoak (*Erregela_multzoa*) abiapuntutzat hartuz, fenomeno linguistiko posibleak sortzen ditu lehendabizi.

- Lortutako fenomenoak psikolinguistak berresten ditu.
- Fenomenoak, beroien jabetze-mailaren arabera (*fen_zer_jabetze_mailaz*) sailkatu egiten dira (*barneratua*, *erdi_barneratua*, *gutxi_barneratua*), edota bestela sailkatu gabe geratzen dira (*fen_zer_ez_sailk*), hau da, jabetze-maila *ezezaguna* (ikus V.4 eta V.5 irudiak).
- Aldez aurretik ikaslearen tartehizkuntzan dauden fenomenoak honela izendatzen dira (*fen_zer_ik_eredu*), eta ikaslearen ereduan detektaturiko fenomenoaren egonkortasuna saioaren azkenean eguneratzeko helburua izan ohi du.
- Ikaskuntza-mailaren tartehizkuntz ereduan aldez aurretik bildutako fenomenoak detektatzen direnean ere, markatu egiten dira (*fen_zer_hizk_eredu*).
- Prozesuaren bukaera lantzen ari garen saio-motaren arabera da: ikasle jakin baten azterketa edota ikaskuntza-maila berean dagoen ikasle-talde baten azterketa egokitzen denaren baitan alegia. Honela bada, testu bakoitza eskuartean hartu eta prozesua behin eta berriro errepikatzen da, eta bukatutzat emango dugu aztergai dugun ikaslearen testuak oro edota aztergai dugun ikaskuntza-mailako ikasleen testuak oro prozesua burutua dutenean.
- Aztertzen ari garen ikasleak idatzitako testu bakoitzean fenomeno linguistikoak prozesuan zehar zenbat aldiz detektatu diren kontatu egiten da, V.6 irudian agertzen denaren arabera, (*detekzio_zerrenda*).
- Ikasle jakin baten fenomeno linguistikoak aztertzerakoan, fenomeno bakoitza hautatua ez izatetik hautatua izatera aldatuz joan liteke ikaslearen eredu propioan sartu ahal izateko, edo dagokion ikaskuntza-mailaren tartehizkuntz ereduan sartzeko edota jakiteko ondoko saioetan kontuan hartzekoa den fenomeno ote den: kasu honetan lan-memoriaren ezagutza iraunkorrean sartuko genuke.
- Fenomeno bakoitza zenbat ikaslerengan detektatu den jasotzen dugu (*ikasle-kopurua*), eta inoiz erabili izan duen ikasleari erabili izana behin kontatuko diogu.
- Prozesuan zehar fenomenoak jazotzen deneko egitura-, gai-, eta hizkuntza-baldintzak detektatzen dira.

```

structure Erregela_multzoa;
  sub-type-of: Saioa_osagaia;
  element-type: instance(Erregela);
  cardinality: min 0 max infinite;
  properties:
    fen_aldaera_ident: string;
    deskribapena: string;
    adibideak: string;

structure Fenomenoa;
  sub-type-of: Erregela_multzoa;
  element-type: instance(Erregela);
  cardinality: min 0 max infinite;
  properties:
    ikaskuntza_maila: integer;
    ikasle_kopurua: integer;
    detekzio_zerrenda: list(Detektatua);
    kandidato: string; value={ez_aktibatua, partikularra, komuna, ezagutza_iraunkorra}
    fenomeno_baldintzak: list(instance-name(Fenomeno_baldintza));

structure Detektatua;
  parts:
    testua: string;
    aldi_kop: integer;

```

V.6 irudia. *fenomeno* kontzeptuaren definizioa.

V.4.2.2 Helburuak eta atazak.

Atal honetan helburu eta atazen antolaketaz jardungo dugu, alabaina ezagutza honen prozesaketa ondorengo atal baterako utziko dugu. Tartehizkuntzaren definizioan gauzatzen diren atazak zeintzu diren deskribatuko dugu, eta horretarako zehazki gai hauek izango dira landuko ditugunak:

- Atazek betetzen dituzten helburu-motak.
- Atazen egitura.

Helburuak

HITES sistemak bigarren hizkuntza ikasten ari diren ikasleek eskuratutako tartehizkuntzaren garapen-prozesuaren azterketarekin zerikusia duen zenbait *helburu nagusi* burutzen du. Prozesu honetan zehar ikasle bakoitzari dagokion eta ikaskuntza-maila bera konpartitzen duen ikasle-multzoari dagozkion tartehizkuntz egiturak detektatzeaz arduratzen da (hnag4). Behin tartehizkuntz egiturak detektatutakoan, ikaslearen tartehizkuntzan (hnag1) eta/edo ikasle jakin horri dagokion ikaskuntza-mailaren tartehizkuntzan (hnag2) aldaketaren eragile izango dira tartehizkuntz egitura horiek. Gogora ekarri nahiko genuke ikaskuntza-mailaren tartehizkuntza ez dela maila horretan lortu beharko luketen ezagutza, berez duten ezagutza baizik. Hau dela eta, esango genuke HITES ezagutzaren eskuratze-sistema dela, eta zentzu honetan ikasten duen sistema bat da. Dena den, HITESen ikaskuntza ez da tartehizkuntz eredu jabetzera mugatzen. Izan ere, HITESen ikaskuntza-elementuak bai baitu bigarren funtzio bat ere, jabetze-prozesuan erabiltzen diren tresna linguistikoen ezagutza eguneratzea

alegia (hnag3). Tresna hauek aztergai dugun ikaskuntza-mailari egokitzen zaion ezagutza izan ohi dute.

Laburbilduz, sistemaren ikaskuntzaren helburu nagusiak honako hauek ditugu:

- hnag1. • Ikasle-ereduaren tartehizkuntza eguneratzea.
- hnag2. • Ikaskuntza-mailaren ereduaren tartehizkuntza eguneratzea.
- hnag3. • Ikaskuntza-maila bakoitzean diharduten tresna linguistikoen ezagutza eguneratzea.

Ikaskuntza-helburuez gainera zuzenean ez baina zeharka bederen ikaskuntzarekin zerikusia duten hiru helburu nagusi definitu ditugu. Hauexek dira beraz:

- hnag4. • Tartehizkuntz egitura berriak eraikitzea.
- hnag5. • Ikaskuntza-maila baterako ondorengo saioetarako lan-memoriaren modulu iraunkorra eguneratzea.
- hnag6. • Ikasle batentzat ondorengo saioetarako lan-memoriaren modulu iraunkorra eguneratzea.

Helburu nagusiak helburu zehatzak burutzearekin erdiesten dira; eta erabakitze- eta arrazonatze-moduluak edota § V.5-n azalduko dugun sistemaren ikaskuntza-moduluak finkatuko dituzte. Helburu nagusien artean saioaren zernahi unetan hauta daitezkeenak berezi behar ditugu batetik (hnag1, hnag2 eta hnag4) eta saioaren azkenean bakarrik aktibatzen direnak bestetik (hnag3, hnag5 eta hnag6).

Helburu bat honelaxe definituko genuke (V.7 irudia).

```
concept Helburu ;
  sub-type-of: Datu_modulua;
  properties:
    identifikadorea: string;
    mota: symbol; value={nagusia, espezifikoa}
    helburu_azalpena: string;
    modulua: symbol; value={planifikatzailea, gainbegiratzailea}
    bete: symbol; value={bai, ez, zai}
    atazak: list(string-value);
```

V.7 irudia. *helburu* kontzeptuaren definizioa.

Helburuak lortu ahal izateko sistemak ataza batzuk aurreikusten ditu. Helburu hauek ataza bakar baten bidez edo ataza askoren bidez lortzen dira. Jarraian helburu zehatz bakar bat bi ataza desberdinen bidez nola lortzen den ikusiko dugu. Atazaren aukeratzea edo alboratzea ataza bakoitzaren onarpen baldintzak betetzearen baitan dago. Azkenik, *helburu* kontzeptuaren *bete* izeneko atributua ez bezala markatuko dugu dagozkion atazak identifikatzen ez diren bitartean. Alabaina identifikatu orduko *zai* egoerara pasatuko dira

helburua bete artean, eta egoera honetan iraungo dute dagozkien atazen exekuzioa burutu bitartean.

Eman dezagun helburu nagusia hnag4 dugula (tartehizkuntzaren egitura berriak eraikitzea). Helburu hau lortzeko ekiten zaion prozesuan, fenomeno linguistikoa behin detektatuta, bere egonkortasun-balioa kontuan harturik jabetze-mailaren arabera sailkatu egiten da (helburu zehatza hesp1: *Fenomeno_hipotesiak_diskriminatzea*). V.8 irudiak erakusten digu barneratua, erdibarneratua edo gutxi barneratua fenomenoaren sailkapena bi atazen bidez gauza litekeela: *Fenomenoa_diskriminatzea_psikolinguistaren_laguntzaz* atazaren bidez edota *Fenomenoa_diskriminatzea_ereduen_laguntzaz* atazaren bidez. Helburu batak nahiz besteak, hnag1 eta hesp1-ek, *bete* eremuan "ez" balioa izango dute, helburuak lortuak izateko bete behar dituzten atazak betetzen ez dituzten bitartean .

<p>([hesp1] of Helburu (is-a Datu_modulua) (identifikadorea "Fenomenoa_diskriminatua") (mota "espezifikoa") (helburu_azalpena "Fenomenoaren jabetze-maila ikasle batentzat.") (modulua "planifikatzailea") (bete "ez") (atazak ("Fen_diskriminatzea_ereduen_laguntzaz" "Fen_diskriminatzea_psikolinguistaren_laguntzaz"))))</p>	<p>([hnag1] of Helburu (is-a Datu_modulua) (identifikadorea "Tartehizkuntz_egitura_sortua") (mota "nagusia") (helburu_azalpena "Fenomenoaren jabetze-maila ikasle batentzat, tartehizkuntz ereduaren gorderik dagoen ezagutza erabiliz")) (modulua "planifikatzailea") (bete "ez") (atazak ("Ikasle_baten_tartehizkuntz_egituren_sorkuntza"))))</p>
---	---

V.8 irudia. Helburuen adibideak.

Ondorenean ataza kontzeptua definituko dugu eta halaber aipatutako atazen esanahia.

Atazak

HITESen ataza saio bateko helburuak lortzeko sistemak burutu behar duen prozesua genuke. Atazen izaera arras bestelakoa da batetik bestera. Badira barnean jarduten duten atazak eta beste zenbaitek psikolinguistarekin edo sistemarekin komunikatzen den beste osagaien batekin komunikatzea izan ohi dute helburu.

Lehen multzoan ditugun atazak hauek dira:

- Tartehizkuntz egituren eraketa-prozesuan parte hartzen dutenak.
- Tresnen ezagutza linguistikoa eguneratzen dutenak.

- Atazen artean lehentasunak finkatzen dituztenak.

Bestalde, komunikazio-atazek:

- Interfazearen bidez psikolinguistarekin komunikazioa garatzen dute.
- Tresna linguistikoekin komunikatzen dira.
- Euskararen ikaskuntzan aholkularitza sistemarekin komunikatzen dira: IDAZKIDEREkin.
- Diagnostikoaren laguntza sistemarekin komunikatzen dira: IRAKAZIREkin.

V.9 irudian ataza baten errepresentazioa definitzen da. Ataza-mota (*ataza_mota*) atazaren egitura adierazten da eta egitura honek bere baitan honako hau ere badarama:

- Atazaren izena (*izena*).
- Atazaren helburua (*helburua*).
- Sarrera- eta irteera-parametroak eta hauen motak.
- Atazaren erabilerarako baldintzak (*onartze_baldintzak*) eta ataza hau erabili ondoreneko post-baldintzak (*post-baldintzak*).
- Guraso ataza (*guraso*).
- Azpiataza edo ume atazak (*umeak*).
- Kontrol-egitura. (*kontrol_egitura*).

```

structure Ataza;
  subtype-of: Saioa_osagaia;
  parts:
    helburua: instance-name(Helburu);
    sarrera_parametroak: set(instance(Parametroa));
    sar_parametroen_mota: set(instance(Mota));
    irteera_parametroak: set(instance(Parametroa));
    irt_parametroen_mota: set(instance(Mota));
    onartze_baldintzak: set(instance(Egoera));
    post_baldintzak: set(instance(Egoera));
    guraso: instance-name(Ataza);
    umeak: instance-name(Ataza_multzoa);
    kontrol_egitura: instance-name(Kontrol_egitura);
  properties:
    izena: string;
    ataza_mota: string;      value={ interfazea,          tartehizkuntza_eraiki,
ezag_ling_komunikazioa,      IDAZKIDE_komunikazioa,
IRAKAZI_komunikazioa,      atazen_lehentasuna, ezag_ling_eguneratzea }

set Ataza_multzoa;
  element-type: Ataza;
  properties:
    mota: string;
      value={konposatu_sekuentzialak, konposatu_paraleloak, disjuntiboak}

```

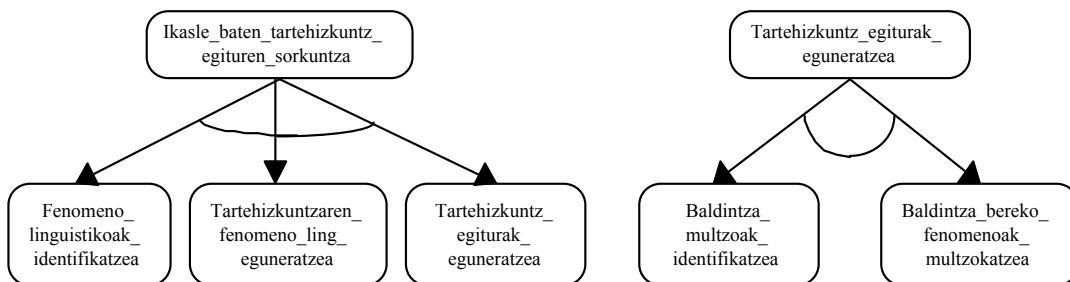
V.9 irudia. Ataza baten egitura.

Ataza baten kontrol-egituran burutzekotan diren azpiatazen artean honako bereizketa hau egiten dugu:

- Azpiatazaren baten gauzatzearekin guraso atazaren post-baldintzak lor ditzaketen atazak; *Fenomeno hipotesiak diskriminatzea* (ikus V.12 irudia) izeneko atazan esaterako hau gertatzen da. Hauei ataza disjuntiboak deitzen diegu.
- Post-baldintza hauek lortzeko azpiataza guztien gauzatzea ezinbesteko duten atazak. Adibide gisa *Fenomeno hipotesiak osatzea* (ikus V.11 irudia). Ataza-mota honi konposatuak deitzen diegu.

Ataza jakin baten ataza konposatuak gauzatzekoan, hurrenkerari dagokionean ez dute beti exekuzioan hurrenkera bera zergatik jarraitzen, horregatik tartehizkuntz egituren eraikitze-prozesuan ataza hauek bereziko genituzke:

- Konposatu sekuentzialak: exekuzioan segida bat jarraitzen duten azpiatazak ditugu. Adibidez, fenomeno linguistikoak identifikatu behar baditugu batera jazotzen diren erregela linguistikoak identifikatu beharko ditugu lehen-lehenik, era honetara geroago diskriminatuko ditugun fenomenoaren hipotesiak osatu ahal izango baititugu. (ikus V.11 irudia. *Fenomeno hipotesiak osatzea*).
- Konposatu paraleloak: azpiatazak hurrenkera jakinik gabe nolana ere egikari daitezke. Ikaslearen tartehizkuntzaren eguneratze-prozesuan esaterako, ikasle jakin baten fenomenoak identifikatzearen azpiataza, egin liteke beste ikasle baten fenomenoak tartehizkuntz egituren eguneratzean dihardutenean ere. Era berean, tartehizkuntz egituren eguneratzea (*Ikasle baten tartehizkuntz egituren sorkuntza*) burura daiteke (ikus V.10 irudia) adibidez fenomeno berrien identifikazioa (*Fenomeno linguistikoak identifikatzea*) mamitzen ari den bitartean. Hau bideragarri izatearen arrazoia litzateke *Ikasle baten tartehizkuntz egituren sorkuntza* atazak aldi berean ikasle jakin baten tartehizkuntz egituren testuinguru posibleak identifikatzea (*Baldintza multzoak identifikatzea*) azpiataza bezala duelako; eta azpiataza hau beste fenomenoaren identifikatzearekin batera paraleloki gauza litekeena da.



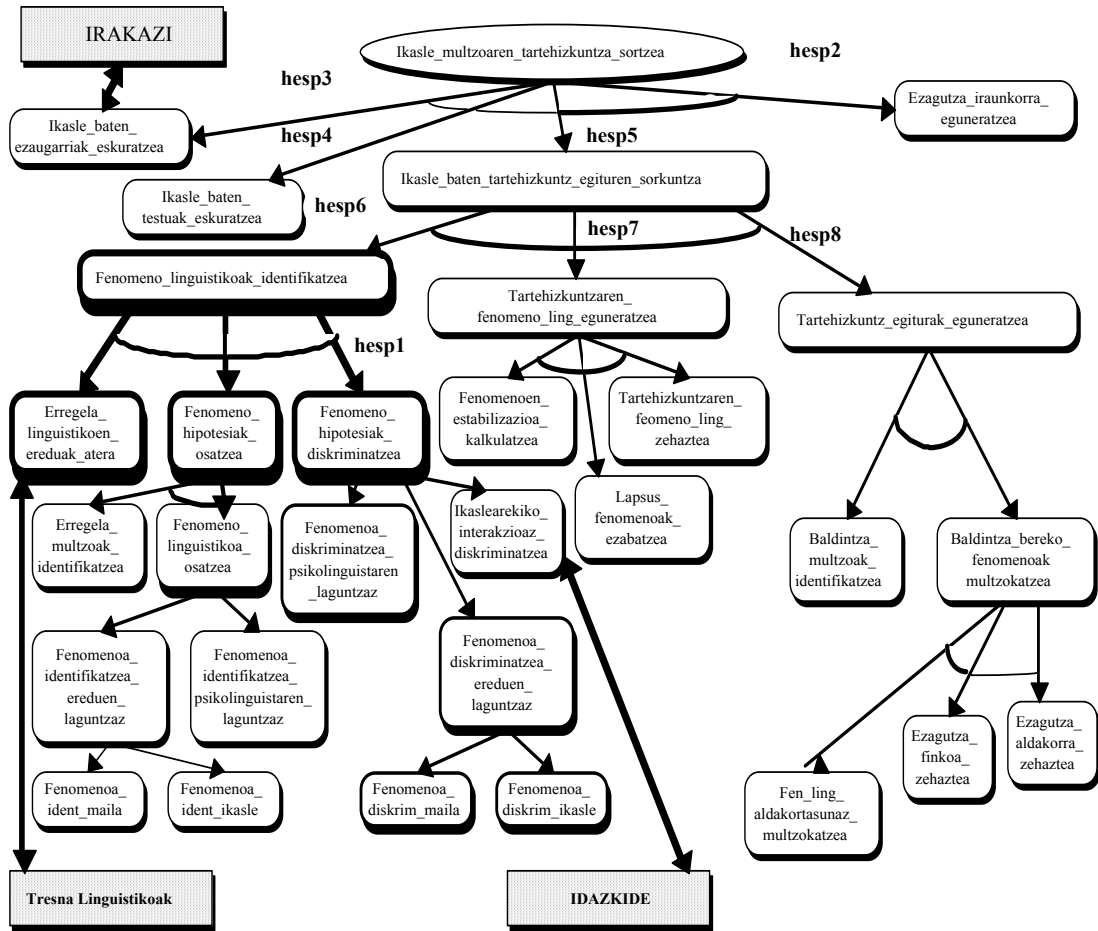
V.10 irudia. *Ikasle baten tartehizkuntz egituren sorkuntza*.

V.11 irudian ikasle-talde baten tartehizkuntz ereduaren sormenarekin zerikusia duen ataza-multzoei buruzko eskema bat azaltzen dugu. Eskema honetan exekuzio sekuentziala eta paraleloa duten ataza konposatuez gainera, ataza disjuntiboak ere ageri dira. Grafiko honetan bereziki *Fenomeno_hipotesiak_osatzea* eta *Fenomeno_hipotesiak_diskriminatzea* atazak azpimarratu ditugu jarraian atazaren adibide bat aditzera eman nahi baikenuen.

Adibidea

V.11 irudian argi asko dagerkigun bezala, tartehizkuntzaren eraketa-prozesuan tartehizkuntz egiturak eraiki aurretik fenomeno linguistikoak eta hauek gertatzen diren testuinguruak identifikatzeari ekin behar izaten zaio. Horretarako hurrenez hurren hiru ataza hauek burutzen dira:

- Corpusaren azterketan oinarritutako erregela linguistikoen identifikazioa.
- Psikolinguistaren laguntzarekin fenomeno linguistikoei buruzko hipotesien definizioa.
- Psikolinguistaren laguntzarekin hipotesien diskriminazio erdiautomatikoa.



V.11 irudia. Atazen antolaketa.¹

¹Irudian agertzen diren atazak honela interpretatzen dira:

(loturarik gabekoak)



ataza disjuntiboak

ataza paraleloak



ataza sekuentzialak

Arkuetan helburuen identifikazioa ere agertzen da. Dena den, kapituluaren zehar aipatutako helburuak besterik ez dugu adierazi irudian.

Fenomeno linguistikoei buruzko hipotesien definizio erdiautomatikoak gauzatzen den prozesuan *Fenomeno_hipotesiak_diskriminatzea* (ikus V.12 irudia) izena duen azpiataza bat badago, honetan fenomeno-multzo bat emanda fenomenoak jabetze-mailaren arabera sailkatu egiten ditu. Ataza honek azpiataza gisa hara zein dituen: *Fenomenoa_diskriminatzea_ereduen_laguntzaz*, *Fenomenoa_diskriminatzea_psikolinguistaren_laguntzaz*, *Ikaslearekiko_elkarrekintzaz_diskriminatzea*. Honek kasu honetan adieraziko liguke fenomenoaren sailkapena egiteko hiru modu badaudela:

- a) Aztergai dugun ikaslearen tartehizkuntz ereduaren arabera nahiz dagokion ikaskuntza-mailaren tartehizkuntz ereduaren arabera.
- b) Psikolinguistaren iritziaren baitan.
- c) Ikaslearekiko elkarrekintzatik jasotzen den informazioan oinarritua.

Azken kasu honetan HITES sistemak IDAZKIDE sistema lagungarriari deia egiten dio ikaslearekin fenomeno konkretuak lan ditzan. Aukera hau, etorkizunean sakonki aztertzeko ataza dugu, bada sistema desberdinen arteko protokolo baten garapena eskatzen baitu. Hortaz, ikaslearekiko elkarrekintzan oinarritutako diskriminatzea garatu gabe dagoenez, lehenengo biei buruz mintzatuko gara. Hala eta guztiz ere, psikolinguistarekiko elkarrekintzari dagokion zatiaz § V.4.4-n zabalago jardungo dugu.

Bi azpiatazen definizioan (ikus V.12 irudia) argi dakuskegun bezala, atazen erabilera baldintzak ez dira gero berdinak eta post-baldintzak ere batetik bestera desberdinak direla esan behar dugu. Aldez aurretik sortutako tartehizkuntz eruedetan oinarrituriko fenomenoaren diskriminazioa aktibatu ahal izateko, fenomenoak eredu horietan aldez aurretik aurkitu behar da. Psikolinguistak laguntzen duen diskriminazio kasuetan ordea, baldintza hau ez da beharrezkoa, hau da fenomenoak ezezaguna izan daiteke. Bestalde, post-baldintzak ere aldatu egiten dira, izan ere lehen kasuan fenomenoaren diskriminazioa beti burutzen bada ere, psikolinguistaren kasuan litekeena da diskriminatu gabe geratzea.

task-knowledge

```
task: Fenomeno_hipotesiak_diskriminatzea
task-specification
task-definition
goal: hesp2 - hipotesiak diren fenomeno linguistikoen zerrenda batetik hipotesi faltsuak
ezabatzea bai ikaslearen ereduaren arabera, bai ikasleari dagokion ikaskuntza-
mailaren eredu komunaren arabera. Honetaz gain psikolinguistaren laguntza izango
dugu.
input: fen_zer_jabetze_mailaz(LM-D), fen_zer_ez_sailk (LM-D), fen_zer_ik_eredu (LM-D),
fen_zer_hizk_eredu (LM-D),
ik_tartehizkuntza (TA), hm_tartehizkuntza (MAITA)
output: fen_zer_jabetze_mailaz(LM-D), fen_zer_ez_sailk(LM-D)
task-body
type:
parents: Fenomeno_linguistikoak_identifikatzea
sub-tasks: Fenomenoa_diskriminatzea_ereduen_laguntzaz
Fenomenoa_diskriminatzea_ereduen_laguntzaz

additional-roles: fenomenoak
acceptance-conditions:
(not hutsa(fen_zer_ik_eredu) or not hutsa(fen_zer_hizk_eredu) or not
hutsa(fen_zer_ez_sailk)) and (edozein fenomeno -> jabetze_maila = ezezaguna)
competence-conditions:
(edozein fenomeno and not dago (fenomeno, fen_zer_ez_sailk) -> jabetze_maila /=
ezezaguna)
control-structure:
begin
fenomenoak <- hutsa;
aztertu_beharreko_fenomenoak_hartu (in fen_zer_ik_eredu,
in fen_zer_hizk_eredu, in out fenomenoak);
fenomeno_ez_sailkatuak (in fen_zer_ez_sailk, in out fenomenoak);

loop

exit when bukaera (in fenomenoak)
hurrengoa (in fenomenoak, out fenomenoa);
sortu_helburua (in hesp1, in fenomenoa,
in ik_tartehizkuntza,
in hm_tartehizkuntza, in fenomenoa,
in out fen_zer_jabetze_mailaz,
in out fen_zer_ez_sailk);

end case;

end loop;

end
```

task-knowledge

```
task: Fenomenoa_diskriminatzea_ereduen_laguntzaz
task-specification
task-definition
goal: hesp1 - fenomenoak identifikatzean ezagunak diren ala ez zehaztea. Gainera,
fenomeno ezagunak kokatzen ditu: barneratuak, erdibarneratuak edo gutxi
barneratuak.
input: fen_zer_ik_eredu (LM-D), fen_zer_hizk_eredu (LM-D), ik_tartehizkuntza (TA),
fenomenoa
hm_tartehizkuntza (MAITA), fen_zer_jabetze_mailaz(LM-D), fen_zer_ez_sailk (LM-D)
output: fen_zer_jabetze_mailaz(LM-D), fen_zer_ez_sailk (LM-D)
task-body
type:
parents: Fenomeno_hipotesiak_diskriminatzea
sub-tasks: nil

additional-roles:
acceptance-conditions: (dago(fenomenoa, ik_tartehizkuntza) or
dago(fenomenoa, hm_tartehizkuntza))
and not dago(fenomenoa, fen_zer_jabetze_mailaz)
competence-conditions: (dago(fenomenoa, ik_tartehizkuntza) or
dago(fenomenoa, hm_tartehizkuntza))
and dago(fenomenoa, fen_zer_jabetze_mailaz)
control-structure:
begi
egonkortasuna_jaso (in ik_tartehizkuntza, in hm_tartehizkuntza, in fenomeno, out
egonkor);
case
0.0 <= egonkor <= 0.29:
jabetze_maila <- gutxi_barneratua;
0.29 <= egonkor <= 0.59:
jabetze_maila <- erdi_barneratua;
0.59 <= egonkor <= 1.0:
jabetze_maila <- barneratua;
end case;
gehitu-dagokion-fenomeno-zerrendan(in out fen_zer_jabetze_mailaz, in fenomeno,
in jabetze_maila);
end
```

```

task-knowledge
task: Fenomenoa_diskriminatzea_psikolinguistaren_laguntzaz
task-specification
  task-definition
    goal: hespl - fenomenoak identifikatzean ezagunak diren ala ez zehaztea. Gainera, fenomeno
        ezagunak kokatzen ditu: barneratuak, erdibarneratuak edo gutxi barneratuak.
    input: fen_zer_jabetze_mailaz (LM-D), fen_zer_ez_sailk (LM-D), fenomenoa
    output: fen_zer_jabetze_mailaz (LM-D), fen_zer_ez_sailk (LM-D)

task-body
  type:
  parents: Fenomeno_hipotesiak_diskriminatzea
  sub-tasks: nil

  additional-roles: sail, daki
  acceptance-conditions: dago(fenomenoa, fen_zer_ez_sailk )
                        and not dago(fenomenoa, fen_zer_jabetze_mailaz)
  competence-conditions:
    (not daki and dago(fenomenoa, fen_zer_ez_sailk)
     and not dago(fenomenoa, fen_zer_jabetze_mailaz))
    or
    (daki and not dago(fenomenoa, fen_zer_ez_sailk )
     and dago(fenomenoa, fen_zer_jabetze_mailaz))
  control-structure:
    begin
    PSIKOLINGUISTARI_SAILKAPENA_GALDETU
      (in fenomenoa, out, daki, out sail);
    if daki then
      gehitu-dagokion-fenomeno-zerrendan (in out
      fen_zer_jabetze_mailaz, in sail,
      in fenomenoa);
      ezabatu-ez-sailkatuen-zerrendatik (in out fen_zer_ez_sailk, in fenomenoa);
    end if;
  end loop;
end

```

V.12 irudia. Fenomeno_hipotesiak_diskriminatzea.

Aurreko atazotan deskribitu dugun prozesuan bi puntu azpimarratzea beharrezkoa da:

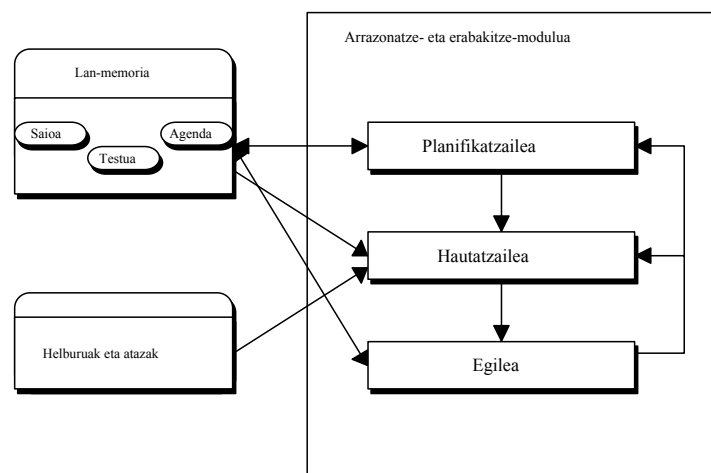
1. Fenomeno bat jabetze-mailaren arabera sailkatzea egonkortasunaren balioaren baitan dago. Balio hau kalkulatzen da ondoko ezaugarrien arabera: testu jakin batean sailkatzekotan dugun fenomenoa agertu den aldi-kopurua, fenomeno honen aldaerak testuan zenbatetan egokitzen den, eta neurri-unitatearen kopurua (hitza, esaldia, etab.). Hauetaz gain, ikaslearen berezko ezaugarriak ere kontuan hartzen dira egonkortasun-balioan. HITESen ikaskuntza-maila eredian zenbait tartehizkuntz egitura ikasle-motaren arabera antolaturik dago, ikaslearen ikaskuntza-ezaugarriak kontuan izanik. Honela, *Fenomenoa_diskriminatzea_ereduen_laguntzaz* atazan *egonkortasun-balioa lortzen dugunean, ikaslearen berezko ezaugarriak jadanik balio honetan kokatzen dira*. Fenomenoa sailkatzerakoan, beraz, aski zaigu tarte numerikoak bakarrik kontuan hartzea.

2. Helburu zehatz bat (hespl) erdiesteko (ikus V.8 irudia) *Fenomenoa_diskriminatzea_psikolinguistaren_laguntzaz* ataza hautatuko bagenu jadanik eredian zeuden fenomenoaren kasuetarako, eredian aldeztu aurretik gordea zegoen egonkortasun-balioaren aurrean psikolinguistak emandako oraingo balioak hartzen du lehentasuna. Gisa honetan sistemaren premisa nagusia are gehiago sendotzen da, *psikolinguista sekula nahasten ez denaren premisa* alegia. Dena den, psikolinguistak bere balioa eman aurretik beti izango du aldeztu aurretik gordea dagoen balioa kontsultatzeko aukera.

V.4.3 Arrazonatze- eta erabakitze-modulua.

HITESen arrazonatze- eta erabakitze-moduluak hasierako helburu nagusi multzo bat abiapuntu bezala hartuz lan egiten du eta helburu bakoitza lortzeko burutu beharreko atazak zeintzu diren identifikatzen ditu. Atazak gauzatzean arrazonatze- eta erabakitze-prozesuaren gidari diren helburuak lortzea bilatzen da. Haatik, prozesuaren gidari bakarrak ez dira helburuak. Izan ere, ataza bakoitzak baditu bere baitan bere gauzatua izan dadin beharrezkotzat jotzen diren onarpen-baldintzak. Beraz, ataza-hautatzaileak nahiz eta lantzeko helburu bat hautatua egoki, unean-uneko egoeran onarpen-baldintzak beteko dituen ataza bat izatekotan bakarrik jarraituko du aurrera.

HITESen arkitekturaz mintzatu ginenean arrazonatze- eta erabakitze-moduluan hiru maila bazirela aipatu genuen: estrategi, taktika- eta exekuzio-mailak hain zuzen. Horregatik, modulu honetan hiru osagai nagusi topatzen ditugu, hain juxtu maila bakoitzari horietako bana egokitzen baitzaie: Planifikatzailea, Hautatzailea eta Egilea (ikus V.13 irudia).



V.13 irudia. Arrazonatze- eta erabakitze-modulua.

• Planifikatzailea

Une oroz planifikatzailearen ardura zera da:

- Saioaren helburu nagusiak eguneratzea.
- Ataza-hautatzaileari helburuekin bat datozen atazak hauta ditzala eskatzea.

Saioaren helburu nagusiak eguneratzea

§ V.4.2.2-n helburu nagusi hauek definituak ditugu:

- hnag1. • Ikasle-ereduaren tartehizkuntza eguneratzea.
- hnag2. • Ikaskuntza-mailaren ereduaren tartehizkuntza eguneratzea.

- hnag3. • Ikaskuntza-maila bakoitzean diharduten tresna linguistikoen ezagutza eguneratzea.
- hnag4. • Tartehizkuntz egitura berriak eraikitzea.
- hnag5. • Ikaskuntza-maila baterako ondorengo saioetarako lan-memoriaren modulu iraunkorra eguneratzea.
- hnag6. • Ikasle batentzat ondorengo saioetarako lan-memoriaren modulu iraunkorra eguneratzea.

Planifikatzaileak helburu hauek honela hautatzen ditu:

Trataera-unitate bakoitzeko (testua) planifikatzaileak lan-memoriaren ezagutza dinamikoan tartekatuko diren tartehizkuntz egiturak sor daitezen agintzen du. Prozesuan sortzen diren egiturei egonkortasun-balioa eransten zaie eta planifikatzaileak balio hau esanahikorra dela irizten duenean, hnag1 edota hnag2 helburua lanean jarraraziko du, kontuan hartuz ikasle baten tartehizkuntza aztertzen ari ote garen, edo ikasle-talde batena ote den lantzen ari garena. hnag3, hnag5 eta hnag6 helburuak, ordea, egiturak sortzen diren bitartean ez dira hautatzen. Ikaste-moduluaren gainbegiratzaileak saioaren azkenean aktibatzen ditu eta helburu hauek.

Ataza-hautatzaileari atazak hauta ditzan eskatzea

Planifikatzaileak helburu bat zehaztu eta lan bat buru dadin deia luzatzen duenean, ataza-hautatzaileari igortzen dio. Lana lortu behar den helburuak eta helburu hori lortzeko hautatua izan den ataza-multzoak definitzen du. Lan batek lehentasuna izan lezake ez-primitiboa den ataza baten azpiatazaren exekuzioa adierazten duenean.

Unean-uneko edota zai dauden saioko helburu eta atazei buruzko informazioa lan-memoriaren agendan izan ohi da. Agenda hau (ikus V.14 irudia) hara nola osaturik dagoen:

- Unean-uneko lan batez.
- Burutzekotan den lan aktibatuen multzo batez.
- Zai dagoen lan-multzo batez: inoiz egileari bidali izandako lanak direnak baina honek uneren batean gelditu izan dituenak.


```

structure Agenda;
  sub-type-of: Lan_memoria_osagaia;
  parts:
    uneko_lana: instance-name(Lana);
    lan_aktibatuak: set(instance(Lana));
    lanak_zai: set(instance(Lana));
structure Lana;
  sub-type-of: Saioa_osagaia;
  parts:
    helburua: instance-name(Helburu);
    atazak: set(instance-name(Ataza));
    lehentasuna: boolean;

```

V.14 irudia. Agendaren egitura.

• Ataza-hautatzailea

Planifikatzaileak agindutakoa

Planifikatzaileak helburu bat hautatzaileari bidaltzen dionean, hautatzailea bera arduratzen da helburu honi dagozkion atazak zeintzu diren hautatzen, hau da, helburu eta atazen egitura jasotzen duen ezagutza-basea kontsultatu egiten du. Helburua lortzeko dauden atazen artean hautatzaileak ataza bat edo zenbait hautatzen du kontuan hartzen duelarik lan-memoriaren egoera zein den. Landu behar dugun helburuari dagozkion atazen onarpen-baldintzak betetzen diren ala ez aztertzen du beraz, eta unean-uneke egoeran erreparatu eta, instantzia litezkeen haiek zeintzu izan litezkeen identifikatu ere egiten du. Azken finean, hautatzaileak aktibatu bezala markaraziko duen lan bat sortzen du. Une honetatik aurrera planifikatzaileak bidalitako helburua *zai* egongo da. Honek adierazten du helburuaren lorpena hautatzailearen esku dagoela une horretatik aurrera. Instantzia litekeen atazarik ez balego helburua lantzeke utziko da eta hurrengo helburua garatzeari ekingo zaio.

Atazak konposatuak direnean hautatzaileak konposatutako ataza-multzoaz osatutako lan bakar bat sortzen du. Alabaina, ataza hautatzaileak disjuntiboak diren atazekin topo egiten duenean, haietako bakoitzeko eraikitzen du lana. Honela bada, ataza disjuntiboen egikaritzapenari eskerrak helburu bat betetzen denean, hautatzaileak helburu horri dagozkion gainontzeko lanak ezabatu egiten ditu.

Egileak agindutakoa

Egileak planifikatzaileari eta hautatzaileari lana bukatua dela jakinarazten dionean, hautatzaileak aktibaturik dauden lanen multzoa (*lan_aktibatuak*) nahiz geldiarazitakoak oro (*lanak_zai*) gainbegiratzeko dituzte. Aktibaturik den lan bakoitzean alabaina, lan-memoriaren egoera berrian aktibagarri segitzeko baldintzak betetzen direnez aztertuko du. Halakorik gerta ez baledi aktibatua zen lana geldiarazitakoen artera pasako da. Era berean,

geldiarazitako lan bakoitzean lan-memoriaren unean-uneko egoera kontuan hartuz, aktibagarri bihurtzeko baldintzak betetzen dituzenez aztertuko du.

Egileak halaber, jakinaren gainean jartzen du hautatzailea ez-primitiboa den ataza baten barruan helburu espezifiko berriren batekin topo egiten duenean. Hala, une honetan berean, lan-memoriaren egoeraren baitan helburu espezifiko horiek lor ditzaketean atazak (guraso atazaren azpiatazak) arakatuko ditu hautatzaileak eta ondorenean aktibaturiko lan-zerrendan edo geldiarazitako lanen artean kokatuko ditu. Kasu batean nahiz bestean hautatzaileak azpiataza hauei lehentasuna emango die ahalik eta azkarren exekutatuak gerta daitezzen. Gainera, guraso atazaren azpiatazak arakatzeko dagozkien lanak sortuko ditu. Honek helburu espezifikoaren identifikazioa ekarriko du.

• **Egilea**

Egileak ausaz erabakitzen du hautatzaileak aktibatutako lanen zerrendan kokatutako lanetatik zein lan gauzatu. Honez gainera lan bat burutua izan den bezain pronto hautatzaileari zein planifikatzaileari lanari zegokion helburua jadanik bete izan dela jakitera ematen dio.

Egilearen jokatzeko modua bi motatako atazen baitan da aldakorra: ataza primitiboak edota ez-primitiboak izatearen baitan alegia.

- Ataza primitiboak ataza sinpleagoetan deskonposa ezin daitezkeenak dira. Ataza burura dadinean helburua bete gertatuko da eta egileak hautatzaileari nahiz planifikatzaileari honen berri emango dio.

- Ataza ez-primitiboen kasuan, egileak bere gain hartuko du atazaren kontrol-egitura eta geldiarazia izan zen une hartan utzia izan zen puntu hartatik aurrera heldu eta gauzatzeari ekingo dio; edota lehen aldiz aktibatua suerta baledi, jakina, kasu honetan hasiera-hasieratik heldu eta mamitzeari ekingo dio. Ez-primitiboa den ataza bat garatzerakoan litekeena da egileak ez-primitiboak diren atazekin ere topo egitea. Kasu hauetan, oraingo ataza bere horretan uzten da eta egileak atazen hautatzaileari deia luzatzen dio atazen berregituratze bat egin dezan. Hautatzaileak azpiatazak zeintzu diren identifikatzen du eta lehentasuna dutenen artean izendatzen ditu, behin aktibatuak izan eta lehen-lehenik haiek exekutatuak izan daitezzen. Ataza ez-primitibo bati dagokion helburua ez da betetzat emango bere kontrol-egituraren gauzatzea buka dezan arte. Izan ere, kontrol-egitura erabat gauzatzen denean izango baita helburua bete, eta orduan honen berri hautatzaileari eta planifikatzaileari emango dio.

V.4.4 Interfazearen kontrolatzailea: psikolinguistarekiko elkarrekintza.

Tartehizkuntz egituren eraketa-prozesuan sistemak fenomeno linguistikoak eta fenomeno hauez osatutako tartehizkuntz egiturak pantailan agerian jartzen ditu. Prozesu hau garatzerakoan zenbait aldiz sistemak psikolinguistari bere iritzia galdetzen dio, baina beste zenbaitetan alde aurretik eraiki dituen eruedetan oinarritzen da zuzen-zuzenean. Litekeena da, ordea, azken kasu honetan psikolinguista sistemak hartutako erabakiarekin bat ez etortzea. Honelakoetan beraz, psikolinguistak desadostasunak nabarmen ditzake beti ere tartehizkuntz ereduaren hainbat egitura pertsonalki egunera ditzan. Eguneratze hauek saioa bukatzen denean egingo dira, saio baten barruan datuen trinkotasuna ez dadin hauts. Hasiera batean, beraz, saioan zehar psikolinguistak sistemak eskatutakoari erantzun besterik ez dio egingo. Sistemaren erabakia atazen onarpen-baldintzek erabakitzen dute. V.12 irudian esaterako, ikus dezakeguna da *Fenomenoa diskriminatzea psikolinguistaren laguntzaz* ataza, alde aurreko erueden bidez tratatu izan ez bada (*not dago(fenomenoa, fen_zer_jabetze_mailaz)*) bakarrik aktibatuko dela, eta tratatu bada aipatu ereduak alde aurretik emandako egonkortasun-balioak eta unean-unean saioan sistemak kalkulaturakoak elkarrengandik aski urrun suertatzen badira.

Interfazea hiru mailatan dago diseinaturik: hitz-mailan, esaldi-mailan eta hiztegi-mailan. Den maila delarik ere, psikolinguistak fenomeno linguistikoekin topo egiten duenean sistemak fenomenoak aurkezten dizkio bereiztuta goitik beherako metodologiaren aplikazioaren arabera:

1. Gorantzako ordenean segidan datorren ikaskuntza-mailan gertatzen diren fenomenoak eta aztergai ditugun ikasleengan detektatu direnak.
2. Gorantzako ordenean segidan datorren ikaskuntza-mailan gertatzen diren fenomenoak eta aztergai ditugun ikasleengan detektatu ez direnak.
3. Gorantzako ordenean segidan datorren ikaskuntza-mailan gertatzen ez diren fenomenoak eta aztergai ditugun ikasleengan detektatu direnak.

Hasiera batean HITES diseinatu zenean testu idazketarako irakasleak bidali ohi zituen testu-mota eta ariketetan gertatutako fenomenoak hartu ziren oinarri gisa. Dena den burututako frogek (VII. kapituluaren azalduko ditugunak) ongi asko erakusten digute ariketen arabera egituraturako interfaze batek ez ziola deus eskaintzen ahal psikolinguistari. Horregatik hain zuzen, interfazea diseinatzeko orduan ariketa-mota ez dugu irizpidetzat hartu. Ikaskuntza-maila bakoitzari dagokion hiztegiak eta erregela linguistikoek berriz, eragina adierazi dute interfazearen diseinuari dagokionez. Honela izanik, analisi bat egokia ez denean sistemak psikolinguista behartu egiten du analisi honetan parte hartzen duten eta aztergai dugun ikaskuntza-mailan ezezagun diren lekuak zein erregelak fin ditzan. Behin ezagutza fintzen denean, sistemak datuak berraztertzen eta, prozesuari aurrera ekiten dio.

Psikolinguistak fenomeno linguistikoei buruzko informazioaren aurrean hiru aukera desberdin ditu:

1. Dagokion fenomeno ezabatzea.
2. Dagokion fenomeno eguneratzea.
3. Fenomeno berri bat eranstea.

Erantzuna eman ahal izateko, ordea, kontsulta-mota desberdinak egiterik izango du, besteak beste:

- a) Ikaslearen tartehizkuntz ereduari buruzko eta ikaskuntza-mailari buruzko kontsulta.
- b) Saioan detektatu arren eruedetan gehitu ez den ezagutzaren egoerari buruzko kontsulta.
- c) Ikaslearen ezaugarriari buruzko kontsulta
- d) Kanpoko tresna linguistikoen bidez ezagutza linguistikoaz HITESi egindako galderak.

a, b eta c motatako kontsultek zenbait objekturen ezaugarriek dituzten balioez baizik ez dezakete galde, esaterako tartehizkuntz egitura baten egonkortasun balioei buruz, beratu osatzen duten fenomeno linguistikoei buruz, etab. Gisa honetako kontsultak psikolinguistaren interfazearen barruan kokatu nahi den GBC interfazearen bidez (Elorriaga et al., 95) burutzen dira.

d motatako kontsulteren artean IV. kapituluan garatu ditugun tresna linguistiko batzuetarako deia kokatu beharko genuke eta VI. kapituluan IDAZKIDERI buruz mintzatzerakoan aditzera emango ditugun tresna linguistikoetarik zenbait ere kontsulta-mota honetan kokatzekoak dira. Psikolinguistak zehazki analizatzaile morfologikora sarbidea izateaz gain, desanbiguatzaile morfologikora, EDBL lexiko datu-basearen kontsulta-interfazera, *estatistikak* programara (ikus IV. kapitulua) eta *ereduak_atera* programara ere (ikus IV. kapitulua) sarbidea izango du.

V.5 Ikaskuntza HITESen.

V.5.1 Ikaskuntza-elementua HITESen.

HITESek arras desberdinak diren bi helburu betetzeko erabiltzen ditu baliabideak: burutzekotan den atazan eragin dezan eta geroan bere jarduera-ahalmena hobetu dezan. Geroan jarduera ahalmena hobetze hau alabaina, desberdinak diren hiru helburuk baldintzatzen dute:

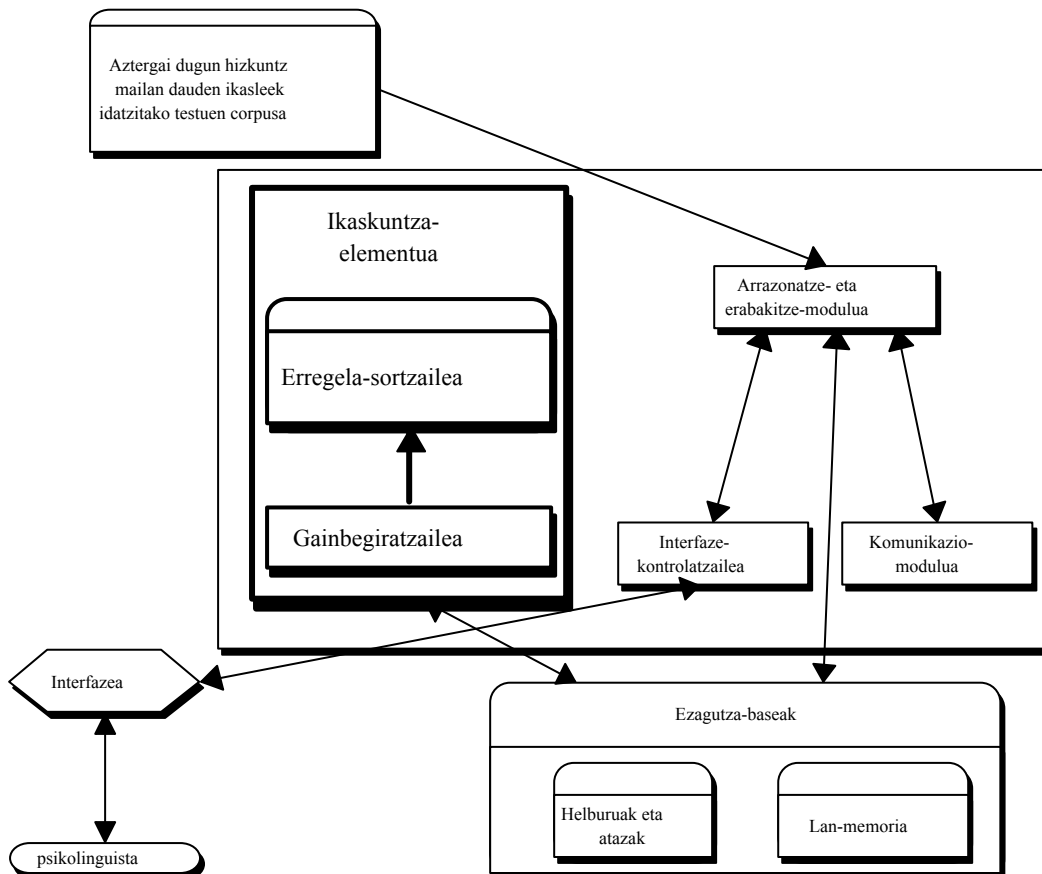
- Ikaskuntza-maila berean dagoen ikasle-talde bati dagozkion tartehizkuntz egituren finketak (ez dago alde aurretiko ezagutza abiapuntu gisa).
- Ikasle bati dagokion tartehizkuntz egituren finketak (ez dago alde aurretiko ezagutza abiapuntu gisa).

- Ikaskuntza-maila bati dagozkion tresna linguistikoetan erabilitako ezagutza linguistikoaren finketak (aldez aurretiko ezagutza dago abiapuntu gisa).

Bi ezagutza-baseen eguneratzea: tartehizkuntz ereduak eta ezagutza linguistikoa HITESen ikaskuntza-elementutzat jo behar dira. Ikaste-prozesu honetan HITESek eredu gainbegiratua jarraitzen du. HITESen ikaskuntza gainbegiratua dela esaten dugu kanpoko agente batek gainbegiratzen dituelako ikaskuntza-elementuaren aldaketen erabakiak, gure kasuan jakina, kanpoko agente hau psikolinguista genuke.

HITESen ikaskuntza-elementuaren osagai kontzeptualak (V.15 irudia) honako hauek genituzke:

- Barnekoa den gainbegiratzailea.
- Kanpoko den gainbegiratzailea (psikolinguista).
- Ezagutza eguneratzeko erregelen sortzailea.



V.15 irudia. Ikaskuntza-elementua.

Gainbegiratzaileak lan-memoriaren egoera zein-nolakoa den behatzen du eta saioan zehar detektatzen diren tartehizkuntz egituren eta/edo fenomeno linguistikoaren baitan ezagutza berri bat sor dezakeen erabakitzen du. Egia esan, barruko gainbegiratzailearen

funtzioa arrazonatze- eta erabakitze-moduluaren planifikatzaileak burutzen du, hauxe baita helburu berriak definitzeaz arduratzen dena (kasu honetan helburua ikaslearen eta/edo ezagutza kolektiboan tartehizkuntz egitura berriak tartekatzea da).

Erregela-sortzaileak barruko gainbegiratzaileak duen informazioa kontuan hartzen du eta kanpoko tresnek duten ezagutza linguistikoaren eguneratze-erregelak sortzen ditu. Tartehizkuntz egituren sormenean ez bezala, ezagutza linguistikoaren eguneratze-erregelen sorkuntza, saio bakoitzaren azkenean gauzatu ohi da. Halere, sortuak izanagatik kanpoko gainbegiratzaileak, hau da, psikolinguistak berretsi artean ez dira erregela hauek aplikatuko.

V.5.2 Ikaskuntza-ataza: adibidea.

§ V.4-n fenomeno linguistikoaren hipotesien diskriminazioa landu dugu eta ikusi dugu nola sistemak inoiz edo behin zenbait fenomeno ezezagun bezala sailkatzen duen (*jabetze_maila: ezezaguna*). Era honetako fenomenoaren detektatzea IV. kapituluan azaldu dugun goitik beherako metodologiaren ondorioa da. Metodologia honen arabera, ikaskuntza-maila bateko tartehizkuntz ereduaren eraketak hasierako ezagutza gisa gaineratik segidan datorren mailak duen tartehizkuntz eredu erabili ohi du. Ondorioz, beheragoko mailako ikasleen testuak fenomenoak detektatzeko abiapuntutzat hartzen direnean, maila horri ez dagozkion interpretazioak egon egongo dira. Hain zuzen, interpretazio hauek aztergai dugun beheko maila honi ez dagozkio, gaineratik dagoen mailan sortutako ezagutza zati haren berri ematen baitigute.

Ondorenean HITESen ikaskuntza-elementuaren eguneratze-erregelen sortzailea denaren funtzio bat CLIPS programazio-lengoaian definiturik azalduko dugu.

Fenomeno-ezezagunak-aztertzea funtzioak (ikus V.16 irudia) aztergai dugun ikaskuntza-mailan ezezagun diren fenomenoak hautatzen ditu, hain zuzen lan-memoriaren modulu dinamikoan pilaturik daudenak. Fenomeno hauen artean ikaskuntza-mailako tartehizkuntz eredu sartzeko hautagai ez direnak bakarrik ditu hautatzen. Eta azkenean tresna linguistikoek erabiltzen duten ezagutza linguistikoaren eguneratzeko erregela sortzen du automatikoki ezagutza-ezabatze-agindua-gehitu izeneko atazaren bidez (ikus V.16 irudia), ondorenean psikolinguistak berretsiko badu ere. Eguneratze-erregela hau erabiliz HITESek gaineratik segidan datorren ikaskuntza-mailaren eredu oinarritzat hartuz uneko mailaren eredu finduko du. Honi buruz azalpen gehiago emango ditugu § V.7.2-n V.27 irudian adieraziko dugunez.

```
(deffunction Fenomeno-ezezagunak-aztertzea (?eguneratze_erregelak ?lan_memoria)
  (open ?eguneratze_erregelak fitxategia "a")
  (bind ?fen_ezez_zerr (send ?lan_memoria fen-ezez-jaso))
  (bind ?fenomenoak (send ?fen_ezez_zerr zerrenda-instantziak-jaso))
  (while (not (hutsa ?fenomenoak))
    (bind ?kandidato (send (lehen-osagaia ?fenomenoak) kandidato-jaso))
    (if (eq ?kandidato komuna) then
      (ezagutza-ezabatzeko-agindua-gehitu
        (zerrenda-lehen-osagaia ?fenomenoak) fitxategia ?lan_memoria)
      )
    (bind ?fenomenoak (rest$ ?fenomenoak))
  )
  (send ?lan_memoria fen-ezez-ezabatu)
  (close fitxategia))
```

V.16 irudia. Fenomeno-ezezagunak-aztertzea.

V.6 HITES: ikaskuntzarako ahalmena duen agente bat.

V.6.1 HITES agenteen ikuspegitik.

Agenteak ingurune jakin batean pertzepzio-sekuentzien arabera ekintzen bidez erantzuna ematen duten entitateak direla esango genuke. Agenteak autonomoak dira, ikasteko ahalmena ere badute eta bere jokabidea norbere esperientzia propioaz gain, diseinatzaileak ingurunearen gainean eransten duen ezagutzaren arabera ere bada. Agente ideiaren inguruan sistema asko garatzen ari dira ikerkuntzaren arlo desberdinetan, besteak beste, agente pedagogikoak (Frasson & Gouardères, 98) sortzen ari dira ikaskuntza eta irakaskuntza eremuetan.

Müller-ek (1998) gaur arte agente adimendunen arloan garatu diren aplikazioak eta arkitekturak aztertu, eta azterketa horretatik abiatuta agenteen taxonomia proposatzen du bi ezaugarri desberdinetan oinarrituta:

- Agentearen egoera materiala.
 - HW agenteak: ingurunearekin elkarreragite fisikoa dute.
 - SW agenteak: software ingurune erreal edo birtualekin elkarreragiten duten programak.
- Agentea eta ingurunearen arteko elkarrekintza-mota.
 - Agente autonomoak: bi entitateren arteko erlazioaz arduratzen dira, hau da, agentea eta ingurunea.
 - Agenteanitzak: ingurunean beste sistema batzuk agertzen dira, horiek izango dira agenteak edo ez-agenteak.
 - Agente lagungarriak: agente-mota konkretu batekin elkarreragiten dute, adibidez giza agenteekin.

Definitutako taxonomiatik abiatuta ondoko agente-motak sortzen dira: HW agente autonomoak, SW agente autonomoak, HW agenteanitzak, SW agenteanitzak, HW agente lagungarriak eta SW agente lagungarriak.

Gauden kapituluan garatutako HITES sistema SW agente autonomoen artean kokatzen da, bere jarduera autonomoa delako eta ingurune batean erabakiak hartzea bere helburua delako. Dena den, ezin dugu ahaztu HITES sistemaz aparte badirela beste bi sistema ingurune berean aritzen direnak: IDAZKIDE eta IRAKAZI. IDAZKIDE, (Johnson & Rickel, 97)-n azaltzen diren agente hezitzaileen moduan, SW agente lagungarria da bere exekuzioa erabiltzailearen nahien arabera egokitzen delako. IRAKAZI ere SW agente lagungarrien artean kokatzen da: kasu honetan irakasleari laguntzen dio ikasleak egindako desbideratzeen diagnosi-prozesuan.

Edozein kasutan, tesi honetan proposatzen dugun MUGARRI ingurune informatikoaren azken helburua ondokoa dugu: agenteanitzeko ingurunea non aipatutako hiru sistemek elkarrekiko komunikazio zuzena izango luketen, portaera sozial-kooperatiboa oinarri izanik. Ezin dugu ahaztu HITESek ikasleen corpusetatik eskura dezakeen informazioa osatua ez denez, ikaslearengandik informazioa jaso beharra duela (IDAZKIDE), eta irakaslearengandik diagnostikoari buruzko iritzia ezagutu behar duela (IRAKAZI). Horregatik, tesi honetan proposatzen dugun ingurunearen azken helbura litzateke SW agenteanitz elkarreragileak garatzea.

Ildo honetatik COSY (Burmeister & Sundermeyer, 92) sistema agenteanitz elkarreragilea agertzen zaigu. Agente honentzat Haddadik (1996) arkitektura orokorra definitzen saiatzen da sistema agenteanitzen azterketarako metodologia definitzeko asmoz. Hau dela eta, COSYren arkitektura oinarritzat hartu dugu HITES sistema etorkizunean bihur daitekeen agenteanitz elkarreragilea aztertzeko.

V.6.2 COSY agenteanitzaren arkitektura.

Komunikazio-protokolo sail baten bidez agenteak inguruneke beste agente batzuekin elkarrekintzan izaten direnean agenteanitzak diren sistemak direla esaten dugu. Agenteanitzak diren sistema hauek autonomoak diren giza agente balira definitzen dira (Cuenca, 98). Hauek, alde batetik, heterogeneoak dira bakoitzak bere arkitektura propioa duelako, besteekiko desberdina gehienetan, eta, beste alde batetik, homogeenak dira helburu bakar bat lortze-bidean koalizioak osatzen dituztelako lankidetzan jarduten dutenean.

Aurreko atalean esan dugun bezala, COSY agenteanitza da, eta bere arkitektura baliagarria izan daiteke HITES sistemarentzat agenteanitza gisa planteatzen dugunerako.

Horregatik, ondoren, COSY arkitekturaren moduluak labur azaltzera ekingo diogu, eta hurrengo atalean HITES eta COSYren arteko konparaketari buruz arituko gara.

COSY arkitekturaren diseinuak goitik beherako metodologia jarraitzen du, eta hauek dira kanpokoan diren moduluak: aktoreak, sentsoreak, komunikazioa, motibazioak eta modulu kognitiboa. Hiru lehenak domeinuaren menpe daudenak dira. Sentsoreak ingurunearen egoerari buruzko informazioa eskaintzen dutenak dira; aktoreak agentearen ekintza fisikoak burutzen dituztenak dira; komunikazioa uneko domeinuaren komunikazio fisikoari dagokio; motibazioak agenteak epe luzera dituen helburuak dira; azkenik, modulu kognitiboa ingurunearen egoera ebaluatu ezezik agentearen ekintzak hautatu, exekutatu eta monitorizatu ere egiten du. Modulu hau ezagutzan oinarritutako sistema bat da eta honako osagaiez osaturik dago.

- Ezagutza-baseak.
- Scripts-en exekuzio-modulua.
- Protokoloen exekuzio-modulua.
- Arrazonatze- eta erabakitze-modulua.

Ezagutza-baseek agentearen arrazonatze-prozesurako beharrezkoak diren datu-egiturak dituzte (helburuak, planak etab.).

Scripts-en exekuzio-moduluak aktoreek zuzenean egin ditzaketen planak monitorizatu eta exekutatzeaz gain, agentearen barruko exekuzioa eskaintzen duten plangintzak ere egiten ditu. Bigarren kasu honetan beste scripts-etarako deiak izaten dira (protokoloetarako deiak besteak beste).

Protokoloen exekuzio-moduluak agenteek beste agenteekin dituen mezu-trukeak kontrolatzen ditu.

Arrazonatze- eta erabakitze-moduluar dagokio agentearen arrazonatzearen ardura: ingurunearen egoeraren arabera planak hautatzearena, eta agenteak bere nahiak ase ditzan helburuak lortzearen ardura ere. Bere funtzioen artean planen ataza administratiboak ere badaude, besteak beste plangintza baten exekuzioa etetea eta erabakiak monitorizatzea. Modulu hau da beraz scripts eta protokolo-exekuzio moduluen exekuzioa erabakitzen duena.

V.6.3 COSY: HITESi egoki dakioken arkitektura.

HITESen arkitektura COSY arkitektura berari aski ongi egoki dakiokiena da. Hau bideragarria zaigu COSYren arkitekturak mekanismo orokor bat eskaintzen digulako, hain zuzen bertan arazoaren ebazpenaren ezagutzak, erabilpen zehatzari espezifiko zaionak, planen

egituren ezagutza-basean gordea eta egituratua egon behar du. Eta gure kasuan ezagutza hau atazen egiturak gordetzen dituen ezagutza-basean kokaturik dago.

Bai COSYn eta baita HITESen ere badago arrazonatze- eta erabakitze-modulua, ezagutza-basea eta komunikazio-modulua, baina hainbat desberdintasun badituzte:

- Batetik, HITESen interfazearen kontrolatzailea deitzen duguna ez da COSYren kontrolatzailearen gisakoa, bere kanpoan baitago kokatua. COSYn kontrolatzaile hau sriptsen exekuzio-moduluan tartekaturik dago, eta berari dagokio interfazeari bere baitan dagozkion ekintzen betearaztearen kontrola. Exekuzio-modulua agentearen barneko ekintzak kontrolatzeaz arduratzen da.

- Bestalde, HITESen barruko ekintzen kontrolatzailea osoki arrazonatze- eta erabakitze-moduluaren barruan dago. Atazak indarrean jartzen dituen modulua da eta exekutatu ere egiten ditu. COSYn sheduler-ak agentearen barne-egoera kontuan hartuz aurrera eramatekoak diren atazak ilaran jarri eta bere horretan uzten dituen gisa berean, HITESen arrazonatze- eta erabakitze-moduluaren hautatzailea izan ohi da atazak bere horretan uzteaz eta aktibatzeaz arduratzen dena.

- Gainera, HITESek garatzen ez duen COSYren alderdietako bat zera gence: nahi eta gertaeretan oinarritzen den helburuen antolaketa. Agenteak gertaeren monitore bat sartzen du eta berak sistemaren kanpoko nahiz barruko gertaerak detektatzen ditu, hau da, schedulerrek bidalitakoak. Zernahi modutan alderdi hau HITESen integra liteke.

- Azkenik aipatu nahiko gence HITES eta inguruan diren agenteen arteko komunikazioa bi modu desberdinetan lortzen dela:

- Zeharka, ezagutza-baseen bidez.
- Zuzenean, sistemen arteko deiak eginez.

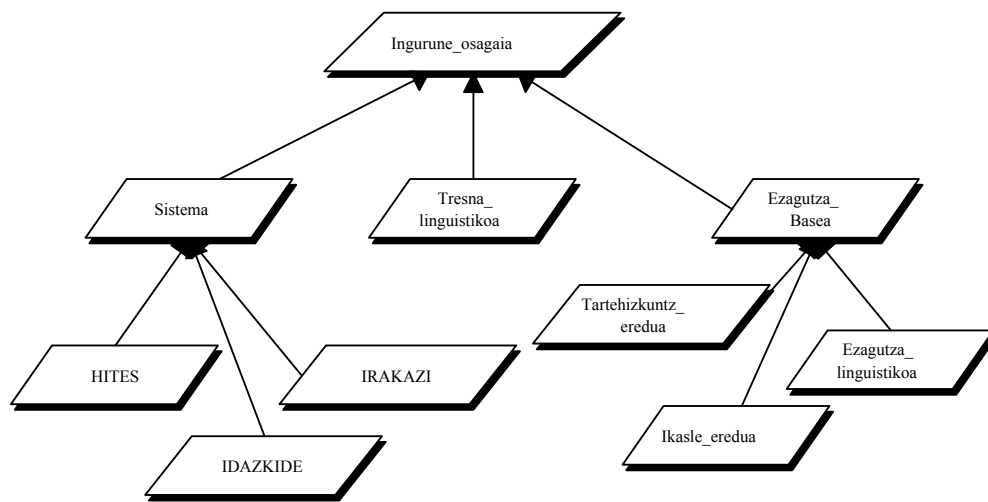
Azken hau garatu gabe dagoen zatia izan arren, tesi honetan aditzera emandako sistemetan ezin egokiago integra daiteke. Gauza jakina da agente batek kanpoarekiko duen komunikazioa baliabide eta ekintzetatik haruntzago doala. Adibidez, agente bat litekeena da beste agenteekin komunikatzea informazioa eskatu asmotan. Gure kasuari helduz, § V.2-n aipatu genuen HITES beste bi agenterekin komunikatzen zela: alde batetik IDAZKIDEREkin, agente honen helburua ikasleari bere tartehizkuntz egiturak hobetzen laguntzea da jatorrizko hiztun batek duen ezagutza linguistikoaren egiturara are gehiago hurbil dadin, eta bestetik IRAKAZIrekin komunikaturik dago, agente honek bigarren hizkuntza ikasten ari diren ikasleen ikaste-prozesuaren diagnostikoan laguntzen du. Komunikazio honetan mezuak hartzea eta igortzea komunikazio-moduluaren bitartez burutuko da. Une honetan moduluaren diseinua lantzen ari da eta luzaro gabe hiru agenteen arteko komunikazioa egiazkoa bilakatuko zaigu.

V.7 Objektuei zuzendutako programazioan oinarritutako implementazioa.

V.7.1 Ikaskuntza-ingurune objektuen hierarkia: garapena HITESen.

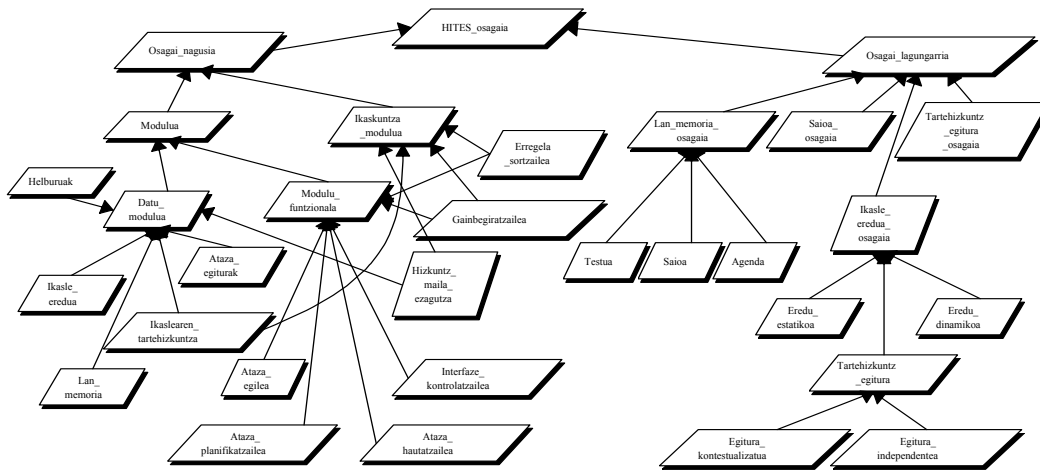
HITESen definitutako objektuen hierarkia laburki adieraztera mugatuko dugu atal honetako zeregina, izan ere kapituluaren zehar objektu desberdinak, bere ezaugarriak eta bere baitan biltzen dituzten funtzioak identifikatu baititugu.

V.17 irudian dagerkigunez, HITES beste sistemekin ezagutza-base multzo bat konpartitzen duen ikaskuntza-ingurunearen osagai bat da.



V.17 irudia. Ikaskuntza-ingurune objektuen hierarkia.

HITESen osagai nagusiak eta bigarren mailako osagaiak bereizten ditugu (ikus V.18 irudia). Nagusiak sistemaren arkitekturaren moduluak ditugu, osagai lagungarriak berriz, modulu hauen osagaiak genituzke.



V.18 irudia. HITES sistemaren objektu-hierarkia.

Osagai nagusiak

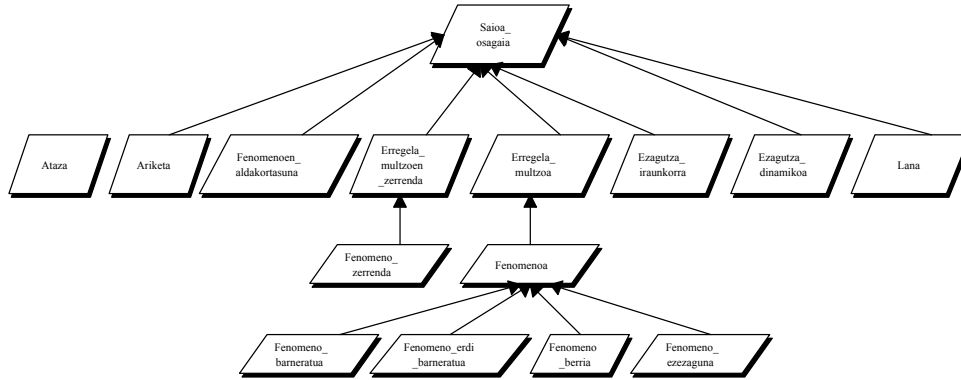
HITESen osagai nagusietan sistemaren ikaskuntza-moduluak gainontzeko moduluetatik berezi egiten ditugu:

Ikaskuntza-moduluak sistemaren ezagutzaren eguneratzean parte hartzen dutenak dira. Adibidez, atazen *gainbegiratzailak* maila bakar bateko ikasle-taldearentzat komunitat jo litekeen tarte hizkuntz egitura bat detektatzen bada, *Erregela_sortzailea* izeneko modulua jakinaren gainean jartzen du, honek ikaskuntza-mailaren ezagutza linguistikoari dagozkion eguneratze-erregelak era ditzan. Ezagutza hau honelaxe legoke osaturik: aztergai dugun ikaskuntza-mailako tarte hizkuntz ereduaz eta egitura linguistikoen detektatze-tresnek erabiltzen duten ezagutza linguistikoa. Honela bada, ezagutzaren eguneratzean detektatutako tarte hizkuntz egitura komuna balitz bezala markatuko du eta maila honetako ikasle-multzoaren banakako tarte hizkuntzak berrikusteko erregela bat sortzen du, eta aztertutako egitura propioa dutenen kasuan, ez da gehiagotan hala agertuko behin une hau igaro eta gero. Bestalde, sortzaileak egituren detektatze-tresnek erabiltzen duten ezagutza linguistikorako eguneratze-erregelak ere sortzen ditu.

HITESen moduluak funtzionalak diren moduluen artean eta datu-moduluen artean bereizten ditugu. Datu-moduluek erazagupen-ezagutza bezainbat (ikasle-eredua, tarte hizkuntz eredu) jasotzen dute prozedura-ezagutza (atazen egitura).

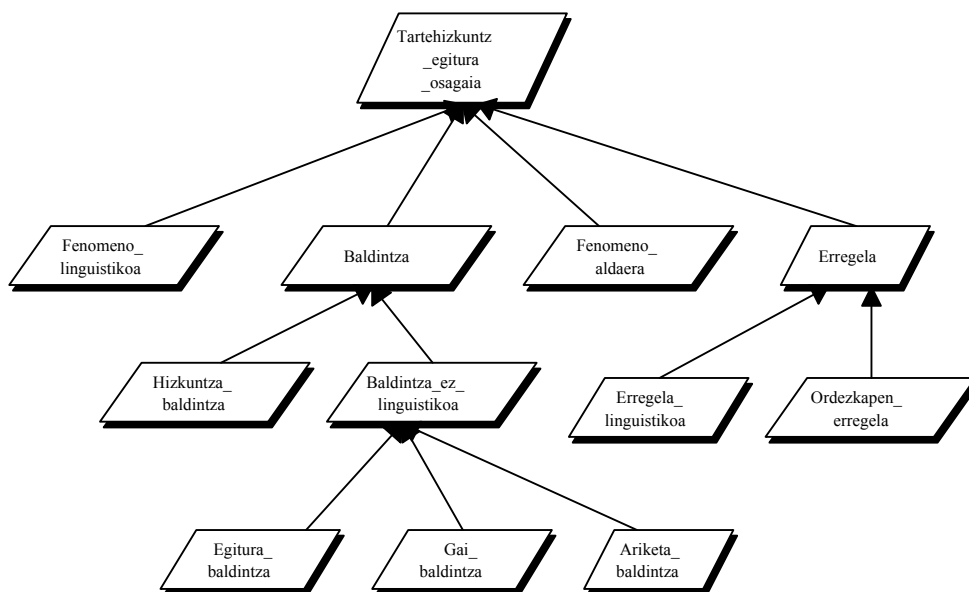
Osagai lagungarriak

HITESen osagai lagungarrien artean lan-memoriaren osagaiak, ikasle-eredua, tarte hizkuntz egitura eta saioa definituko ditugu. Azken bi hauek V.19 eta V.20 irudietan hurrenez-hurren landu ditugu.



V.19 irudia. Saio baten osagaien objektu-hierarkia.

§ V.4.2.1-n azaldu genuen bezala lan-memoriaren osagaiak testua, saioa eta agenda dira. *Testuak* bere ezaugarriak jasotzen ditu (hitz-kopurua, testu-mota (narrazioa, elkarrizketa...), ...). Eta *agendak* unean-uneko atazak, aktibaturik daudenak eta zai daudenak gordetzen ditu. Saio baten osagaien artean (ikus V.19 irudia) bada zenbait fenomeno erregela-multzoa baizik ez dena eta geroan tartehizkuntz egituretan sartzeko fenomeno linguistiko bilaka daitekeena. § V.4.2.1-n aipatu genuen ezagutza iraunkorra eta dinamikoa ere aurkitzen dugu. Azkenik, saioaren une bakoitzean saioaren ibilbidea markatuko diguten *lanak* izango ditugu, hau da, burutu beharreko atazak. Amaiera gisa *ariketa* genuke, nolabait saioaren input-a definitzen duena. Honela bada, saio batek ikasle-talde baten tartehizkuntzaren azterketaz dihardunean, ikasleek testu hura idatzi zuten arrazoiaren baitan testuak sailkatuz burutu ohi da azterketa, hau da, irakasleak adierazitako ariketaren baitan. Bukatzeko, V.20 irudian IV. kapituluaren sakon landu genituen tartehizkuntz egituren osagaiak laburbiltzen dira: *fenomeno_linguistikoa*, *baldintza*, *fenomeno_aldaera* eta *erregela*.



V.20 irudia. Tartehizkuntz egituraren osagaien objektu-hierarkia.

V.7.2 Exekuzio-adibide bat HITESen.

Ondoko adibidean HITES eta psikolinguistaren arteko saio baten zati bat azalduko dugu. Bi agenteen arteko elkarrekintza erakutsi ahala, sistema barruan martxan jartzen diren atazak eta helburuen prozesua jarraituko dugu.

HITES: Menu Nagusia	
Tartehizkuntza kontsultatu	<input type="checkbox"/>
Tartehizkuntza eratu	<input checked="" type="checkbox"/>
Bukatu	<input type="checkbox"/>

HITES: Tartehizkuntza eratzen	
Noren tartehizkuntza landu nahi duzu?:	
Ikasle bat	<input type="checkbox"/>
Ikasle talde bat	<input checked="" type="checkbox"/>

V.21 irudia. Menu nagusia.

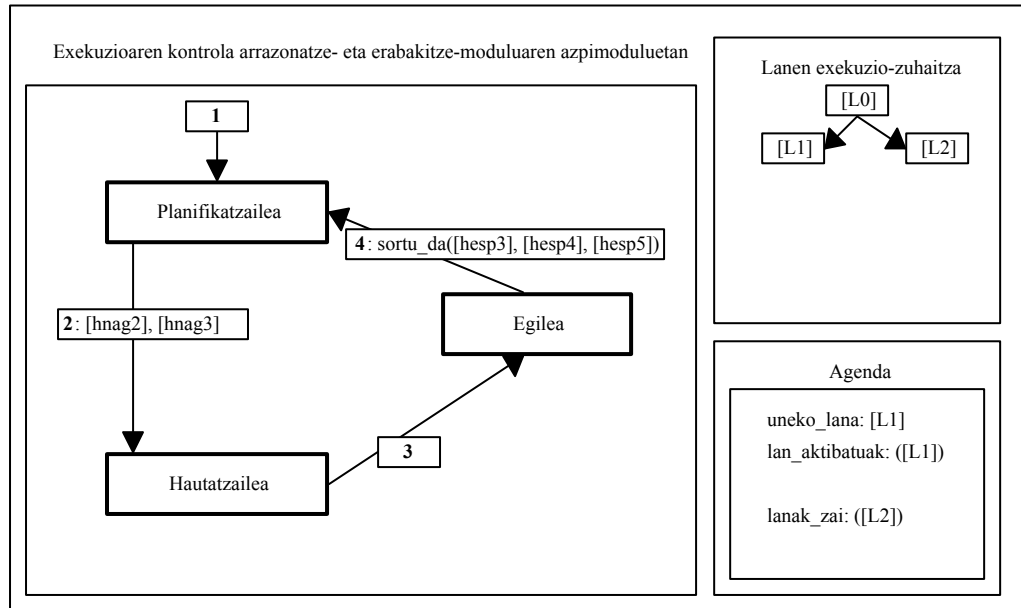
V.22 irudia. Tartehizkuntzaren eraketa.

V.21 eta V.22 irudietan ikusten denez saio konkretu honetan psikolinguistak ikasle-talde baten tartehizkuntzaren eraketa-prozesua lantzea erabakitzen du. Une honetan arrazonatze-eta erabakitze-moduluaren planifikatzailea saioaren helburu nagusiak eguneratzen ditu (ikus V.23 irudia). Ondoren, agendaren garapenarekin batera, momentu horretatik aurrera ondoko agenteek nola egiten dute lan azalduko dugu:

- Planifikatzailea.
- Hautatzailea.
- Egilea.

Planifikatzaileak abisua ematen dio hautatzaileari esanez helburuak sortu direla eta ondorioz lanak sor ditzakeela. Irudian ikusten denez bi helburu nagusi sortu dira: *ikaskuntza-mailaren ereduaren tartehizkuntza eguneratzea [hnag2]* eta *ikaskuntza-maila bakoitzean diharduten tresna linguistikoaren ezagutza eguneratzea [hnag3]*.

Hautatzaileak ataza eta helburuen ezagutza-baseak aztertuta lan bana sortzen du helburu bakoitzeko. Lehenengo helburuarentzat [hnag2] sortutako lana, [L1], agendako lan aktibo gisa identifikatzen du, [hnag3] izeneko helburuarentzat sortutakoa, [L2], berriz, zai dauden lanen artean kokatzen du.



V.23 irudia. Ataza eta helburuen prozesua elkarrekintzan zehar (I).

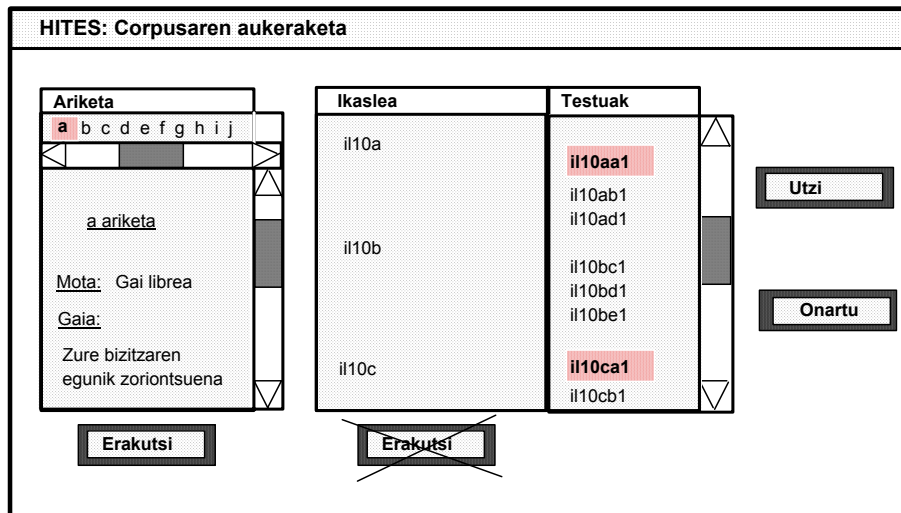
Egileak ordutik aurrera exekuzioaren kontrola bere esku izango du. Honek [L1] lana martxan jartzen du, hau da, uneko lana bezala identifikatzen du, eta dagokion ataza exekutatzera pasatzen da: kasu honetan ez-primitiboa den ataza dugu (*Ikasle_multzoaren_tartehizkuntza_sortzea*). V.11 irudian adierazten dugunez, ataza honen kontrol-egituraren une batean egileak helburu espezifiko berrien sorrera topatzen du, hain zuzen ere [hesp3], [hesp4] eta [hesp5]. Hau dela eta, egileak bi gauza egiten ditu: alde batetik, martxan dagoen [L1] izeneko lana gelditzen du zai dauden lanen artean kokatuz, eta, beste aldetik, planifikatzaileari abisatzen dio helburu espezifiko berriak sortu direla esanez.

Planifikatzaileak helburuak berrantolatzen ditu. Kasu honetan [hesp3] izeneko helburua zai bezala markatzen du planifikatzaileak. Ikasle baten ezaugarri pertsonalei buruzko informazioa lortzea da helburu horren zeregina, horretarako IRAKAZI sistemaren laguntza behar duenari. IRAKAZI sistema, ordea, ez dago aktibatua une horretan. Planifikatzailea jakinaren gainean dago, kanpoarekiko harremanari buruzko informazioa baitu. Beraz, [hesp3] helburua betetzerik ez dagoenez, planifikatzaileak ez du helburua aktibatzen hautatzailea helburu honentzako lana sortzen saia ez dadin.

Hautatzaileak, honen guztiaren ondoren, [hesp4] eta [hesp5] helburuak sortuko ditu, planifikatzaileak abisatuta. Lan berriak sortuko ditu, eta agendan dauden lanak berrantolatu.

Egileak une horretan [L4] lana martxan jartzen du [hesp4] eta [hesp5] helburuak lortzeko behin lanak sortuta (ikus V.25 irudia), hauek dira [L4] eta [L5] hurrenez hurren.

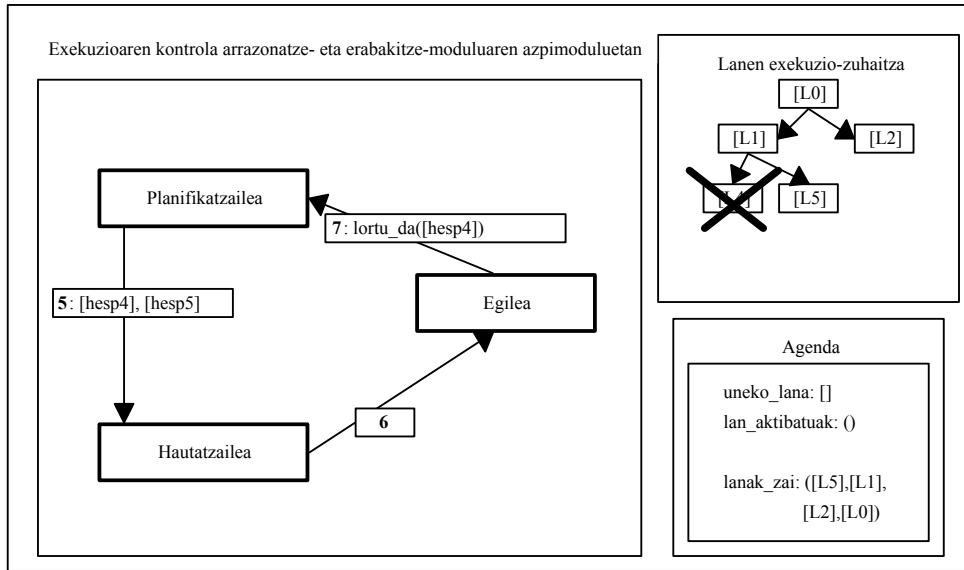
Momentu honetan, *psikolinguistak* corpusaren aukeraketa egin dezan interfazea aktibatzen da une horretan. V.24 irudian elkarrekintza honetan erabiltzen den pantaila erakusten da.



V.24 irudia. Corpusaren aukeraketa.

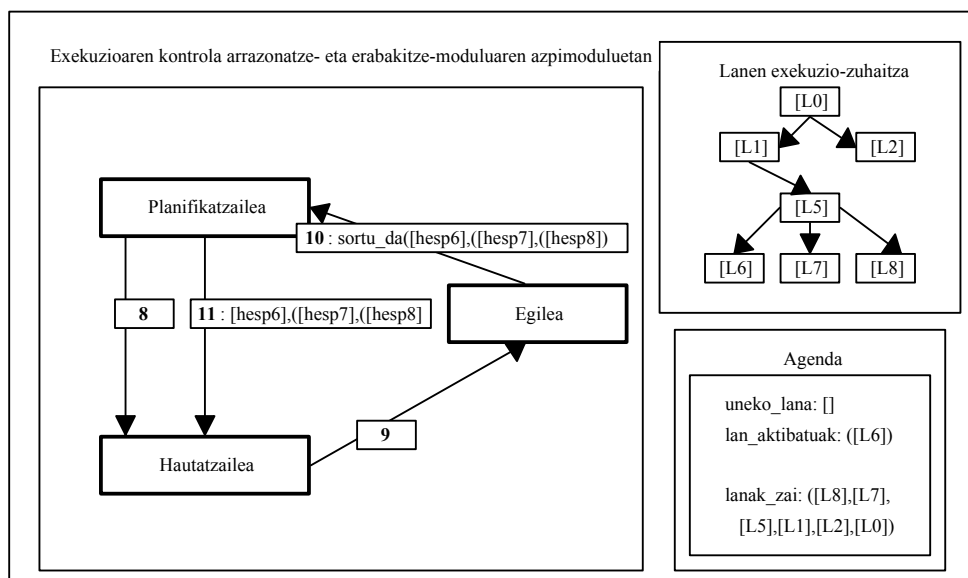
Pantaila honen bidez psikolinguistak, testuak aukeratzeaz gain, ikasleek testuak idatzi aurretik irakasleak bidalitako ariketari buruzko informazioa jaso dezake. Horretaz gain, ikasle bati buruzko informazioa ere eska dezake. Dena den, azaltzen ari garen adibidean, une honetan IRAKAZI sistemarekiko komunikaziorik ez dagoela suposatu dugunez, ikasleen informazioa eskuratzeko *Erakutsi* izeneko botoia desaktibatuta dago pantailan.

V.25 irudian [L4] lana martxan jartzen denetik burutzen den arteko prozesua adierazten da. Aurretik azaldutako pantailan psikolinguistak *onartu* botoia sakatzen duenean [L4] lana bukatutzat ematen da. Momentu horretan egileak abisatzen dio planifikatzaileari [hesp4] helburua bete dela esanez. Planifikatzaileak helburua *betetzat* ematen du, eta, ondoren, hautatzaileak [L4] lana ezabatzen du lanen exekuzio-zuhaitzetik. Kontuan izan, ezabaketa hau beti zuhaitzaren hosto bat izango dela. Hala ere, ezabaketa ez da mugatzen hostoa ezabatzeraz, hostoaren gurasoaren beste umeak (hots, lanak) ere aztertzen dira, eta, hauetako baten helburua betetako helburu bera izanez gero, ume hori ere ezabatuko da.



V.25 irudia. Ataza eta helburuen prozesua elkarrekintzan zehar (II).

Eman dezagun psikolinguistak corpora aukeratu duela dagoeneko, eta *Ikasle_baten_tartehizkuntz_egituren_sorkuntza* izeneko atazaren exekuzioan zehar une batean lana gelditzen dela [hesp6], [hesp7] eta [hesp8] izeneko hiru helburu espezifiko paralelo (ikus V.11 irudia) sortzen direlako. Esan bezala, planifikatzaileak helburuak tratatzen ditu; gero, hautatzaileak aipatu helburuei dagozkien atazak identifikatu, eta lan berriak sortzen ditu, hauek dira, [L6], [L7] eta [L8]. Ondoren, egileak exekutatu beharreko lana hautatzen du (ikus V.26 irudia).



V.26 irudia. Ataza eta helburuen prozesua elkarrekintzan zehar (III).

Suposa dezagun hainbat exekuzio-pausu eman ondoren, *Erregela_linguistikoen_ereduak_atera* (ikus V.11 irudia) izeneko ataza martxan dagoela. Ataza honen zehar, besteak beste, testuen hitzen analisiak lantzen dira. Konkretuki, psikolinguistak aktibatutako erregelak lantzen ditu interpretazio bakoitzeko (ikus V.27 irudia). Goitik beherako metodologia aplikatuz, aztertzen ari den (*maila_aztergai*) eta honen gorago dagoen hurrengo maila (*goiko_maila*) konparatu eta desberdinak diren interpretazioak aurkezten zaizkio psikolinguistari. Hau da, adibidez, *goiko_maila* izenburupean dauden analisiak goiko mailan baziren, baina aztertzen ari garen ikaskuntza-mailan baztertu dira. Hau bi modu desberdinetan irakur daiteke: alde batetik, aztertzen ari garen mailan desagertutako interpretazioa ez dela egokia; edo, beste aldetik, egokia izan arren, oker analizatu dela; azken kasu honek esan nahi du aztergai dugun ikaskuntza-mailaren ezagutza findu behar duela psikolinguistak.

Ezagutzaren finketa dela eta, hitzen analisi morfologikoak erakusten dira desanbiguaziorako erregela morfosintaktikoak aplikatu aurretik, eta erregelak aplikatu ondoren. V.27 irudian agertzen den adibidean, erregelak aplikatu ondorengo analisiari erreparatzen badiogu, *arraio* hitzaren kasuan ikusten dugu aztertzen ari garen ikaskuntza-mailaren kasuan *IZE* interpretazioa desagertu dela. Honen aurrean psikolinguistak bi irtenbide ditu:

- Hiztegiaren ezagutza aztertzea
- Erregela morfosintaktikoen ezagutza aztertzea

Suposa dezagun aztergai dugun ikaskuntza-mailan *jarraiki* lema ezaguna izan arren ikasleek ez dakitela aditz trinko gisa erabiltzen. Kasu horretan psikolinguistak *jarraitu* eta *ADT* (aditz trinkoa) itemak klikatu eta ezagutza ezezagunatzat markatuko du *hiztegiaren finketa* izeneko aukeraren bidez. Ondoren, sistemak analisisa berriro egingo du.

HITES: Tresna linguistikoaren ezagutzaren finketa

Hiztegiaren finketa

Erregela linguistikoaren finketa

Hitzen analisiak erregelak aplikatu ondoren

"<arraio>"
GOIKO_MAILA
 "arraio" IZE ARR + DEK ABS MG @OBJ @SUBJ
MAILA_AZTERGAI **MS_250**
 "jarraiki" ADT A1 NOR_NORI NR_HI NI_HU
 LEM_LEKHH LUZ6 BOK MS_24 @+JADNAG
 "<esan>"

Hitzen analisiak erregelak aplikatu aurretik

"<orduan>"
 "orduan" LOT LOK @LOK LUZ6 KTS
 "<nola>"
 "nola" ADB ADO LUZ4 BOK AORG @ADLG
 "<arraio>"
 "arraio" IZE ARR
 "jarraiki" ADT A1 NOR_NORI

 "<esan>"
 "esan" ADI
 "ezan" ADL MDNC NOR_NORK NR_HU NK_HI LEM_LEOEZS

Onartu Utzi

V.27 irudia. Ezagutza linguistikoaren azterketa.

Eman dezagun, orain, analisi berrian oraindik ere analisi zuzena ez dugula; edo, pentsa dezagun aztergai dugun ikaskuntza-mailan ikasleek badakitela *jarraiki* aditz trinko bezala erabiltzen. Kasu bietan psikolinguistak aplikatutako erregelak aztertzeari ekingo dio *erregela linguistikoaren finketa* aukera hautatuz: azaldutako adibidean *MS_250* identifikadorea duen erregela aukeratu du psikolinguistak.

Hautatutako erregela editatzean (ikus V.28 irudia), psikolinguistak aurkitzen du *arraio* hitza analizatzean zergatik ez den agertu *IZE* interpretazioa aztergai dugun ikaskuntza-mailan: *IZE* interpretazioa ezabatu (*REMOVE*) du erregelak eskubiko hitzaren interpretazioaren batek *ADL* ezaugarria duelako (*1 ADL*). V.27 irudian ikusten denez eskubiko hitza analizatzean *ezan* interpretazio desbideratzeduna (*LEM_LEOEZS*) sortu da ikaskuntza-maila honetan interpretazio horretan *ADL* ezaugarria dago, eta, ondorioz, *MS_250* erregela aktibatuta da.

Guzti honen aurrean psikolinguistak erabakitzen du *MS_250* erregela ez dela baliagarria aztergai dugun ikaskuntza-mailan. Erregela ezabatu eta berria sortu, edo erregela bera egunera dezake. Kasu honetan erregela eguneratzen du (ikus V.28 irudia) ondokoa adieraziz:

IZE interpretazioa ezabatu (*REMOVE*) eskubiko hitzaren interpretazio guztiek *ADL* ezaugarria baldin badute (*IC ADL*).

Bukatzeko, ez ahaztu psikolinguistak edozein momentutan erregelak editatu eta eguneratu ditzakeela prozesuan zehar.

HITES: Erregelen edizioa

Erregela baten ezaugarriak

Identifikadorea:

Deskribapena:

Erregela agindua:

Agindugaia:

Adibidea:
(ADI ...)
(IZE ...)

Egitura baldintzak:

Hizkuntza baldintzak:

Mota linguistikoa:

V.28 irudia. Erregela linguistikoen edizioa.

Azaldutako elkarrekintza burutu ahala *Erregela_linguistikoen_ereduak_atera* izeneko ataza martxan dago. Honetaz gain, saioaren hasieran sortu zen *[L2]* izeneko lana agendan dago *zai* dauden lanen artean. Lan hau, berriz, aktibatuko da psikolinguistak eguneraketa bat egiten duenean ezagutzaren finketan zehar, eta aktibazio honek *[L2]* lana martxan jartzea ekarriko du.

Adibidearen azalpenari amaiera emateko asmoz, saioan zeharreko exekuzioan salto txiki bat emango dugu:

Erregela_linguistikoen_ereduak_atera izeneko ataza bukatutzat emango dugu, eta suposatuko dugu dagoeneko fenomeno linguistikoen identifikazio-prozesura heldu dela psikolinguista (ikus V.11 irudia). Ataza honen barruan *fenomeno_linguistikoen_zehaztapena* izeneko helburua *[hesp1]* dugu. Helburu hau hiru ataza desberdinen bidez lor daiteke: *fenomenoa_diskriminatzea_ereduen_laguntzaz*, *fenomenoa_diskriminatzea_psikolinguistaren_laguntzaz* eta *ikaslearekiko_elkarrekintzaz_diskriminatzea*, – azken atazak IDAZKIDEREKIKO komunikazioa eskatzen du –.

Ondoren azalduko dugu nola burutzen den aipatutako helburua *fenomenoa_diskriminatzea_psikolinguistaren_laguntzaz* izeneko atazaren bidez:

V.29 irudian psikolinguistarekiko interfazea azaltzen da. Honetan psikolinguistari fenomeno linguistikoak aurkezten zaizkio. Honek fenomenoak editatzen ditu, zuzendu eta automatikoki lortu ez den informazioa osatzen du, hots, fenomeno linguistikoaren deskribapena. Behin informazioa aztertuta (hizkuntza-baldintzak, egitura-baldintzak, eta abar), psikolinguistak erabaki bat hartuko du: fenomenoa ezabatu, fenomeno berria sortu, edo onartu, berak egindako zuzenketekin, noski. Azken kasu honetan fenomeno aztertuen zerrendara pasako da fenomenoa.

HITES: Fenomeno linguistikoaren azterketa
Erregelen edizioa

Aztertzeko fenomenoak

FL_25: LEOEAE LEGBE

FL_32: LEOEHMm
MS_18

➔

Onartu

Utzi

Fenomeno aztertuak

FL_13: LEOEKG
MS_16

FL_90: LEGBE

FL_77: LEOETD

FL_10: LEKHH
MS_250

Erakutsi

Ezabatu

Erakutsi

Ezabatu

Berria sortu

Editatu

Fenomeno linguistiko baten edizioa

Identifikadorea	FL_32
Deskribapena	<i>Aditzen erabilera esaldi hasieran</i>
Osagaiak	LEOHMm MS_18
Detekzio-zerrenda	10
Adibideak	(Esan zidan ...) (Ikusi zuen ...)
Hizkuntza baldintzak	(ADI LEM)
Egitura baldintzak	(Esaldi hasiera)
Mota linguistikoa	morfo_sintak

Onartu

Utzi

V.29 irudia. Fenomeno linguistikoaren diskriminazioa.

Irudian ikusten denez, momentu batean, aztertutako fenomeno bat ezabatzea erabaki dezake psikolinguistak: agian, fenomeno berri bat lantzean konturatzen da aurretik landutako bat ez zela egokia. Dena den, hau egiteko muga bat du: HITES sistemak erabakitzen duenean fenomeno onargarria dela (hots, psikolinguistak aspaldian onartu zuen saio berean) fenomeno atzigarri egongo da interfazean, baina ezabatzeko aukerarik ez du izango. Une horretatik aurrera, fenomeno hori aztertzen ari den tartehizkuntzaren egituraren batean integra daiteke.

V.8 Implementazio-egoera eta landu beharrak.

Objektuetara zuzendutako programazioan oinarritutako metodologian inplementatu da HITES sistema. Moduluen inplementazioan erabili den lengoia CLIPS izan da.

Atazen inplementazioak halaber CLIPS-en burutu dira, nahiz eta baditugun C-n eta Perl-en inplementatutako batzuk; honen adibide dugu *Erregela_linguistikoen_ereduak_atera* izeneko ataza. Ataza hau corpusaren trataera osoaz arduratzen dena da: dagozkion testuak hautatu egiten ditu, trataerarako prestatzen ditu (filtroak...), eta testu hauetan gertatzen diren erregela-linguistikoak identifikatzen ditu. Horretarako beharrezko dituen tresna linguistikoekin komunikatzen da.

Kontrol-sistemaren zatiari dagokionean, oraingoz simulazio erdiautomatikoak baizik ez da garatu. Gerora begira alabaina (Arregi, 95)-n garatua izan den atazen aktibazioa gure kontrolean integratzea proposatuko genuke. Izan ere, modu honetara atazen exekuzio bat inplementatuko genuke, non beren kontrol-egiturak ezagutza prozeduralaren zati izango bailirateke. Gure inplementazioan oraingoz ataza objektuaren slot gisa ez dira sartu kontrol-egiturak. Aipatu integrazioak halaber, diseinu-mailan aldaketak eskatzen ditu, izan ere (Arregi, 95)-n atazen aktibazioa atazaren onarpen-baldintzen baitakoa baita, eta gure kasuan aldiz, atazen onarpen-baldintzez gain, baditugu burutu beharreko lanak seinaleztatzen dizkiguten helburuak ere.

Azkenik, psikolinguistarekin izandako komunikazio zati desberdinak inplementatuak izan dira, baina interfaze bakar batean integratzea da falta dena. Integrazio honen alderdirik nagusienak tresna linguistikoekiko komunikazioa eta tartehizkuntz ereduen eguneratzea eta kontsulta genituzke. Psikolinguistaren interfazea inplementatua izan ez bada ere, alderdirik nagusienak beste esparru batzuetan erabili dira. Tresnen integrazioari dagokionean esaterako, VI. kapituluan ikusiko dugu IDAZKIDE sisteman TCL/TK-n interfaze bat garatu dela eta hainbat tresna linguistiko honen barruan integratu dela. Eta tartehizkuntz ereduen eguneratze eta kontsultari dagokionez GBC (Elorriaga et al., 98) erabat baliagarria izan zaigu egindako probetan, CLIPS-en inplementatutako ezagutza-baseen eraketarako tresna dugu GBC. Tresna honen integrazioa (Arruarte, 98)-n ikus dezakegun bezala, jadanik erabili izan da ezagutza-baseen sormeneko interfazeetan.

VI. IDAZKIDE eta IRAKAZI:

idazte-prozesurako eta diagnosirako laguntza-sistemak.

Kapitulu honetan IDAZKIDE eta IRAKAZI sistemak azalduko ditugu. IDAZKIDE idazte-prozesurako ingurune lagungarria da; ikaslearekin egiten du lan eta bai HITESein bai IRAKAZIrekin komunikazioa du. IRAKAZIk ikaslearen ezagutza linguistikoaren diagnosian eskaintzen dio laguntza irakasleari. Esaterako, ikasleen desbideratzeen arazoak aztertzeko erabiltzen dugu sistema hau. IRAKAZIk hizkuntz irakaslearekin egiten du lan eta bere emaitzak bai IDAZKIDEk bai HITESek erabiltzen dituzte. Kapitulu honetan gehienbat IDAZKIDE azalduko dugu garatuagoa baitago eta ekarpen gehiago baititu. IRAKAZIri buruz, ordea, ideia orokor batzuk emango ditugu.

Esan bezala, aurreko kapituluaz azaldutako HITES sistema adimendunak eta kapitulu honetan deskribatuko ditugun bi sistemek bigarren hizkuntzaren ikaste-prozesuaren azterketarako MUGARRI ingurunea osatzen dute. Sistemek elkarrekiko komunikazioa izateaz gain hainbat ezagutza-base komun dute, konkretuki, tartehizkuntz ereduak, diagnosi-ezagutza, ikasle-eredua eta tresna linguistikoaren ezagutza. Aurreko bi kapituluetan tartehizkuntz ereduak eta ezagutza linguistikoak aztertu ditugu; honetan, berriz, ikasle-eredua eta diagnosi-ezagutza landuko ditugu, besteak beste.

Ondoren, IRAKAZI eta IDAZKIDE sistemak deskribatuko ditugu, bakoitzaren interfazea azalduz. Horretaz gain, ikasle-ereduaz, diagnosi-ezagutzaz eta tresna linguistikoaren integrazioaz ere arituko gara.

VI.1 Sarrera.

IDAZKIDE eta IRAKAZI sistemek ondoko bi arazoei erantzuna eman nahi diete:

1. Euskara ikasten dutenek orokorrean oztopo handia aurkitzen dute hizkuntzaren morfologia ikastean, hizkuntza eranskaria baita euskara.
2. Euskaltzaindiak maiz argitaratzen dituen arau berriak ez dira heltzen hizkuntz ikasle guztiengana.

IDAZKIDE ingurune lagungarria da, eta ez du sistema tutoreen filosofia jarraitzen. Bere helburua ikaslearen ikaste-prozesuan zehar laguntza ematea da, beraz, ez dago alde aurretik erakutsi beharreko ezagutza linguistiko antolatua ikasleari irakasteko. IDAZKIDEn lan hori

ikasgelan burutzen dela suposatzen da, hortaz, sistema osagarria da ikaste-prozesuan, eta ez gidaria. Sistemak hainbat tresna linguistiko eskuragarri jartzen dio ikasleari idazte-prozesuan zehar. Inguruneak interfaze grafikoa eskaintzen dio ikasleari hainbat eragiketa egiteko aukera emanez: hitz baten analisi morfologikoa egitea, hitz baten definizioa erakustea, hitz desbideratu baten proposamenak ematea, e.a. Sistemak ikaslearekiko elkarrekintzan honen ikaskuntza-maila, ikaste-estrategiak eta hainbat ezaugarri izango ditu kontuan.

Ikaslearentzako sistema lagungarria izateaz gain ikaste-prozesuaren azterketarako osagai ere badugu IDAZKIDE. Honegatik, ikaslearekiko elkarrekintzetan sortutako informazioa ikaste-prozesuaren azterketarako erabiliko da.

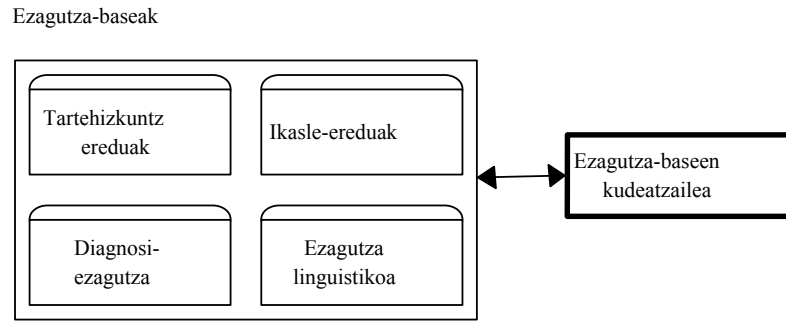
IRAKAZI inguruneak irakaslearekin elkarreragiten du, bai ikasleari buruzko ezaugarriak jasotzeko bai ikasleak erabiltzen dituen egitura desbideratuen diagnostia lantzeko. Prozesu honetan irakaslea da diagnosiari buruzko informazioa ematen duena. Egitura linguistikoak HITESek sortutako tartehizkuntz ereduetatik jasotzen ditu IRAKAZIk. Honek irakasleari eskaintzen dion interfaze grafikoa ikasleari buruzko ezaugarriak eta bere egitura linguistikoak jasotzeko aukera emateaz gain, ikasleek ikasgelan egindako ariketei buruzko informazioa ere jasotzen du. Gainera, informazioa lantzeko hainbat tresna ere eskaintzen du: egitura linguistikoen sailkapenak, egiturak erabiltzeko arrazoi posibleen zerrendak, aldi berean gertatzen diren egitura linguistikoen azterketa, e.a. Azkenik, IRAKAZIren bidez egindako diagnostia IDAZKIDEk erabiliko du.

Laburbilduz esango dugu IDAZKIDE sistema elkarreragileak, alde batetik, HITES sistemak ikaskuntza-maila bakoitzeko sortutako ereduak dituela oinarri, eta, beste alde batetik, IRAKAZIren bidez jasotako diagnosirako informazioa erabiltzen duela.

VI.2 Ikasle-eredu komuna.

VI.2.1 Ezagutza-baseak.

VI.1 irudian MUGARRI ingurunean HITES, IRAKAZI eta IDAZKIDE sistemek konpartitzen dituzten lau ezagutza-base azaltzen dira: tartehizkuntz ereduak, ezagutza linguistikoa, ikasle-ereduak eta diagnosi-ezagutza. Azken honek tartehizkuntz ereduaren desbideratze-egiturak oinarritzat hartzen ditu. Ezagutza-baseak kudeatzeko ezagutza-baseen kudeatzailea dugu. Honek HITES, IRAKAZI edo IDAZKIDEk eskatutako informazioa eskuratu edo eguneratuko du hala eskatuz gero.



VI.1 irudia. MUGARRI ingurunearen ezagutza-baseak

Tartehizkuntzaren eredu kontzeptuala IV. kapituluaren ikusi genuen. Ereduaren gestioa, berriz, hau da, eskuratze-prozesua zein eguneratzea, V. kapituluaren aztertu genuen. Ezagutza linguistikoa ereduaren tratamendua ere aipatutako bi kapituluaren landu genuen. Ikasle-eredua eta diagnosi-ezagutza, ordea, orain arte sakondu gabeko ezagutza-baseak ditugu. Bi hauek HITESek ere erabili arren funtzio garrantzitsuagoa betetzen dute IDAZKIDE eta IRAKAZI sistemetan. Hau dela eta, aipatutako bi ezagutza-baseak kapitulu honetan aztertuko ditugu.

VI.2.2 Ikasle-ereduaren gunea.

Ikasle-ereduaren gunea HITES, IDAZKIDE eta IRAKAZI sistemek erabiltzen dute. Dena den, hirurek ez dituzte ikasle-ereduaren ezaugarri guztiak erabiltzen. Batek ezaugarri bat erabiltzen duen bitartean beste batek eguneratu egin lezake eta hirugarrenak kontuan eduki ere ez. Beraz, ezaugarri bakoitza zein sistematan den erabilgarri adieraziko dugu (ikus *sistemak* ezaugarria VI.2 irudian). Horretaz gain, ezaugarri bakoitzaren portaera ere azalduko dugu definizioan: ikasle-ereduaren ezaugarri bat saioan zehar aldakorra denean dinamiko dela esango dugu, bestela, estatikoa dela kontsideratuko dugu. Ezaugarri bera sistema baterako dinamiko izan daiteke eta beste baterako, ordea, estatikoa. Adibidez, HITESen ikaslearen ikaste-estrategiak estatikoak dira, hau da psikolinguistikak kontsultatu bai, baina ezin ditu eguneratu; IRAKAZIn, berriz, ikaste-estrategiak aldakorak dira irakasleak ikaslea ezagutzen duela, edozein kasutan jarraitzen dute estatikoak izaten saio baten barruan; azkenik, IDAZKIDEn dinamikoak dira, ikasleak erabilitako tresna linguistikoen arabera bere ikaste-estrategiak egunera daitezkeelako saioan zehar. Beraz, *ezaugarria* kontzeptua definitzean zein sistemari dagokion esateaz gain, horien artean dinamiko zein kasutan den ere adieraziko dugu.

```

structure Ezaugarria;
subtype-of: Ikasle_ereduaren_kontzeptua;
parts:
  balioa: instance(Ezaugarri_balioa);
properties:
  identifikadorea: string;
  sistemak: list(symbol); value={HITES, IRAKAZI, IDAZKIDE}
  dinamikoa: list(symbol); value={HITES, IRAKAZI, IDAZKIDE}

concept Ezaugarri_balioa;
subtype-of: Ikasle_ereduaren_kontzeptua;
properties:
  mota: symbol;

concept Ezaugarri_balio_simplea;
subtype-of: Ezaugarri_balioa;
properties:
  balioa: symbol;

concept Ezaugarri_balio_konposatua;
subtype-of: Ezaugarri_balioa;
properties:
  balioa: list(symbol);

```

VI.2 irudia. Ezaugarria kontzeptuaren definizioa.

Ikasle-ereduan *ezaugarriak* izeneko slotak izango dugu, eta slot honek *ezaugarria* motako objektuen zerrenda izango du. Hau dela eta, ikasle-eredua honela definitzen dugu VI.3 irudian:

```

structure Ikasle_ereduaren_gunea;
subtype-of: Ezagutza_basea;
parts:
  ezaugarriak: set(instance(Ezaugarria));
  testuak: set(instance(Testua));
  tartehizkuntza: instance(Tartehizkuntza);
properties:
  identifikadorea: string;
  hizkuntz_maila: integer;

```

VI.3 irudia. Ikasle-ereduaren definizioa.

Ikasle-eredua errepresentatzen duten objektu nagusiak ikusi ondoren, ikasle-ereduaren gunean ze informazio gorde dugun aipatzeari ekingo diogu:

- euskaltegia
- irakaslea
- taldea
- ikasketak
- ama-hizkuntza/k
- ama-hizkuntza_ren_maila
- euskaldun berria (true/false)
- noiztik berria / euskalkia euskaldun zahar izanez gero
- hizkuntzak (entzumena/mintzamena/idazmena/irakurmena)

- euskararen erabilera (0tik 100era -> %)
- non erabiltzen duen
- non erabiliko duen
- ikasteko arrazoi pertsonalak
- hasieratik jarraitutako ikaste-prozesua (garaia/non/maila/ordu_kopurua_astean)
- eguneroko bizitzan dauzkan edo eduki dituen erreferente euskaldunak (nongoa eta noiz "txikitan/gaztetan/helduaroan")

VI.2.3 Diagnosi-ezagutza.

Diagnosi-ezagutzan ikasleak bere tartehizkuntzaren egitura linguistiko desbideratuak erabiltzeko eduki ditzakeen arrazoiak gordetzen ditugu. Arrazoi horiek batzuetan ikaste-estrategietan izaten dute jatorria. Hau dela eta, sistemak ikasleari gomendioak emango dizkio erabili beharreko ikaste-estrategiei dagokienez. Sistema automatiko batek bidera dezake ikaslea ikaste-estrategia konkrituak landu ditzan; bestelako arrazoietan (adib. ikasteko arrazoi pertsonalak, euskararen erabilera, ...), ordea, ezin du eragin handirik sortu.

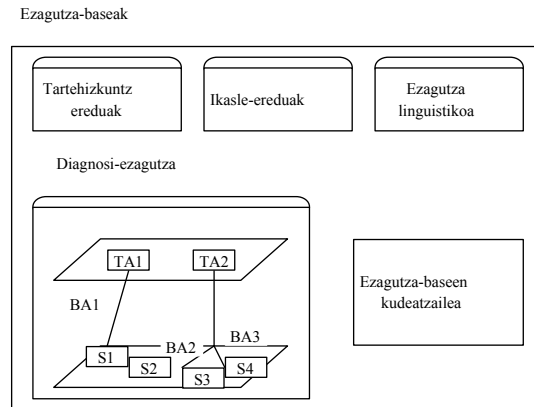
VI.4 irudian ikus daitekeenez bi maila bereizten dira diagnosi-ezagutzan:

- Tartehizkuntz egiturak (TA1, TA2 ...).
- Egiturak erabiltzeko arrazoiak (S1, S2, S3, S4 ...).

Tartehizkuntz egiturak tartehizkuntz ereduen osagaiak dira. IV. kapituluaren ikusi dugunez tartehizkuntz egitura bakoitza fenomeno linguistikoen zerrenda batek definitzen du, konkretuki testuinguru zehatz batean gertatzen diren fenomenoek (fij). Testuingurua egitura erabili den testuaren hainbat ezaugarriak definitzen du:

TA1 = < t1, (f11, f12, f13 ... f1n) >

TA2 = < t2, (f21, f22, f23 ... f2m) >



VI.4 irudia. Diagnosi-ezagutza ezagutza-basea.

Egiturak erabiltzeko sakoneko arrazoiak (S_i) beren aldetik bi multzo nagusitan banatzen dira: ikaslearen barruko faktoreak (B_i) (adib. denbora laburrean hizkuntz maila altua lortu izana) eta ingurunearen faktoreak (I_i) (adib. euskalki batean mintzatzen zen norbaitekin bizi izandako garaia) (ikus C eranskina). Hori dela eta, sakoneko arrazoi hauek bi multzotan bereiztuko ditugu, adibidez:

$$S_1 = \langle B_{11} \dots B_{1n}, I_{11} \dots I_{1m} \rangle \dots$$

Hortaz, diagnosi-ezagutzan tartehizkuntz egitura bakoitzeko egitura hori erabiltzeko arrazoiak biltzen dira. Era honetan, azken finean, tartehizkuntz egiturak ikasle-moten arabera ari gara antolatzen. Bi ikaslek egitura-multzo baten arrazoi berdinak dituztenean ikasle-mota berekoak direla esango dugu.

Irudian ikusten denez tartehizkuntz egituren maila eta egiturak erabiltzearen arrazoi maila bi motatako erlazioek lotzen dituzte: erlazio biunibokoak (adib. BA1) eta erlazio ez-biunibokoak (adib. BA2, BA3).

Erlazio bat biunibokoa da baldin eta tartehizkuntz egitura bat emanda eta sakoneko arrazoi bat emanda ez badago arrazoi gehiagorik emandako egitura erabiltzeko.

Erlazio bat ez-biunibokoa da baldin eta tartehizkuntz egitura bat emanda eta sakoneko arrazoi bat emanda, badago beste arrazoiaren bat ere emandako egitura erabiltzeko.

Diagnosi-ezagutza HITESek erabiltzen du IRAKAZIk eskaintzen duen ezagutza dela medio. IDAZKIDE ikaste-sistema lagungarriak ere diagnosi-ezagutza hori erabiliko du ikaslearekiko elkarrekintzan, baina, ezingo du diagnosi-ezagutza zuzenean aldatu, bai, ordea, zeharka, ikaslearekin izandako saioen berri IRAKAZIri emanez.

VI.3 IDAZKIDE: idazte-prozesurako sistema lagungarria.

VI.3.1 IDAZKIDEn arkitektura.

IDAZKIDE (Díaz et al., 99) ez da hizkuntzak *irakasteko* sistema adimenduna, baizik eta hizkuntzak *ikasteko* ingurune adimenduna. Sistemaren helburua ikaste-prozesuan zehar ikasleari laguntzea da, hauen ikasketa bideratzeko aholkuak emanez. Horretarako, ikasleen ezaugarriak zein ezagutza linguistikoa gordetzen ditugu ikasle-ereduan. Honetaz gain, HITES eta IRAKAZI sistemekin batera osatzen duen MUGARRI izeneko ingurunean tartehizkuntz ereduak gordetzen ditugu ikaskuntza-maila bakoitzeko.

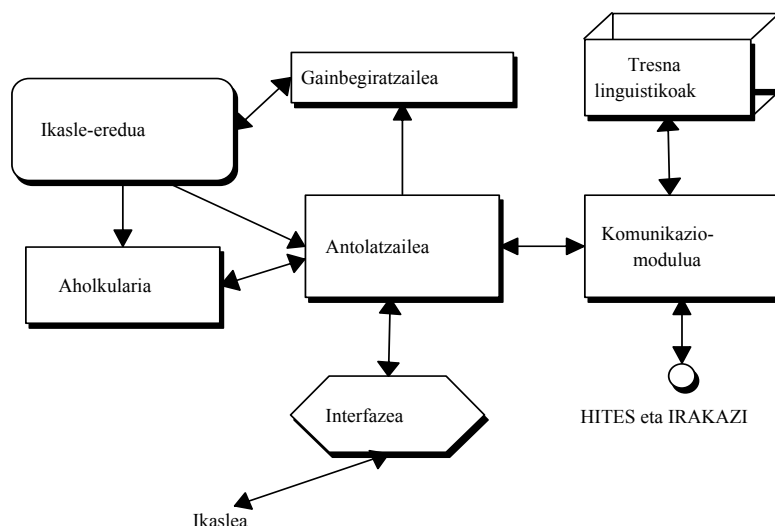
Sistemak ez dauka gordeta alde aurretik erakutsi beharreko materiala, CALL (*Computer Assisted Language Learning*) sistema gehienek duten moduan. IDAZKIDEn oinarria sakonki landutako hainbat sistema independentek (hots, tresna linguistikoak) osatzen du. Tresna linguistikoek komunikazio-modulu baten bidez elkarreragiten dute sistemarekin. Tresnak zeharo independente izateak garrantzi handia du sistemaren hobekuntza inkrementalari begira. Esate baterako, tresna linguistikoak interfazerik gabe exekuta daitezke, hau da, zeharo autonomoak dira, ondorioz interfaze-mota desberdinekin esperimentatzeko aukera ematen digute.

IDAZKIDE zazpi moduluk osatzen dute:

- Tresna linguistikoak.
- Interfaze adaptatibo bat.
- Antolatzailea.
- Aholkularia.
- Gainbegiratzaila.
- Komunikazio-modulua.
- Ikasle-eredua.

Hizkuntz ikasleak interfazearen bidez elkarreragiten du sistemarekin, edozein momentutan tresna linguistikoak erabiltzeko aukera izanik. Antolatzaileak sistema eta ikaslearen arteko elkarrekintza kontrolatzen du: komunikazio-moduluaren kontrola darama. Aholkulariak ikaslearentzako emaitzak egokitzen ditu ikaslearen arabera, beti ere, ikasle-eredua oinarri izanik. Azkenik, gainbegiratzailak antolatzailearen prozesua aztertuko du, eta, egoki balitz, ikaslearen eredua eguneratuko luke.

VI.5 irudian arkitektura orokorra aurkezten dugu moduluen arteko elkarreragiteak adieraziz. Hurrengo ataletan modulu bakoitza sakonago azalduko dugu.



VI.5 irudia. IDAZKIDERen arkitektura.

VI.3.2 Antolatzailea, aholkularia eta gainbegiratzailerarekin.

Antolatzailea

Antolatzaileak sistema eta ikaslearen arteko elkarrekintza kontrolatzen du, hiru ikaste-modu desberdinetarako aukera emanez: ikasketa gidatua, esplorazio askea eta ikasketa aske-gidatu konbinatua.

Esplorazio askean ikasleak edozein unetan tresna linguistiko guztiak erabil ditzake edozein testutan hitzen bat aukeratuaz edo berak asmatutako hitzen bat emanez. Adibidez, ikasleak zalantzaren bat duenean sistemari galde diezaioteko, bai testua idazten ari denean, bai testua zuzentzen ari denean. Esplorazio askean laguntza gidatua ez dago aktibatua.

Antolatzaileak, **ikasketa gidatua** erabiltzen denean, hainbat aholku bideratzen dio ikasleari, ikaslearen ezaugarriak eta ikasleak saioan zehar jarraitutako prozesua kontuan izanik. Ez dugu ahaztu behar erabilitako tresna linguistikoek ikaslearen ikaste-estrategiei buruzko gakoak ematen dizkiotela antolatzaileari.

Ikasketa aske-gidatuan ikasleak sistemaren gomendioak jarraitzen ditu batzuetan, baina beste batzuetan gomendioei kasurik ez egin eta esplorazio askea aukeratzen du edozein tresna erabiliz.

Azkenik, ez dugu ahaztu behar antolatzaileak tresna linguistiko bat noiz exekutatu behar duen ere erabakitzen duela.

Aholkulariak

Antolatzaileak tresna linguistiko bat exekutatzeko eskatzen duenean beharrezko informazioa pasatzen dio aholkulariari eta honek emaitza egokitu du tresna linguistikoa egikaritu ondoren. Egokitzea ikasle-ereduaren arabera da.

Gainbegiratzailea

Gainbegiratzailearen funtzioa antolatzailearen zein ikaslearen prozesua aztertzea da, horretarako ikaslearekiko saioen garapena gordeko du eta egoki deritzon ikasle-eredua eguneratzeko aginduak sortuko ditu.

Ikasleak erabilitako tresnak kontuan izanik, gainbegiratzaileak ikaslearen ikaste-estrategiak ebaluatuko ditu. Esate baterako, batzuetan ikasleak idatzitako testuaren hitz konkretu bat aukeratzen du, hitz horri buruzko ezagutza berria lantzeko. Prozesu honetan jarraitutako kontsultek ikaslearen ikaste-estrategiak islatzen dute neurri batean. Hortaz, saio baten ondoren gainbegiratzaileak kontsulta-prozesu hauek aztertutakoan ikasleari modu egokienean laguntzen dioten ikaste-estrategiak gordeko ditu ikasle-ereduan. Ondorengo saioetan antolatzaileak estrategia hauek erabiliko ditu laguntza gidatuz. Adibidez, ikasleak askotan hitzen deklinabidea kontsultatuz gero, bere ikaste-estrategien artean gramatika kontsultatzearena dagoela ondorioztatuko du. Ikaslearen estrategien artean estrategia hori aurretik ez balego ikaslearen ereduan gehituko luke.

VI.3.3 Tresna linguistikoaren integrazioa IDAZKIDEn.

Komunikazio-modulua

Aipatu bezala komunikazio-modulua tresna linguistikoekiko komunikazioaz arduratzen da antolatzaileak eskatuta. Beraz, ikasleak tresna linguistikoei dei egiten dienean edo sistemak berak egoki deritzon komunikazio-modulua aktibatuko da.

Komunikazio-modulua beste sistemekiko komunikazioaz ere arduratzen da, hots, HITES eta IRAKAZIrekin egon daitezkeen harremanez. Dena den, komunikazio-mota hau garatu gabe gelditu da hemen aurkeztu den lanean, eta etorkizunean, heldu beharreko gaia dugu hau.

Tresna linguistikoak

IDAZKIDEn diseinatu eta inplementatutako tresna linguistikoaren garapenean euskararen tratamendu automatikorako sortutako tresnak abiapuntu izan dira. Ingurune honetan IXA

ikertaldeak hainbat arlo landu du: datu-base lexikalak, morfologia, morfosintaxia, hiztegi eleanitzak, kontzeptu-sareak, sintaxia, eta abar. Kapitulu honetan integratu, egokitu eta sortutako tresnak bai hiztegi-mailan bai morfosintaxi mailan garatuak dira. Arlo hauek garrantzi handia dute euskara bezalako hizkuntza eranskari bat ikasterakoan.

Azpimarratzekoa da IDAZKIDEn integratutako tresnek ez dituztela hitzen zerrendak erabiltzen hitzak analizatu, zuzendu edo sortzeko, hizkuntzak ikasteko sistema askotan gertatzen den bezala. Gure kasuan, tresna linguistiko guztietarako komuna den 70.000 sarreradun datu-base lexikala dugu oinarri (Aduriz et al., 98).

Azkenik, ikasleak ikaste-sisteman eskuragarri duen tresna linguistiko bakoitza azaldu aurretik, ez dugu aipatu gabe utzi nahi aplikatutako berrerabilpenaren filosofia. Sistema informatikoen munduan berrerabilpena oso garrantzitsua izan arren, askotan errealitatean berrerabilpen hau ez da posiblea izaten diseinuak espezifikoegiak egiten direlako edota euskarri fisiko zein sistema logiko konkretuetan egiten direlako, eta gero hauek ez direlako eramangarriak izaten. Hemen aurkezten dugun lanean, ordea, hasieratik tresnak berrerabiltzea izan dugu helburu. Lehen aipatutako IXA ikertaldean tresna bera aplikazio desberdinetarako moldatua izan da askotan. Adibide gisa, datu-base lexikala lematizatzaileko (Aduriz et al., 96) eta hizkuntz ikasleen tartehizkuntzaren modelizaziorako (Maritxalar et al., 97) tresnen oinarri moduan erabilia izan da.

Ondoren IDAZKIDE sistema lagungarrian integratu, egokitu eta sortutako tresnak azaltzeari ekingo diogu:

- XUXEN: testu bat emanda desbideratzeak bilatzen dituen zuzentzaile ortografikoa (Aduriz et al., 97; Agirre et al., 92), edo hitz bat emanda hitz estandarra den ala ez esaten duena. Egokitutako zuzentzaileak ikasleen ikaskuntza-maila zein ezaugarrien arabera egingo du zuzenketa. Adibidez, ikaskuntza-maila konkretuari ez badagokio hitz baten ezagutza, hitza agertu arren sistemak ez du desbideratzea markatuko.

- EDBL: 70.000 sarreradun datu-base lexikala (Agirre et al., 95; Aduriz et al., 98). Datu-base honetan hitzen kategoria eta azpikategoria, eta erabilera-adibideak kontsulta daitezke. EDBLk hitzun adituen ezagutza biltzen duenez, ez dugu egokiena ikusten datu-basea dagoen bezala erabiltzea. Horregatik, ikaslearentzako ingurunean ez dugu interesgarri ikusten datu-base lexikalaren informazio guztia erakustea, zenbait datuk ikaslea nahas dezake eta. Beraz, atzipenerako interfaze berria sortu dugu. Honetaz gain, hitzari buruzko informazioa eskatzean ez zaio ikasleari eskatzen lema jakitea, beraz, hitz deklinatua emanda ere informazioa jasotzeko aukera du. Batzuetan, hitz deklinatu batek lema posible bat baino gehiago izan dezake, kasu horretan sistemak ikasleari lema posibleen zerrenda aurkeztu eta galdetzen dio lema aukera dezan. Azkenik, ikaskuntza-maila bakoitzeko datu-base bana

sortzea proposatzen dugu. Honetarako, atributu berriak beharrezkoak dira, hots, ikaskuntza-maila eta hitzaren aplikazio-testuingurua (tematikoa, linguistikoa eta egiturazkoa).

- Hitzen definizioa kontsultatzeko hiztegia ere badugu gure ingurunean. Tresna hau egiteko oinarri bakarra hiztegia izan da. Kontsultarako tresna sortu dugu eta EDBLn gertatzen den bezala hitzaren bilatu aurretik lematizatu egiten da.

- Proposatzaileak desbideratzedun hitz bat emanda antzeko hitz estandarrak proposatzen dizkio ikasleari. Guk proposatutako egokitzapenean proposatzaileak kontuan izaten ditu bai ikaslearen ezaugarriak, bai saioaren helburuak.

- Analizatzaile morfologikoa (Alegria et al., 96) egoki deritzogu gramatikaren kontsultarako sistema gisa. Analizatzailea Bi Mailatako Morfologian (Agirre et al., 92) oinarritua dago eta hitzaren analisi posible guztiak ematen ditu, batzuetan sistemarentzat lema ezezaguna izan arren. Analizatzailean kontrola eta ezagutza zeharo independenteak direnez posible izan zaigu sistema egokitzea ikasleen sistemari begira. Aldaketak ezagutzan batez ere egin ditugu, eta aldaketa batzuk kontrolean ere beharrezkoak izan zaizkigu. Adibidez, hasierako analizatzaileak ez zituen analisisian aktibatzen ziren erregelak detektatzen, egokitutako analizatzaileak, ordea, analisi bakoitzean aktibatutako erregelak, estandarrak eta desbideratzeak, azaltzen ditu ikaslearentzat adierazgarri den kode baten bidez. Modu honetan ikasleak informazio gehiago du desbideratzeak egiten baditu berak zuzendu ahal izateko. Eta zuzentzeko gai ez bada, proposatzaileari kontsulta diezaioke beti.

- IDAZKIDE ingurunean sortutako azken tresna hitzen deklinabidearen sortzailea dugu. Honek hitz bat emanda bere deklinabide osoa, edo ikasleak eskatutako deklinabide-kasu konkretuak ematen ditu. Sortzailea analizatzaile morfologikoan oinarrituta dago. Hau posible da aipatutako analizatzailea bi zentzutan dabilelako, analisisian eta sorkuntzan: hau Bi Mailatako Morfologiaren (Koskeniemi, 83) ezaugarri nagusietako bat da.

VI.3.4 Ikasle-eredua.

VI.3.4.1 Ikasle-ereduak sistema adimendunetan.

Gaur egun, sistema adimendunen arloan dabilzan ikertzaile askok erabiltzailearen eredua beharrezkotzat hartzen dute. Dena den, ez dago akordiorik eredian kontsideratu beharreko ezaugarriei buruz aritzen denean. Eta batzuetan zalantzan jartzen da osagai beharrezkoa ote den ere (Holt et al., 94). Guk ikasle-eredua osagai garrantzitsutzat hartzen dugu IDAZKIDE bezalako sistema adimendun batean (Díaz et al., 98b).

Hori dela eta, ikasle-eredua bibliografian topatzen ditugun ereduaren sailkapenetan kokatu nahi izanez gero, kontuan izan behar dugu amankomunean duten ikasle-eredua desberdin sailka daitekeela kontuan izandako sistemaren arabera: HITES, IRAKAZI edo IDAZKIDE. Adibidez, ereduaren ezagutzaren eguneratzeari dagokionez bada hainbat ezaugarri IDAZKIDEn estatikoa dena, eta beste bi sistemetan dinamikoa. Hau dela eta, § VI.2.2-n *ezaugarria* objektua definitu dugu dinamismoa adierazi ahal izateko. Ondorioz, ikasle-ereduaz aritzen garenean, hiru sistemek konpartitzen duten ikasle-ereduaz mintzatzeko garen edo sistema konkretu baten ereduaz aztertzen ari garen zehaztea beharrezkoa da. Atal honetan, Verdejoren (1992) sailkapena abiapuntutzat hartuta eredu komunaz aipamenak egingo ditugu; hurrengo atalean, ordea, IDAZKIDEn ikasle-eredua soilik landuko dugu.

Verdejoren (1992) sailkapenean ereduak sei irizpide desberdinen arabera sailkatzen dira:

- Eredu generikoak versus banakoak.
- Ereduaren erabilera.
- Domeinu-ezagutzaren adierazpenaren zabalkuntza.
- Azaleko versus sakoneko ereduak.
- Eredu iraunkorrak versus tenporalak.
- Eredu aurredefinituak versus ondorioztatuak.

Eredu generikoak versus banakoak. Lan honetan proposatzen den MUGARRI ingurunearen ezagutza-baseen artean **tartehizkuntz ereduak** eta **ikasle-ereduak** agertzen zaizkigu. Tartehizkuntz ereduetan ikaskuntza-maila bereko ikasleen ezagutza komuna gordetzen da; ikasle-ereduetan, berriz, ikasle bakoitzari dagokion ezagutza propioa errepresentatzen da, § VI.2.2-n aipatutako ikaslearen ezaugarriak ere hemen biltzen direlarik. Horren arabera, tartehizkuntz ereduak eredu generikoak izango lirateke eta ikasle-ereduak eredu banakoak.

Ereduaren erabilera. Verdejok eredu deskriptiboak eta aurreikusleak desberdintzen ditu. Aurreikusleek erabiltzailearen portaera simulatzeko aukera ematen dute. Gure kasuan, ez tartehizkuntz ereduak ezta ikasle-ereduek ere ez dute ezaugarri hau, eredu deskriptiboak erabiltzen ditugu eta.

Domeinu-ezagutzaren adierazpenaren zabalkuntza. Aztertzen ari garen sailkapenean hiru eredu desberdin aipatzen dira: overlay (Carr & Goldstein, 77), diferentziala (Self, 92) eta grafo genetikoa. Ezagutza-basean aurkezten ditugun ereduak diferentzialak dira sistemaren portaera eta ikaslearen konparatuz sortuak baitira. Selfek (1992) diagnosi-prozesua eta ikaslearen modelizazioa sinonimotzat hartzen ditu. Testuinguru honetan hiru eredu-mota desberdintzen ditu eredu diferentzialen artean: eredu generikoak; eredu anitzak eta eredu generiko anitzak. Guk aurkezten dugun MUGARRI ingurunean ikasleen eredu anitzak definitzen ditugu (ikasuntza-maila desberdinen tartehizkuntz ereduak). Selfek (1992)

dioenez: *A novice's knowledge may be a conceptualization fundamentally different from the expert's. Hence the former cannot be represented by small perturbations of the latter.* Eredu anitzen artean ordenatuak eta ez-ordenatuak definitzen ditu. Lehenengoan ereduaren arteko progresioa errepresentatzen da, bigarrenean, berriz, ez dago progresio-erlaziorik. Guk definitutako tartehizkuntz ereduaren bi aspektuak gauzatzen dira.

Azaleko versus sakoneko ereduak. Bigarren hizkuntzaren ikaste-prozesua aztertzeke proposatzen dugun MUGARRI izeneko ingurunean argi dago sakoneko ereduak hartu ditugula oinarri, hau da, ikasleak egiten dituen desbideratzeen azalpena espreski adierazten dugu ezagutza-baseen artean diagnosi-ezagutza modulua gehituz.

Eredu iraunkorrak versus tenporalak. Konbinaketa desberdinak kontsideratu ditugu gure sistemetan. Adibidez, V. kapituluaren ikusi dugu HITESen tartehizkuntzaren eredu iraunkorra eta tenporalaren konbinaketa erabili dugula, hau da, saioaren bukaeran eredu tenporalaren (ezagutza dinamikoa deitu diogu han) egoera ikusita eredu iraunkorra eguneratzen genuen. Argi dago ezagutza-basean ditugun ereduak iraunkorrak direla. IDAZKIDEn ere bi eredu moten arteko konbinaketa erabili dugu ikaslearen buruzko informazioa adieraztean. IRAKAZIk, berriz, ez du eredu propiorik erabiltzen.

Eredu aurredefinituak versus ondorioztatuak. Gure kasuan, eredu berak, hots tartehizkuntz ereduak, aurredefinituak dira IDAZKIDE sistemarako, eta HITES eta IRAKAZI sistemarako berriz ondorioztatuak dira, sistemek berek sortzen dituzte eta. Ikasle-ereduen kasuan IRAKAZIren ikuspegitik ondorioztatuak dira, eta HITES zein IDAZKIDERentzat aldi berean aurredefinituak eta ondorioztatuak dira, sistemek ereduak eguneratu egiten dituztelako.

VI.3.4.2 Ikasle-eredua IDAZKIDEn.

Aurreko atalean aipatu dugunez, IDAZKIDEk bigarren hizkuntza ikasten ari den ikaslearen ereduak gordetzen du sistemaren erabakiak zehatzago bideratu ahal izateko. Dena den, IDAZKIDE ez da ikasle-eredu osoaren jabe, eta ikasle-eredua bi azpieredu nagusitan banatzen dugu:

- Ikasle-ereduaren gunea: hau IDAZKIDE, HITES eta IRAKAZI sistemek konpartitzen duten ezagutza-baseetako bat da (§ VI.2.2).
- Saioaren ikasle-eredua: ikaslearekiko elkarrekintzan zeresana duten ikaslearen ezaugarriak biltzen ditu.

Atal honetan saioaren ikasle-eredua landuko dugu gehienbat. Lehenengo, saioaren garrantzitsu den informazioa azalduko dugu, eta, ondoren, saioaren ikasle-eredua definituko dugu.

Eman dezagun saio batean ikasleak nahi adina testu eduki ditzakeela aldi berean irekita, eta, gainera, testu batetik bestera mugi daitekeela, testu desberdinen ezagutza aldi berean aztertuz. Honek aldi berean testuinguru desberdinak landu ahal izateko aukera ematen dio ikasleari, konparaketak egiteko ahalbidea izanik. Hau dela eta saioaren ikasle-ereduan gordetzen den aztertze-unitate bakoitzari dagokion testua asignatuko litzaioke.

Testuak aztertze hainbat tresna eskaintzen dio sistemak ikasleari. Honek, bai modu askean, bai sistemak ematen dizkion gomendioei jarraituz, ikaste-prozesu propioa bideratuko du.

Prozesu honen hainbat datu garrantzitsu bildu beharko da saioaren ikasle-ereduan:

- Ikasleak idatzitako testuak.
- Testuen ezagutza lantzeko erabilitako tresnak.
- Testuetan ez dagoen ezagutza lantzeko ikasleak egin dituen kontsultak.
- Sortutako edo aldatutako ikasle-ereduaren gunea.
- Ikasleak erantzundako galderak saioaren hasieran edota bukaeran, erantzun baditu.
- Sistemak emandako gomendioen aurrean ikasleak izandako jokabidea.

Saioaren bukaeran gainbegiratzailerak datu hauek landu eta ikasle-ereduaren gunea eguneratuko du. Honetaz gain, hurrengo saioetarako interesgarri izan daitekeen informazioa ere gordeko du saioaren ikasle-ereduan.

Saioaren ikasle-ereduan bi zati bereiztuko ditugu:

- Ikasle-eredu iraunkorra.
- Ikasle-eredu tenporala.

Ikasle-eredu iraunkorrak hurrengo saioetarako interesgarri izan daitekeen informazioa gordeko du: informazio hau saio bakoitzaren bukaeran eguneratuko du ikasle-eredu tenporalaren datuak erabiliz.

Ikasle-eredu iraunkorra.

Ikasle-eredu iraunkorra ikaste-prozesuak dirauen bitartean gordeko da, hau da, saioa bukatzean eredu honen informazioa gorde egingo da. Eredu iraunkorrean azken saioei buruzko informazioa gordetzen da, beti ere ikasle-ereduaren gunean ez bada informazioa eguneratu.

Ereduaren zati honetan bi ezagutza-mota gordetzen dira: hizkuntzari buruzko ezagutza eta ikaste-estrategiei buruzko ezagutza.

```
structure Saio_eredu_iraunkorra;  
  subtype-of:   Ikasle_ereduaren_kontzeptua;  
  parts:  
    ondorio_uitateak: set(instance(Unitate_iraunkorra));  
  
structure Unitate_iraunkorra;  
  subtype-of:   Ikasle_ereduaren_kontzeptua;  
  properties:  
    erabilera_puntuaketa: float;  
  
structure Ezagutza_iraunkorra;  
  subtype-of:   Unitate_iraunkorra;  
  parts:  
    balioa: instance(Tartehizkuntz_egitura);  
  properties:  
    mota: symbol; value={lema, erregela}  
  
structure Estrategia_iraunkorra;  
  subtype-of:   Unitate_iraunkorra;  
  parts:  
    balioa: instance(Ikaslearen_ikaste_estrategia);
```

VI.6 irudia. Saioaren ikasle-eredu iraunkorraren definizioa CLIPSen.

Erabilera_puntuaketa slotak (VI.6 irudia) saioan zehar ezagutza linguistikoa edo ikaste-estrategiak agertutako aldi-kopuruaren batez-bestekoa gordetzen du. Ezagutza linguistikoaren kasuan gainbegiratzailerak balio hau eta ikasle-ereduaren gunean ezagutza berak duen estabilizazio balioaren arabera ikasle-ereduaren guneko egituraren estabilizazioa eguneratuko du. Ikaste-estrategien kasuan ere *erabilera_puntuaketa* balioa kontuan izanik ikaste-estrategien *lehentasuna* balioa eguneratuko du ikasle-ereduaren gunean. Hurrengo atalean, ikaste-estrategiak lantzean, ikusiko dugu *lehentasuna* balioak ikasleak ikaste-estrategia erabiltzeko nahia neurtzen duela.

Ikasle-eredu tenporala.

Ikasle-eredu tenporalak lansai baten bizitza du. Saioa bukatzean bere informazioa ikasle-eredu iraunkorra edo ikasle-ereduaren gunea eguneratzeko erabiliko da; ondoren, duen informazioa galdu egingo da.

```

structure Saio_eredu_tenporala;
subtype-of:    Ikasle_ereduaren_kontzeptua;
parts:
    aztertze_unitateak: set(instance(Unitate_tenporala));

structure Unitate_tenporala;
subtype-of:    Ikasle_ereduaren_kontzeptua;
properties:
    identifikadorea: string;
    mota: symbol; value={hitza, testua}

structure Testua_unitatea;
subtype-of:    Unitate_tenporala;
parts:
    desbideratzeak: set(instance(Desbideratzea));
    hitz_aztertuak: set(instance(Hitza_unitatea));
properties:
    identifikadorea: string;

structure Hitza_unitatea;
subtype-of:    Unitate_tenporala;
parts:
    tresnak: set(instance(Tresna));
properties:
    balioa: string;
    testuak: list(string-value);
    agerpenak_saioan: integer;

structure Tresna;
subtype-of:    IDAZKIDE_kontzeptua;
parts:
    ikaste_estrategiak: set(instance(Ikaste_estrategia));
properties:
    identifikadorea: string;
    mota: symbol; value={hiztegia, morfosint}

structure Desbideratzea;
subtype-of:    Tartehizkuntz_egitura;
properties:
    mota: symbol; value={lema, erregela}

```

VI.7 irudia. Saioaren ikasle-eredu tenporalaren definizioa.

Hitza_unitatea objektuaren (ikus VI.7 irudia) *tresnak* izeneko slotean *balioa* slotean gordetzen den hitza aztertzean ikasleak erabilitako tresnak adierazten dira. Tresna horiek esango digute ikaslea lema aztertzen ari den (adib. EDBL edo definizio-hiztegia kontsultatzen badu) edo hitz horretan agertzen den fenomeno morfosintaktikoren bat lantzen ari den. Honetaz gain, tresnak erabilitako ikaste-estrategien isla dira.

Aurretik esan dugunez, aztertutako hitzak ikasleak idatzitako testuetan ez du zertan agertu behar izan, hau da, ikasleari ikaste-prozesuan zehar besterik gabe kontsultatutako hitza izan daiteke. Kasu honetan *testuak* izeneko slota hutsa egongo da.

Testua _unitatea objektuan ikusten denez *desbideratzeak* izeneko slota dugu. Honetan testua aztertzean detektatutako desbideratzeak gordeko dira. Desbideratzeren bat ikasleak aztertuz gero *Hitza _unitatea* sortzen da desbideratze horretarako.

VI.3.5 Ikaste-estrategiak.

VI.3.5.1 Ikaste-estrategiak sistema informatikoetan.

Learning strategy is steps or actions taken by learners to improve the development of their language skills.

Hau da Oxford & Cohen-ek (1992:1) ikaste-estrategia definitzeko modua. Ikaste-estrategiak Selinker & Gassek (1994:265) dioten bezala buru barruko ekintzak zein ekintza fisikoak izan daitezke. Literaturan zehar ikaste-estrategien sailkapen asko topatzen ditugu. O'Malley & Chamotek (1990) ikaste-estrategien sailkapen baliagarriak laburbildu zituzten. Hauek egindako sailkapenean ikaste-estrategiak hiru azpimultzotan banatzen dituzte, estrategia metakognitiboak, estrategia kognitiboak eta estrategia sozialak. Sailkapen hau gure sistemaren oinarria izan da, eta egitura linguistikoak erabiltzearen sakoneko arrazoen sailkapena egitean ere O'Malley & Chamoten (1990) sailkapena erabili dugu (ikus C eranskina).

Tesi honetan proposatzen dugun MUGARRI ingurunearen azken helburua ikaslearen gaitasun gramatikala eta honen bereganatze-prozesuaren azterketa egitea da. Gaitasun gramatikala edo tartehizkuntza bereganatzean faktore askok hartzen dute parte, faktore psikolinguistikoak eta soziolinguistikoak. Gure sistema aspektu psikolinguistikoetara (estrategia metakognitiboak eta kognitiboak) murriztu dugu, eta, zehazki, ikasleak bere tartehizkuntzaren gaitasun gramatikala osatzeko erabiltzen dituen ikaste-estrategietara. Faktore soziolinguistikoak (estrategia sozialak) ez ditugu tratatu sisteman.

Eta zergatik lantzen ditugu ikaste-estrategiak, eta ez tartehizkuntzaren osaketa-prozesuan eragina duten komunikazio-estrategiak? Cookek (1993) dio:

In a learning strategy the learner attempts to bring **long-term competence** into being, in a communication strategy, to solve a momentary communication difficulty.

Beraz, gure sistemaren azken helburua ikaslearen gaitasun gramatikala eta honen bereganatze-prozesuaren azterketa egitea denez, gure lana desbidera zezaketen komunikazio-estrategiak oraingoz baztertu egin ditugu eta ikaste-estrategietara murriztu dugu ikerketa. Dena den, honek ez du esan nahi komunikazio-estrategien ondorioz sortutako egiturak detektatzen saiatu ez garenik. Gaitasun komunikatiboa tartehizkuntzaren zati bat ere bada. Hortaz, sistemak tartehizkuntza bere osotasunean aztertuko duenean, bai gaitasun

gramatikala, bai komunikatiboa landuko du. Azken honetarako gaur egun multimedia arloak eskaintzen dizkigun aukerak ekarpena izan litezke (Maritxalar & Díaz, 95). Hala eta guztiz ere, oraingoz behintzat, gaitasun gramatikalean arreta jarri eta ikaste-estrategiak izan ditugu aztergai gehienbat.

Diseinatu dugun ingurunea irekia da ikaste-estrategiak era inkrementalean lantzeko aukera ematen duelako. Baina zein irizpideren arabera hasi gara ikaste-estrategiak aztertzen? Zeintzuk dira ikaste-estrategiak eta ikaslearen tartehizkuntzaren arteko erlazioa lantzeko jarraitu ditugun pausuak?

- 1) O'Malley & Chamoten (1990) ikaste-estrategien sailkapena (*Metacognitive Strategies, Cognitive Strategies eta Social Mediation Strategies*) oinarritzat hartu dugu. Sailkapen hau atal honetan landuko dugu.
- 2) Desbideratzeen azpian dauden sakoneko arrazoiekin lotuta dauden ikaste-estrategiak bereziki aztertu genituen esperimenduetan. Ondoren, sakoneko arrazoiak eta ikaste-estrategien arteko erlazioei buruz hipotesiak formulatu genituen (ikus C eranskina).
- 3) Ikasle konkretu batzuekin aurreko fasean egindako hipotesiak aztertu genituen.

Atal honetan O'Malley eta Chamotek aurkezten dituzten ikaste-estrategiak komentatuko ditugu, sistema informatikoei begira. Ikuspegi honetatik ondo desberdinu beharko ditugu zeintzuk diren sistema adimendunek detektatzen lagun dezaketenak, adibidez, beste hizkuntza baten egitura linguistikoen transferentzia (*transfer*); eta zeintzuk diren sistemak inola ere detekta ez ditzakeenak, hala nola, entzutezko errepresentazioa (*auditory representation*). Azken honetan ikaslea saiatzten da soinua bere buruan gordetzen eta hau sistema batek detekta dezakeenetik urruti dago. Hortaz, aipatutako sailkapenean ikaste-estrategia bakoitzerako tratamendu informatiko lagungarria posiblea den ala ez aztertuko dugu. § VI.3.5.2-n, berriz, IDAZKIDEk eskaintzen dituen aukerak landuko ditugu.

Esan bezala, aipatutako sailkapenean ikaste-estrategiak hiru multzo nagusitan banatzen dituzte:

- Estrategia metakognitiboak.
- Estrategia kognitiboak.
- Estrategia sozial afektiboak.

Gure ustez, sailkapen honetan agertzen diren ikaste-estrategia asko tratagarriak dira adimen artifizialeko edo multimediako tekniken bidez.

Estrategia metakognitiboak

Estrategia metakognitiboen bidez ikasleak bere ikaste-prozesua planifikatu, gainbegiratu eta ebaluatu egiten du. Beraz, estrategia metakognitiboak ikasteari buruzko estrategiak dira,

ikaste-estrategiak baino. Ikus ditzagun estrategia metakognitiboak zeintzuk diren eta nola bidera daitezkeen sistema informatiko bati begira. Lehenengo sistema informatiko batean bideragarrienak direnak aipatuko ditugu:

- Antolamendu aurreratua (*advanced organisers*):
Sistemak aukera eman diezaioke ikasleari aurreikusteko zenbait gai ikaste-saioa hasi baino lehen (hiztegiak, gramatikak, e.a.), berrikusteko aurreko saioetan egindako ariketak, kontsultatzeko entziklopedia elektronikoak, e.a.
- Atentzio selektiboa (*selective attention*):
Ikasleak burutuko duen saioan hizkuntzaren zein arlo konkretu aztertu nahi dituen aukera ditzake (lexikoa, morfologia, sintaxia, ...). Honen arabera sistemaren modulu batzuk aktibatu eta beste batzuk desaktibatuko lirateke.
- Planifikazio funtzionala (*functional planning*):
Ikasleak ataza linguistikoari ekin baino lehen, ataza burutzeko behar dituen ezagutza linguistikoak (hiztegiak, gramatikak, ...) eska diezazkioke sistemari.
- Autoebaluazioa (*self-evaluation*):
Zalantzen kasuan sistemak zuzentzaile automatikoak (ortografikoak, sintaktikoak, e.a.) erabiltzeko aukera eman diezaioke ikasleari.

Ondoren aurkezten diren estrategia metakognitiboak tratatzea, ordea, zailago gertatzen da, ez bazaizkio ikasleari galdera oso zuzenak egiten; eta, hala ere, teknika hau ez da zeharo fidagarria, ikasleak askotan aplikatzen dituen ikaste-estrategien kontzientziarik ez baitu.

- Autoorganizazioa (*self-management*):
Ikaste-ataza baldintza egokienetan eman dadin ikasleak ikasgela egokia, une egokia, e.a. aukeratu du.
- Automonitorizazioa (*self-monitoring*):
Ikaslea ikaste-ataza egiten ari den une berean bere buruan ataza ondo burutzen ari den ala ez aztertzen ari da.

Estrategia kognitiboak

Estrategia kognitiboek ikaskuntza hobetzeko ezagutzaren bereganatzeaz eta manipulazio zuzenaz arduratzen dira. Estrategia kognitibo asko azter daitezke sistema informatiko baten bidez, beti ere, kontuan izanda makinak markatzen dizkigun mugak. Badaude estrategia batzuk, ordea, tratagaitzak edo trataezinak direnak; ikus ditzagun hauetako batzuk. Adibidez, entzutezko errepresentazioa (*auditory representation*) ezin da inola ere tratatu sistema informatiko baten bidez. Ikaslea bere buruan soinua gordetzen saiatzen ari den ala ez detektatzea ezinezkoa zaigu. Dedukzioa (*deduction*) ere detektagaitza agertzen zaigu. H2-ren erregelak *kontzienteki* aplikatzen ari den jakitea oso zaila zaigu. Hitz gakoaren teknika (*key*

word) ere, beste aldetik, zaila gertatzen da harrapatzea. Ikasleak memorizatzeko erabil dezakeen teknika denez, ezinezkoa da sistema informatikoak ikaslearen buruan gertatzen ari dena jakitea. Arrazoi beragatik, inferentzia (*inferencing*) tratagaitza dela esango dugu. Edozein kasutan, hitz gakoak, inferentzia eta honelako ikaste-estrategiak ariketa-moten bidez implementa daitezke, gero aukera emateko ikasleari hauek erabil ditzan nahi duenean. Ikasleak egindako ariketen aukeraketaren arabera estrategien erabilpena azter daiteke neurri batean.

Beste estrategia batzuk, ordea, neurri handiago edo txikiago batean tratagarri gertatzen dira sistema informatiko batean. Ikus ditzagun, beraz, hauek tratatzeko lehenengo ideia batzuk.

- Errepikapena (*repetition*):
Detektatu beharko genuke ea ikaslea saiatzen den pertsona batek dioena errepikatzen. Honetarako, multimedia arloan aukera politak agertzen zaizkigu. Bideo bat ikusi ondoren, ikasleak bideoa berrius dezake berak pelikulako pertsonaia baten papera hartuz. Ikasleak dioena graba daiteke eta gero aditu.
- Baliabideen erabilpena (*resourcing*):
Hizkuntzarako baliabideak erabiltzeko aukera eduki dezake sistemak (hiztegi eta entziklopedia elektronikoak, e.a.).
- Itzulpena (*translation*):
Linguistika konputazionalako tekniken bidez ikasleak bere ama-hizkuntzaren egitura linguistikoak erabiltzen dituen detekta daiteke neurri handi batean, maila morfologiko zein sintaktikoan.
- Taldekatzea (*grouping*):
Sistemak aukera eman diezaioke ikasleari zenbait gai aztertzeke atributu komun batzuen arabera (adib. deklinabidea aztertzean, alde batetik, bokalez eta, beste aldetik, kontsonantez bukatzen diren lehen deklinabidea ikasi).
- Oharrak hartzea (*note-taking*):
Edozein unetan oharrak hartzeko aukera emateak ez du inolako zailtasunik alde informatikotik.
- Irudiak erabiltzea (*imagery*):
Multimedia bidez tratatzeko estrategia aproposa dugu hau. Ikasleak, hiztegian begiratu beharrean, aukera izan dezake ikusteko hitz bati dagokion argazkia, edo aditz baten definizioari dagokion bideozko ekintza-sekuentzia.
- Kontestualizazioa (*contextualisation*):
Ikaste-teknika hau erabiltzeko aukera eman diezaioke sistemak multimedia-sistemen bidez. Egitura linguistikoak ariketa linguistikoen bidez ikasi beharrean, egitura horiek aztertzen dituzten elkarriketak, dokumentalak, e.a. ikusteko aukera eman diezaiokegu ikasleari.

- *Elaborazioa (elaboration):*
Informazio berria ikasleak memorian aurretik dituen kontzeptuekin erlazionatzen saiatzen den detekta daiteke zenbait kasutan (adib. saiatu eratorpenaren erregelak aplikatzen hitz berriak sortzeko).
- *Transferentzia (transfer):*
Itzulpenean esan dugun bezala, zenbait egitura transferitu den ala ez detekta daiteke, beti ere, kontuan hartuz transferitutako hizkuntzaren errepresentazioa sisteman integratu beharko litzatekeela. Hiru transferentzia-mota desberdintzen ditugu gure sailkapenean:
 - a) Ama-hizkuntza ez den hizkuntza baten transferentzia.
 - b) Bigarren hizkuntzaren beraren euskalkien eraginez sortutako transferentziak.
 - c) Tartehizkuntzaren beraren barruan sortutako transferentziak (*Internal Interlanguage Transfer*).
- *Laburpenak (summarization):*
Jasotako azken informazioaren laburpena egiteko lan-tresnak eskain diezazkioke sistemak (adib. ikasleak idatzitako testutik aditzen zerrenda eman eta ikasleak sailka ditzan eskatu). Bestela, sistemak ikasleari ikasitakoaren laburpenak egiteko leiho berezia eskaini dakioko.

Ondoren aipatuko dugu, gure ustez, estrategia kognitiboen artean sartu beharko genukeen estrategia, (O'Malley & Chamotek, 90)-n estrategia sozial afektiboen artean sailkatu eta (O'Malley et al., 85)-n, berriz, estrategia kognitiboen artean kokatu zutena.

- *Argitzeko galdeketa (question for clarification):*
Sisteman askatasuna eman diezaiokegu ikasleari nahi duenean kontsultak egin ditzan.

Estrategia sozial afektiboak

Ikaste-estrategia hauetan beste pertsonetikiko elkarrekintzak eta aspektu afektiboak sartzen dira. Gure sistemak arlo hauek tratatuko ez lituzkeenez, ez ditugu hemen aztertuko. Beste pertsonetikiko elkarrekintzei dagokienez, informatikaren munduan sare bidezko sistema informatikoetan azter litezke. Baina argi utzi nahi dugu garatu dugun sistema adimendunak, oraingoz behintzat, ikasle bakar batekin lan egingo duela une bakoitzean.

VI.3.5.2 Ikaste-estrategiak IDAZKIDEn.

Ikaste-sistemetan batzuetan diseinatzaileak landutako erabilerak ikasle gutxik erabiltzen ditu, eta diseinatzaileak garrantzirik eman gabe eskaintako aukerak, ordea, ikasle askok erabiltzen dituzte. Honetaz gain, ikaste-sistemak ikasleak ikaste-estrategia konkretuak gara

ditzan prestaturik egon arren, zaila egiten da ikaste-estrategien erabileraren detekzioa. Ikaste-estrategien detekzioa eta garapena lantzen duen Mr. Collins sistema (Bull, 97)-n aurkitzen dugu. Sistema honek bigarren hizkuntza ikasten ari den ikaslearen ikaste-estrategien azterketa sakona egiten du, ikasleari bere ikaste-estrategien erabilpenaren azalpen esplizitua azalduz. Dena den, guk uste, ikaste-estrategien azalpen esplizituak ikasle batzuk besterik ez dute nahi izaten; honegatik, IDAZKIDEn ikaste-estrategien erabilpena, Mr. Collins-en bezain landua ez izan arren, ikaslearentzat atseginagoa da ezkututzen baitzaio. Beste alde batetik, Mr. Collins sistemak ezagutza eta galdera aurredefinituak aurkezten dizkio ikasleari, eta, gure ustetan, ikaslearen output librea beharrezkoa da ikasleak nahi dituen ikaste-estrategia guztiak gara ditzan. IDAZKIDEk ikasleari output librea adierazteko aukera ematen dio.

IDAZKIDEk, inguruneak ikasleari eskaintzen dizkion tresnak direla medio, ikaste-estrategien erabilpena aztertzen du. Honetaz gain, saioaren hasieran eta bukaeran galdera batzuk egiten zaizkio ikasleari, berak nahi izanez gero. Galdera hauek saio konkretuan erabilitako hainbat estrategia argitzen dute. Kontuan izan behar dugu, ikasle-ereduaren gunean ikaslearen ikaste-estrategiak gordetzen diren arren, askotan une desberdinetan estrategia berriak aukeratzen dituztela ikasleek, ikaste-prozesuan zehar konturatzen baitira estrategien aldaketa lagungarria dela. Beraz, ikasle-ereduaren gunean ikaslearen ikaste-estrategia orokorrak gordeko ditugu eta saioaren ikasle-ereduan saioan zehar erabilitako estrategiak landuko ditugu: hauei buruzko informazioa erabilitako tresnek emango digute. Honela, saioaren bukaeran gainbegiratzailerak ikasle-ereduaren gunea eguneratuko du saioaren eredia eta aurreko saioetako ereduak kontuan izanik.

Oraingoz, ikasle-ereduaren definizioa aztertzean ez dugu ikaste-estrategiaren definizioa eman. Ondoren, ikaste-estrategia objektuaren definizioa azaldu eta ikasle-ereduan kokatuko dugu. Horretaz gain, IDAZKIDEk detektatzen dituen ikaste-estrategiak komentatuko ditugu.

Ikaste-estrategia objektuaren definizioa

§ VI.2.2-n ikasle-ereduaren gunearen deskribapena azaltzean *ezaugarriak* izeneko multislota agertu zaigu. Slot honen balio posibleen artean ikaste-estrategiak topatzen ditugu, besteak beste. Ondoren ikaste-estrategia objektuaren definizioa emango dugu.

Lehentasuna balioak (VI.8 irudia) ikaste-estrategiari dagokionez, ikaslearen nahia adierazten du. Balio hau ikaslearen eredia sortzean hasieratuko da. Ondoren, IDAZKIDEk berak eguneratuko du ikasleak erabilitako tresnen arabera.

```
concept Ikaste_estrategia;  
  subtype-of: ikasle_ereduaren_kontzeptua;  
  properties:  
    identifikadorea: string;  
    deskribapena: string;  
    mota: symbol; value={kognitiboa, metakognitiboa}  
  
concept Ikaslearen_ikaste_estrategia;  
  subtype-of: Ikaste_estrategia;  
  properties:  
    lehentasuna: float;
```

VI.8 irudia. Ikaste-estrategiaren definizioa CLIPSen.

Ikaste-estrategiak IDAZKIDEn

Esan dugun bezala, tresnak ikaste-estrategien isla izango dira IDAZKIDEn. Hortaz, tresnek ikasleari eskaintzen dizkioten aukerak oinarritzat hartuz ikaste-estrategia detektagarriak definituko ditugu. Honetaz gain, Mr. Collins (Bull, 97) sistemarekiko konparaketa egingo dugu. Mr. Collins sistemaren helburua da ikaslea konturaraztea bere ezagutzaz eta bere ikaste-estrategiez. Horretarako, hainbat ikasleren elkarrekintza eskaintzen du. IDAZKIDEn, berriz, ikaste-estrategiak prozesuaren bideratzailatzat hartzen dira, eta ikasle bakarrarekin lan egiten da.

Estrategia metakognitiboak

Antolamendu aurreratua eta planifikazio funtzionala erabiltzeko aukera ematen zaio ikasleari IDAZKIDEn. Saio baten hasieran, testuen tratamenduarekin hasi aurretik, ikasleari zuzenean gogorarazten zaio tresnak erabil ditzakeela ezagutza kontsultatzeko edota gainbegiratzeko. Estrategia hauek posible dira tratatzen ari garen domeinuaren errepresentazioa eta tratamendua zabala baita. Mr. Collins sisteman, ordea, estrategia hauek ez dira lantzen, domeinu murrizta dela eta.

Era berean, saioaren bukaeran sistemak ikasleari *autoebaluazioa* egiteko aukera eskaintzen dio. Saioan zehar ikasleak landutako hitzak eta sistemak detektatutako desbideratzeak aurkezten zaizkio ikasleari, zalantzaren bat sortzen dizkioten hitzak markatzeko eskatuz. Ezagutzaren autoebaluazioa egiteko aukera ematen zaio ikasleari, ez, ordea, ikaste-estrategien autoebaluazioa egiteko aukera Mr. Collins sisteman gertatzen den moduan.

Automonitorizazioa bi sistemetan ematen den estrategia metakognitiboa dugu. Bietan ikasleak bere output-a analizatu edo zuzendu dezake testua idazten duen heinean, sistemak ezer gomendatu aurretik.

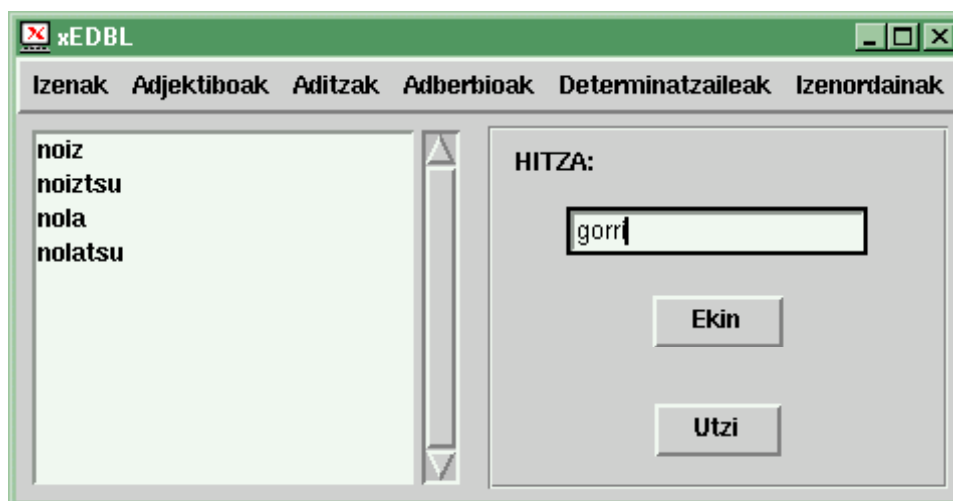
Atentzio selektiboa erabiltzeko aukera emanez gero, ikasleek sistemaren hainbat aukera desaktiba ditzakete. Modu honetan, ikasleen prozesua aztertuta sistema ere azter genezake. Oraingoz, gure helburua apalagoa denez, aukera hau ez dugu gehitu, implementatzea nahiko sinplea izan arren.

O' Malleyren sailkapenetik, *autoorganizazioa* gelditzen zaigu. Ikasle-ataza burutzeko baldintza egokiak aukeratzea ez dago gure baliabideen artean. Estrategia hau ez da tratatzen ez IDAZKIDEn, ez Mr. Collinsen.

Estrategia kognitiboak

Aurreko atalean aipatu dugunez, inferentzia, dedukzioa, elaborazioa eta hitz gakoien tekniken azterketa baztertu dugu. Dena den, horrek ez du esan nahi IDAZKIDEk honelako estrategien erabilpena ukatzen dionik ikasleari. Esate baterako, euskararako datu-base lexikalean ikasleak aditzen erabileren adibideak kontsultatzean gerta liteke inferentzi estrategia lantzen aritzea. Edo hitzen analisi morfologikoan aplikatutako erregelei buruzko informazioa eskatzea, dedukzioa aplikatu ahal izateko. Modu honetan justifikatzen da inferentzia eta dedukzioa estrategien erabilpena Mr. Collinsen. Gure ustez, estrategia hauek hori baino konplexuagoak dira, eta azterketa sakonagoa beharko lukete.

Baliabideen erabilpena, taldekatzea, oharrak hartzea, laburpena eta argitzeko galdeketa direlako estrategia kognitiboak IDAZKIDEn eta Mr. Collinsen eskaintzen zaizkio ikasleari; eskaintzeaz gain, detektagarriak ere badira. IDAZKIDEn taldekatzea ez da informazioa taldekatzeko leiho berezia antolatzea, oharrak hartzeko leihoan ikasleak informazioa ere taldeka dezakeen arren. Sistemaren tresnek taldekatze-aukerak eskaintzen dizkiote ikasleari, adibidez, EDBL kontsultatzean hitzaren kategoriaren arabera informazioa eska dezake ikasleak, VI.9 irudian ikusten den bezala.



VI.9 irudia. Taldekatze-estrategia IDAZKIDEn.

Azkenik, itzulpena zein transferentzia tratatzen dira bi sistemetan. Mr. Collinsen itzulpena hitzetik hitzera murrizten da; IDAZKIDEn, honetaz gain erregela-mailako itzulpenak tratatzen dira (adib. LEOEOU, LEAC... erregelak. - ikus III. kapitulua -). Tranferentziari dagokionez, Mr. Collinsen ama-hizkuntza ez den hizkuntzen transferentzia konkretuak lantzen dira (hizkuntz domeinu murrizt batean); IDAZKIDEn, berriz, bigarren hizkuntzaren beraren euskalkien eraginez sortutako transferentziak detektatzen dira, erregela orokorren bidez.

Aipatu eta tratatu ez ditugun estrategia kognitiboak ere badira: *errepikapena*, *berreraiketa*, *irudiak erabiltzea eta kontestualizazioa*. Estrategia hauek inplementatzeko ideia batzuk aurreko atalean komentatu ditugu. Hauek Mr. Collinsen ere ez dira lantzen.

VI.3.6 IDAZKIDEn interfazea.

Atal honetan IDAZKIDEn interfazeak eskaintzen dituen aukerak azalduko ditugu. Honetaz gain, inguruneak ikasleari eskaintzen dizkion tresnetako bat komentatuko dugu adibideren bat erabiliz. Atal honetan ikusten da nola aldeztatik aurretik egindako ikaste-estrategien azterketak interfazearen diseinua bideratu duen.

Bi plataforma desberdin eskaintzen ditu inguruneak tresnekin lan egin ahal izateko. Batean ikasleak hitz konkretu bati buruzko aztertze-lana burutu dezake, eta bestean testu oso batekin lan egiteko posibilitatea proposatzen zaio. Bi kasuetan ikasketa gidatua (ikasle-ereduan oinarritua), ikasketa librea eta bien konbinaketa erabiltzen da. Ikasketa gidatuan zein konbinatuan ikasle-ereduaren eguneratzea suertatzen da hainbat kasutan.

Aurretik azaldu den bezala, ikasle-ereduan hizkuntz ezagutzaz aparte ikaslearen ikaste-estrategiak eta saioetan jarraitutako prozesuak gordetzen dira. Prozesua diogunean ikasleak egindako desbideratzeak eta erabilitako tresnak adierazi nahi ditugu. Hori dela eta, sistemak ikaslearekin duen elkarrekintza egokituko du aipatutako informazioaren arabera, beraz, interfaze adaptatiboa (Downton, 91) du IDAZKIDEk.

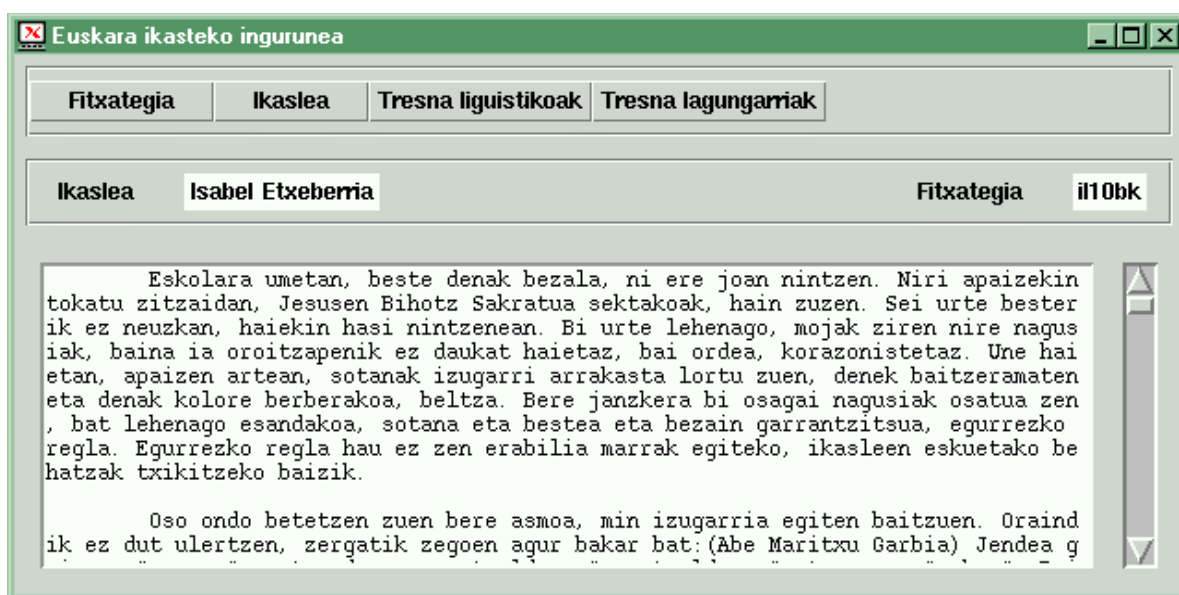
Sistema bi helburu desberdinekin erabil dezake ikasleak:

- Hainbat zalantza argitzeko.
- Idatzitako testuen zuzenketa bermatzen laguntzeko.

IDAZKIDEn interfazea Motif estiloko interfaze grafikoa da. Hau X Window aplikazioetarako sistema estandarretako bat denez, erraz suertatzen zaio ikasleari elkarrekintzaren ikasketa. Gainera, sisteman egindako diseinuan ikasleak edozein unetan erabil dezake edozein tresna leiho nagusitik pasa gabe.

Saio baten hasieran sistemak ikasleari kodea eskatzen dio eta laguntza gidatua nahi duen ere galdetzen dio. Ikasleak baietza emanaz gero ikasle eredu sortzeko aukera azaltzen dio, edo eguneratzeko aukera, ikaslea aurretik noizbait sistemarekin lan egina bazen era gidatuan. Honetaz gain, saioa hasi aurretik, sistemak ikasleari gogorazten dio hiztegiak eta gramatika erreparatzeko aukera duela zuzenean leiho nagusian agertzen diren tresna linguistikoak erabiliz. Ondoren, ikaslea testuen lanketan hasiko da. Saio bukaeran sistemak galdetzen dio ea idazte-prozesuan zehar gramatika kontzienteki aztertu duen edo esanahian gehienbat arreta jarri duen. Informazio honen bidez ikasleak jarraitutako prozesua aztertzei aukera aztertuko du sistemak etorkizunean. Edozein kasutan, une honetan saioaren bukaeran sistemak ikasleari autoebaluazioa egiteko aukera ematen dio, saioari buruzko datu batzuk aurkeztuz; beti ere ikasleak nahiko balu.

Ondorengo irudian sistemaren leiho nagusia azaltzen da.



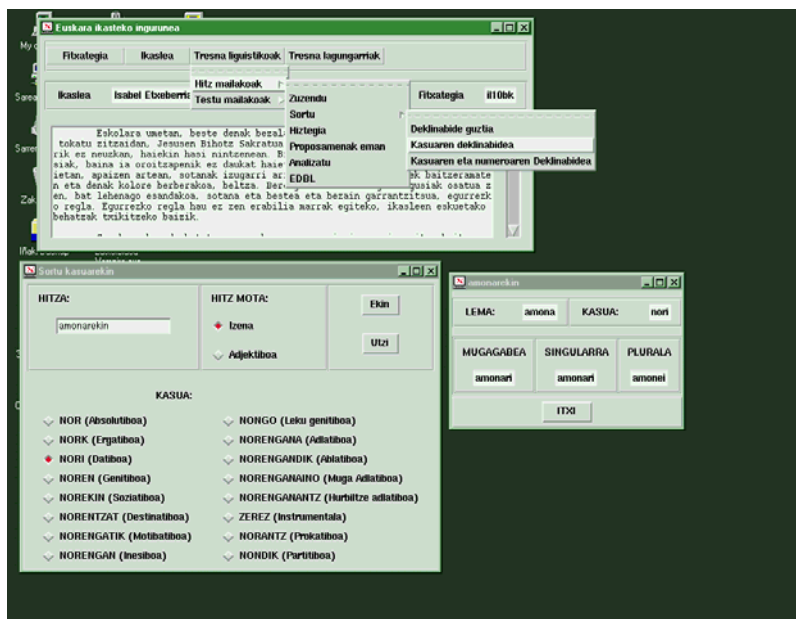
VI.10 irudia. IDAZKIDERen leiho nagusia.

Aurreko leihoan ikusten denez, ikasleak testu batekin, edo batzuekin, lan egiten duenean testua aldeztu aurretik idatzia egon daiteke, edo unean bertan idazten ari. Editore baten aukera tipikoak zein fitxategien maneiuak erabil ditzake ikasleak. Behin testua aurrez aurre izanik, sistemak idazten lagunduko dio eskaintzen dizkion tresnak direla medio. VI.10 irudian agertzen denez, bi motatako tresnak erabil ditzake ikasleak: tresna linguistikoak eta tresna lagungarriak. Tresna linguistikoak § VI.3.3-n azaldu ditugunak dira. Tresna lagungarriek, berriz, ez dute informazio linguistikorik erabiltzen. Oraingoz kontuan izan ditugunak ondokoak dira: oharrak hartzeko aukera, ikasleak berak bere laburpenak egin ditzan sortutako leihotxoak, eta saioaren bukaeran autoebaluazioa egiteko tresna. Azken honetan sistemak aztertutako zein detektatutako desbideratzeak aurkeztu dizkio ikasleari autoebaluazioa egiteko.

Irudian agertzen denez, ikaslea izeneko leihotxoak ere badugu leihorearen goiko aldean. Honen bidez, ikasleak bere ezaugarriak buruzko informazioa emango dio sistemari. Saio hasieran sistemak gordeta duen ikasle-eredua hartuko du abiapuntu hala izanez gero. Ikaslearen esku geldituko da ikasle-eredua sortzea ala ez sortzea. Ikasleak eredurik ez erabiltzea aukeratzen badu ikasketa librea aktibatzen da. Honegatik, sistemak saioaren hasieran informatuko dio zein den eredua izatearen ondorioa, hau da, laguntza gidaturik ez izatea.

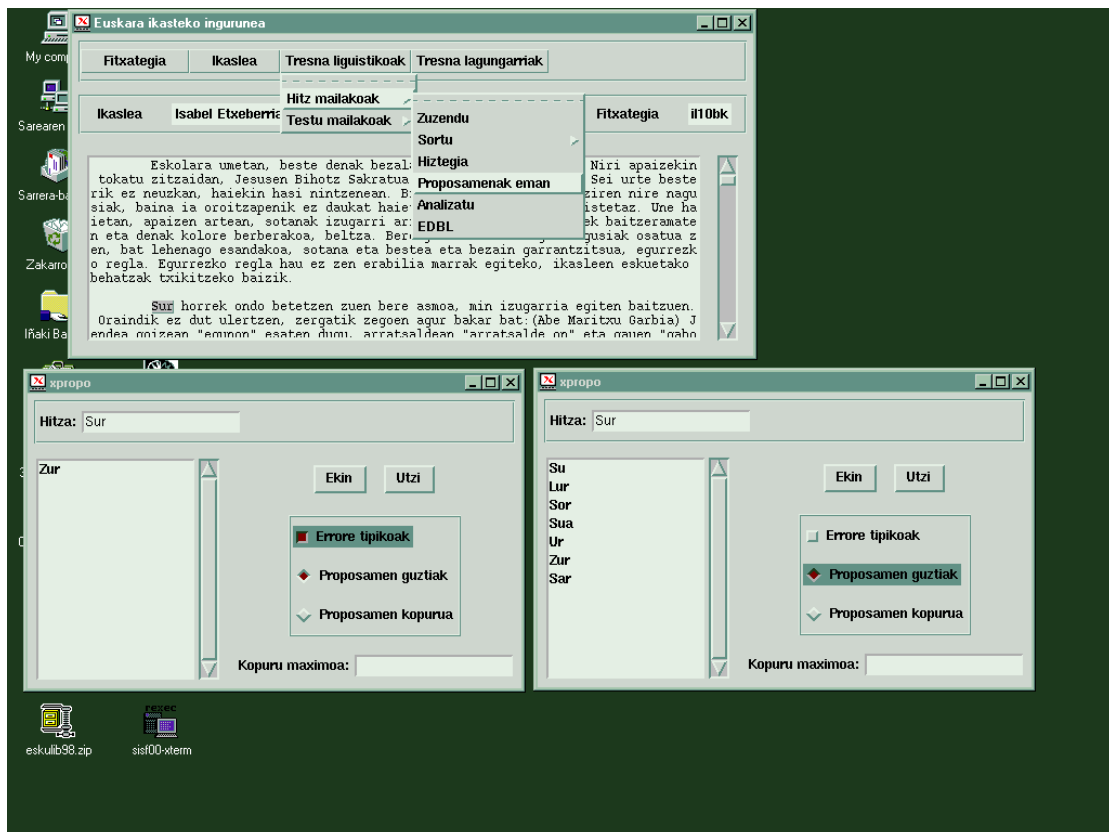
Erabilera-adibideak

Tresnen artean ikasleak hitzen sortzailea eta proposatzailea aukera ditzake, besteak beste.

VI.11 irudia. *amona* hitzaren deklinabidea.

Ikasleak sortzailea aukeratzen duenean hainbat aukera izaten du (hitz baten deklinazio osoa, kasu baten deklinazioa, edo kasua eta zenbakia emanda hitz deklinatua). VI.11 irudian ikusten den adibidean, adibidez, *amonarekin* hitza emanda ikasleak bere deklinazioa jakin nahi du datibo kasuan. Sistemak deklinazioa sortu egiten du, hau da, ez du aldezturik hitz deklinatua gordea. Beraz, sistemak hitzaren lema behar du, eta ikasleari zuzenean galdetu behar du sistemak berak *amonarekin* hitza lematizatuko du. Lematizatu ondoren sistemak lema datibo kasuan deklinatuko du irudian ikusten den moduan. Dena den, lematizazio-prozesuan lema topatzeaz gain lema aukeratu beharra izaten da askotan. Adibidean bi lema desberdin agertzen zaizkigu: *ama* eta *amona*. Aukera egiteko heuristikoen beharra dugu. Oraingoz, lema luzeana aukeratzen du sistemak. Hala ere, sakonago aztertu beharreko heuristikoa da hau.

Sortzailea tresna interesgarria da hitz estandarren kasurako. Ikasleak desbideratzea erabiliz gero, zalantzen kasuan egokia izaten zaio interfazeak eskaintzen dion proposatzailea.



VI.12 irudia. *sur desbideratzerako proposamenak.

VI.12 irudian ikasleak testuan markatutako desbideratze baten azterketa egitea erabakitzen du, konkretuki *sur hitza. Horretarako, proposatzaileari deitzen dio hitzaren forma estandarra ez baitu ezagutzen. Irudian ikusten denez, ikasleak proposamen guztiak, kopuru konkretu bat edo ikasle-ereduaren araberrako proposamenak (desbideratze tipikoak) eska ditzake. Adibidean ikusten denez, proposamen guztiak eskatuz gero zazpi desberdin agertzen zaizkio ikasleari; desbideratze tipikoak bakarrik kontuan izanik, ordea, proposamen bakarra (kasu honetan ikasleak bilatzen zuena).

VI.4 IRAKAZI: diagnosirako sistema lagungarria.

IRAKAZI inguruneak irakaslearekin elkarreragiten du, bai ikasleari buruzko ezaugarriak eskuratzeko bai ikasleak erabiltzen dituen egitura desbideratuen diagnostia lantzeko. Egitura linguistikoak HITESek sortutako tartehizkuntza erduetatik eskuratzen ditu IRAKAZIK. Honek irakasleari eskaintzen dion interfaze grafikoan ikasleari buruzko ezaugarriak eta bere egitura linguistikoak jasotzeko aukera emateaz gain, ikasleek ikasgelan egindako ariketei buruzko informazioa ere lortzen du. Gainera, informazioa lantzeko hainbat tresna ere eskaintzen du: egitura linguistikoaren sailkapenak, egiturak erabiltzeko arrazoi posibleen

zerrendak, aldi berean gertatzen diren egitura linguistikoen azterketa, e.a. Azkenik, IRAKAZIren bidez aztertutako diagnostia IDAZKIDEk erabiliko du.

HITESen bidez detektatutako desbideratze linguistikoak IRAKAZI sistemari pasa eta honen bidez irakasleak egituren diagnostia egingo du. Aipatutako egitura linguistikoak hainbat kasutan testu batetik jasoak izaten dira. Honegatik, IRAKAZI sisteman testuari buruzko informazioa jasotzen da; konkretuki ondoko datuak gordetzen dira:

- testu-kodea.
- hitz-kopurua.
- testu-mota (idazlana, eskutitza ...)
- idatzitako data.
- ikasle-kodea.
- egoera-kodea.

Testua idatzi duen ikaslearen kodearekin batera testua idatzi deneko egoera ere gordetzen da. Egoera honek irakasleak agindutako ariketa eta ikasgelan une horretan lantzen ari diren hizkuntz egiturak biltzen ditu. Ariketa izaten da: gai bat emanda idazlana idatzi; irakurritako testu baten laburpena; ikasgelan entzundako istorio baten kontakizuna, eta abar. Ikasgelan lantzen ari diren egiturak garrantzitsuak dira diagnosi-prozesu batean, ikasle batzuk landutako egiturak espreski erabiltzen saiatzen baitira testuak idaztean.

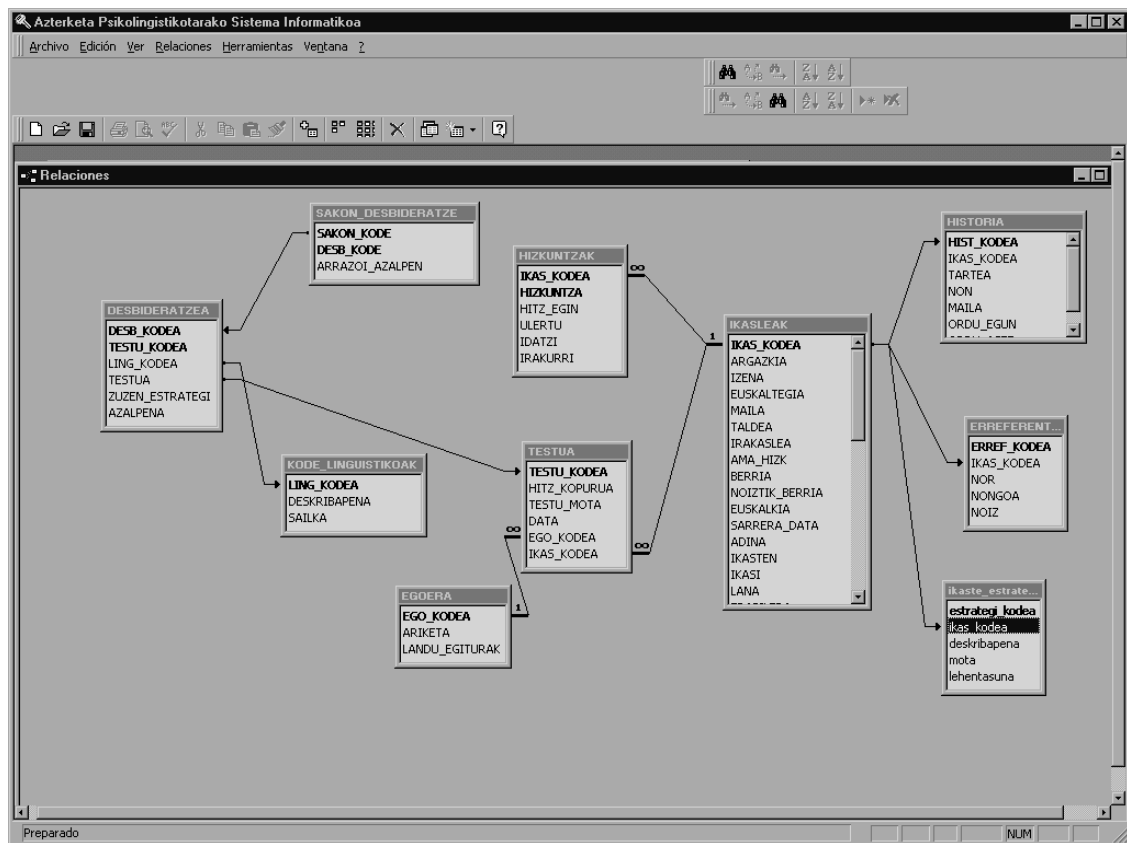
Desbideratzeen diagnostia lantzean irakasleak sakoneko arrazoiak azaldu eta sailkatzen ditu alde zuzenetik emandako sailkapen baten arabera (ikus C eranskina). Aztertu beharreko egitura testuaren testuinguruarekin batera aurkezten zaio, nahi izanez gero, zuzentzeko aukera emanik.

Laburbilduz, IRAKAZI sistema ondoko objektuek definitzen dute (ikus VI.13 irudia):

- Irakasleak.
- Testuak.
- Egoerak.
- Desbideratze-egiturek.

Ikasleei buruzko informazioan ikasleen ezaugarriak biltzen dira, tartehizkuntza eta ikaste-estrategiak izan ezik. Ezaugarriei buruzko informazioa HITESek erabiltzen du, eta IDAZKIDEk erabili eta alda dezake. Ondorioz, IRAKAZI eta IDAZKIDE sistemen arteko informazioaren sendotasuna lantzeko komunikazioa landu beharko litzateke. Guk hau ez dugu landu, baina etorkizunean hurbileneko garapenen artean ikusten dugu.

Ondorengo irudian aipatutako objektuen arteko erlazioak adierazten dira.



VI.13 irudia. IRAKAZI sistemaren objektuen arteko erlazioak.

VI.4.1 IRAKAZIren interfazea.

IRAKAZI sistemaren leiho nagusian aurkitzen ditugun aukerak ondokoak dira: datuei buruzko **kontsultak, aldaketak eta azterketak**.

Kontsultak zein aldaketak egitean ikasleei eta ariketei buruzko informazioa landu daiteke.

Ikasleei buruzko informazioa lantzean ondoko datu orokorrak kontsulta daitezke:

- Ezaugarri orokorrak.
- Egindako ariketak eta idatzitako testuak.
- Desbideratzeak.

VI.14 irudian ikasleei buruzko ezaugarri orokorren leihoetako bat azaltzen da, konkretuki ikaslearen ikaste-historia biltzen duena.

IKASLE KODEA: il9g

Izena: Pepi

Ondoren euskara non eta zein eritmotan ikasi duzun azaltzea eskatzen dizugu. Euskaldun zaharra izanik ikastaro edo antzekorik egin baduzu erantzun ezazu ere.

HISTORIALA				
TARTEA	NON	MAILA	ORDU EGUN	ORDU ASTE
78-82 (5 ur)	eskolan			2
83-86 (4 ur)	eskolan			2
92-93 (3 hil)	ilazki euskaltegian	4	2	

MENUA Hizkuntzak Ariketak Fitxak Inkesta Erreferentziak

VI.14 irudia. Ikasle baten ikaste-historiaren bilketa.

VI.15 irudian, berriz, ikaslearen desbideratzeak aurkezteko modua azaltzen da. Desbideratzearen kode linguistikoa eta testuaren testuingurua automatikoki azaltzen zaio irakasleari. Honek sakoneko kodea definitu behar du: sistemak aldeztu aurretik definitutako sakoneko arrazoiaren zerrenda aurkeztu dio irakasleari, honek sailkapen hori eguneratzeko aukera duelarik. Honetaz gain, irakasleak nahi izanez gero, desbideratzearen zuzenketa ere egin dezake, zuzenketa oharrak jartzeko leihotxo ere bertan duela. Irudian ikusten denez, irakasleak ikasleari berari eta beste ikasleei buruzko informazioa kontsultatzeko aukera du edozein unetan. Aukera honek konparaketak egiteko posibilitatea ematen dio.

IKASLE KODEA: ilAb

TESTU KODEA: ilAbm1

TESTU MOTA: DATA: 97/05/30

HITZ-KOPURUA: 294 EGOERA KODEA: Am1

Beste ikasle Batzuk
Beste ikasleak ikus ditzakezu

DESBERATZEAK DESB. KODEA: 1

KODE LING. MENLA_BE1 HITZA edo ESALDIA Batzuen ustez, mendeen zehar guregana etorri den ondare kulturala dela eta hitzegiten diren lekuetan gorde egin behar dugula, gal ez dadin, baina ez dute etorkizunik ikusten

SAKON KODEA #zNombre?

ZUZENKETA NOLA?

AZALPENA

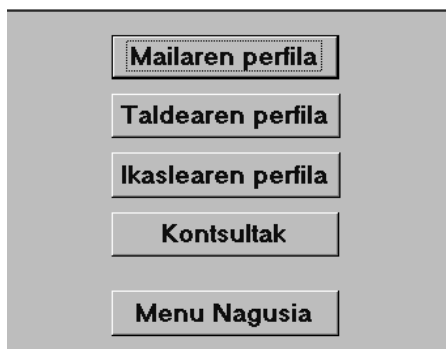
Registro: < > >> >>>

MENUA Hizkuntzak Fitxak Inkesta Historiala Erreferentziak

VI.15 irudia. Ikasle baten desbideratzearen diagnosi-azterketa.

Informazioaren azterketa ingurune honen hirugarren baliabidea dugu. Datuen esplotazioa lantzen den aukera honetan hiru motako azterketak egin daitezke, VI.16 irudian azaltzen denez:

- Ikaskuntza-mailaren perfila.
- Ikasle-talde baten perfila.
- Ikasle konkretu baten perfila.



VI.16 irudia. Datuen esplotazioa.

Edozein perfil lantzean hiru azterketa-mota egin ditzake irakasleak:

- Azterketa linguistikoa.
- Azterketa psikolinguistikoa.
- Azterketa konparatibo psikolinguistikoa.

Ondoren, hauei buruzko adibide batzuk emango ditugu:

Azterketa linguistikoa eta psikolinguistikoa

Bi azterketa hauetan diseinatutako esplotazioa antzekoa dugu. Irakasleak eska ditzakeen sailkapenak ondokoak dira, besteak beste:

- Ikaskuntza-maila baten kode linguistiko (*desbideratzea*) edo sakonekoen (*arrazoia*) zerrenda zein kode bakoitzaren erabilpen-kopurua maila osoan.
- Ikaskuntza-maila baten ikasle bakoitzak zenbat alditan erabiltzen duen kode konkretu bat.
- Data bat emanda zenbat ikasle dauden data horren aurretik euskara ikasten hasi zirenak eta kode konkretua erabiltzen dutenak.
- Kode bat emanda adibideak eman testuka ordenatuta.
- Eta abar.

Azterketa konparatibo psikolinguistikoa

Azterketa konparatiboaren helburua kode linguistiko eta sakoneko kodeen arteko erlazioak lantzen ditu. Honetan diagnosiaren azterketa egiten da zuzen zuzenean:

- Kode linguistiko baten arrazoen azterketa.
- Kode linguistiko konkritu bat eta sakoneko kode konkritu bat emanda zein ikaslerengan betetzen den diagnosi hori.
- Data bat emanda zein ikasle dauden data horren aurretik ikasten hasi zirenak eta diagnosi konkritu bat betetzen dutenak (hau da, kode linguistiko baten arrazoi konkritua betetzea).
- Eta abar.

VI.5 Inplementazio-egoera eta landu beharrak.

IDAZKIDE zein IRAKAZI sistemak inplementatuta daude, eta zenbait erabiltzailek erabili ditu. IRAKAZI bere osotasunean inplementatuta egon arren, interfazea barne, ez du irakaslea gidatzeko modulurik. Hau etorkizuneko lana gisa proposatzen dugu azken kapituluan. Honetaz gain, diagnosi-ezagutza izaneko ezagutza-basearen eguneraketa ere lantzeke dago. Ez da prototipo hutsa, hau da, eskuliburua egin eta bukaera batzuk eginez kalera daitekeen sistema dugu. IDAZKIDEn kasuan interfazea eta tresnen integrazioa eginda dago. Oraingoz ikasle gutxi batzuek erabili dute beren testuak lantzeko. Tcl-TKn (Brent, 95) egindako interfazea atsegina zaie ikasleei. Laguntza gidatua eskaintzen duen modulua integratu gabe dago, eta modulu honetan ikasle-eredua zati garrantzitsuena izanik, ereduaren eragina ezin izan dugu aztertu esperimenduetan. Simulazioen bat besterik ez da egin oraingoz. CLIPS lengoaiari inplementatuta dago eta interfazearekin integratzea falta da. Azkenik, IDAZKIDEn tresnen azkartasuna hobetu beharko litzake, oso garrantzitsua baita ikasleentzat.

VII. MUGARRI ingurunearen garapenean zeharreko ebaluazioa.

Kapitulu honetan, tesian zehar bigarren hizkuntzaren ikaste-prozesua aztertzeke aurkeztu dugun MUGARRI ingurunearen ebaluazioari buruz arituko gara. Azpimarratzekoa da ebaluazioa ez dela izan bukaeran bakarrik egindako azterketa, baizik eta ingurunekeo sistemen garapena gidatu duen eragilea. Ebaluazioak planifikazioan laguntzen du egindakoa hobetuz, produktu hobea ateratzeko helburuarekin. MUGARRI ingurunearen garapenean zehar hainbat esperimentu burutu da sistema ebaluatzeko asmoz. Esperimentuetan erabilitako metodologia III. kapitulan azaldu genuen esperientzia konkretu baten azalpenarekin batera. Han ere aipatutako hainbat esperimenturen emaitzak emango ditugu kapitulu honetan. Honetaz gain, kapituluaren azken atalean ingurunearen funtzionamenduari buruzko datuak emango ditugu. Datu hauek MUGARRI ingurunearen erabilpenean psikolinguistik edo irakasleak atera ditzaketen ondorio-motak emango dizkigute.

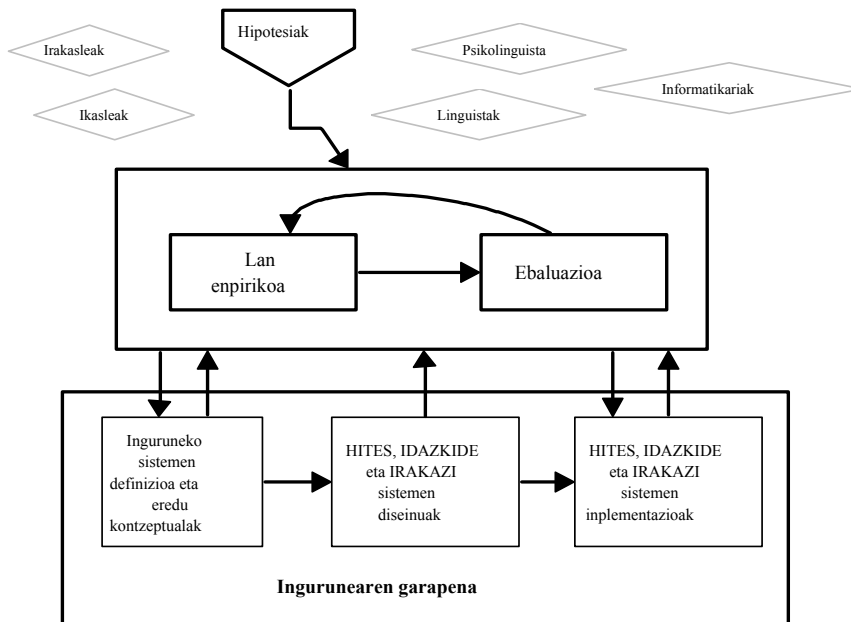
VII.1 Sarrera.

Ebaluazioa eta emaitzak aztertzean kontuan izan behar dugu garatutako MUGARRI izeneko ingurunean hiru sistema diseinatu eta inplementatu ditugula, hots, HITES, IRAKAZI eta IDAZKIDE. Hori dela eta, bi fase bereiztu ditugu nagusiki ebaluazioa azaltzeko: lehenengoan, hiru sistemen diseinuen eta eredu kontzeptualen definizioa, eta, bigarrean, sistema hauen inplementazioa. Hiru sistemen eredu kontzeptualak eta diseinuak aldi berean egin dira; inplementazioak, ordea, une desberdinetan landu dira. Hori dela eta, *lan enpirikoaren ebaluazioa eta emaitzak* eta *ingurune informatikoaren emaitzak* izeneko azpiatalak banatu ditugu kapitulu honetan. Lan enpirikoa sistemen eredu kontzeptualetan eta diseinuetan laguntza izan den bezala ingurune informatikoaren ebaluazioak inplementazioaren hobekuntza-prozesuan lagundu du.

Azkenik, ebaluazioaren partehartzaileak aipatuko ditugu: hizkuntz irakasleak, ikasleak, psikolinguista, informatikariak eta linguistak. Denen iritziak oso garrantzitsuak izan dira prozesu osoan, baina iritziak jasotzea ez da izan lan erraza. Jakintza-arlo desberdinetakoak ditugu ebaluatzaileak, hizkuntza desberdina hitz egiten dutenak; aipatzekoa da, esaterako, 1994/95 ikasturtean ILAZKI euskaltegiko bi irakaslerekin izandako esperientzia. Kasu hartan, beraien ekarpenekin desbideratzeen kodeketa linguistikoak hobeto zehaztu genituen. Esperimentuak eta partehartzaileen papera III. kapitulan zabalago azaldu genituen, hortaz, kapitulu honetan emaitza konkretuagoi buruz arituko gara.

Esan bezala, ebaluazioa MUGARRI ingurunearen garapenaren eragilea izan da, eta ebaluazioan parte hartu duten agenteak jakintza-arlo desberdinetakoak izan dira.

Ondoko irudian ingurune garapenaren eskema azaltzen da. Bertan islatzen da ebaluazioak prozesuan izandako garrantzia.



VII.1 irudia. Ingurunearen garapenaren eskema: ebaluazioa hobekuntzaren eragilea gisa.

VII.1 irudian ikusten den bezala MUGARRIn sortu diren sistemen garapena prozesu zikliko gisa planteatu da, non ebaluazioa ziklo bakoitzeko fase bat den; hau da, ez gaude prozesu lineal baten aurrean, non ebaluazioa bukaerako pausoa den. Lan-metodo honetan hiru ziklo nagusi agertu dira:

- Sistemen definizioa.
- Sistemen diseinua.
- Sistemen implementazioa.

Sistemen definizioa zikloaren helburu nagusia beharren identifikazioa izan da, konkretuki, bigarren hizkuntzaren ikaskuntza arloan dauden beharren artean zeintzuk landu daitezkeen informatikaren teknikak erabiliz.

Sistemen diseinua sistemaren definizioa zein eredu kontzeptualak zehaztu ondoren hasten da garatzen. Bigarren ziklo honen helburu nagusia da, behin beharrak identifikatuta, hauek asetzeko sistema informatikoa diseinatzea.

Sistemen implementazioaren helburua definizioan zehaztutako beharrei erantzuteko sistema erdiautomatikoaren gauzatzea da.

VII.2 Lan enpirikoaren ebaluazioa eta emaitzak.

Atal honetan ikaskuntza-mailaren araberrako azterketari ekingo diogu, bai hitz-mailan, bai esaldi-mailan. Honetaz gain, ezagutza linguistikoaren egonkortasunaren azterketa ere komentatuko dugu. Honetarako, egitura linguistikoren bat *gaizki* erabiltzearen ondorioz gerta daitezkeen desbideratze linguistikoak aztertuko ditugu batez ere; gainera, III. kapituluari desbideratze kontzeptua definitzean aipatutako lehenengo kategoria landu dugu gehienbat, hau da, zentzu klasikoan errorea deitutakoa. Dena den, ez dugu ahaztu behar kasu batzuetan erregela estandarrak ere landu direla.

Gure helburuetako bat egitura idatzi desbideratuen sailkapen bat definitzea izan da, eta sailkapenak ezin du eduki bera izan ikaskuntza-maila guztietan. Bi maila agertzen zaizkigu hemen, azaleko desbideratzeak (adib. *haiek zeunden haiek zeudenen* ordez erabiltzea) eta sakoneko desbideratzeak (adib. bi hitzek antz handia dute). Hau dela eta, sailkapen-mota bakarria ez da nahikoa izango, bi sailkapen-mota nagusi behar ditugu: azaleko sailkapena eta sakoneko sailkapena. Honegatik, desbideratze bat aztertzerakoan azaleko sailkapenean dagozkion kodeez gain sakoneko sailkapenean dagozkionak ere identifikatuko ditugu. Gainera, kontuan izan behar dugu desbideratze bati ez dagokiola beti azaleko mailako desbideratze-mota bakarria, ezta sakoneko bakarria ere. Desbideratze bera egitearen arrazoiak bat baino gehiago izan daitezke, dudarik gabe.

Jarrai dezagun aztertzen nondik abiatu sailkapen hauek. Lehengo adibidean, sakoneko arrazoi gisa hitzen arteko antza aipatu dugu. Azaleko mailan, desbideratze berari bi ikuspegi desberdin eman dakioke: ikuspegi linguistiko batetik singularreko bigarren pertsona eta pluraleko hirugarrena nahastu ditu ikasleak, eta azaleko ikuspegi huts batetik bi hitzen arteko ordezkapena egin du. Beraz, azaleko mailan beste bi maila bereizi ditugu: azaleko maila hutsa eta maila linguistikoa. Honegatik, azkenean, hiru sailkapen desberdin definitu behar izan ditugu: azaleko mailan bi eta sakoneko mailan bat. Sailkapen hauek III. kapituluari azaldu genituen. VII. kapituluari, ordea, sailkapenak lortzeko egindako lan esperimentalaren emaitzak emango ditugu.

VII.2.1 Hitz-mailako azterketa ikaskuntza-mailaren arabera.

Hitz-mailako azterketaren helburua erabilitako erregela linguistikoaren sailkapena lortzea izan zen. Gehienbat, azaleko mailen sailkapenak aztertu genituen; egitura linguistikoak erabiltzeko arrazoiak lantzen dituen sakoneko maila, berriz, esaldien azterketarako utzi genuen. Sakoneko arrazoiaren azterketan hitz-mailan egindako saiakeretan ez zen informazio handirik ondorioztatzen, esaldi-mailan, berriz, diagnosirako arrazoi ugari identifikatzen ziren.

Hitz-mailako lan enpirikoan, hasiera batean, hainbat testu hitzez hitz aztertu genuen eskuz, inolako tratamendu informatikorik aldeztirik egin gabe. Geroago egindako esperimentuak, ordea, erdiautomatikoak izan ziren. Edozein kasutan, azterketa hauetan landa-lana genuen abiapuntua, eta informatikaren betebeharra gehienbat lanak erraztea izan zen.

Azaleko maila hutsean lortutako sailkapenaren lan esperimentalean, hiru fase nagusi definitu genituen:

1. Sailkapenaren definizioa.
2. Sailkapenaren ulermena eta baliagarritasuna.
3. Sailkapenaren azterketa automatikoa.

1. Sailkapenaren definizioa

Sailkapenaren definizioaren lehen azterketa (Maritxalar et al., 93) 1991/92 ikasturtean zehar egin zen, IRALE programaren 6. mailan zeuden 29 ikasleek idatzitako 47 testu (7956 hitz) landu genituelarik, guztira 14 langai desberdin. Azterketa honetatik A eranskinean agertzen den sailkapenaren lehen zirriborroa lortu genuen.

Kontuan izan behar da sakatze-erroreak ez zirela sartu aipatutako sailkapenean. Honetarako, ikasleak eskuz idatzitako testuak eta ordenagailura pasa zirenak konparatu egin ziren banan banan, sakatze-erroreak bereizteko asmotan. Konparaketa honetan 42 sakatze-errore detektatu ziren (%7,8), ondoren aipatuko ditugun desbideratze-motez gain. Aipatu beharrekoa ere bada desbideratze batzuk sailkagaitzak (% 5,2) gertatu zitzaizkigula abiapuntu gisa hartu genuen sailkapena oinarri izanik. Ondoren desbideratzeak identifikatzeko kodeketa azalduko dugu.

Azalpenean, desbideratze konkretu bakoitzari jarritako kodea azalduko dugu, eta mota bakoitzean topatutako desbideratzeen portzentajea aipatu:

a) LE = letra / LM = lema / MO = morfema, lema ez dena.

(Kodifikazio honetan morfema (MO) diogunean lema ez diren morfemak kontsideratzen ditugu)

b) LEAX = letra bat (X) Asmatu du.

LETX = letra bat (X) hitzaren barruan Tokiz aldatu du.

LEKPX = letra bat (X) Kendu du hitzaren P posizioan. (% 11,7)

LEGPX = letra bat (X) Gehitu du hitzaren P posizioan. (% 2,7)

LEOPXY = letra baten (X) Ordez Y letra jarri du hitzaren P posizioan (X eta Y letra bat edo *ts*, *tz*, *tx* letra-pareak izan daitezke). (%32,6 = LEOPXY + LEAX + LETX)

non P = H (hitzaren lemaen Hasieran) /

B (hitzaren lemaen Bukaeran) /

- T (hitzaren lemaen barrualdean, hau da Tartean) /
E (hitzaren lemaen Edozein tokitan)
- c) MOA-X = morfema bat (X) Asmatu du. (%4,6)
MOK-X = morfema bat (X) Kendu du. (%2,6 = MOK-X + MOG-X + MOO-X-Y)
MOG-X = morfema bat (X) Gehitu du.
MOO-X-Y = morfema baten (X) Ordez Y morfema jarri du.
- d) LMA-X = lema bat (X) Asmatu du. (%31,5)

Bigarren azterketa batean, 1992/93 ikasturtean zehar beste euskaltegi batetik ere testuak jaso genituen. Jasotakoan artean 8. eta 9. mailako 20 testu (3777 hitz) ausaz hautatu genituen. Hauek 10 ikaslek idatzitakoak ziren, eta langai desberdinak zituzten. Ondoren, testu hauek erdiautomatikoki landu ziren aurretik lortutako sailkapena ebaluatzeko. Sailkapena baliagarria suertatu zen, aurretik definitutako fenomenoak berriro azaldu zirelako, fenomeno berriak ere sortu baziren ere.

2. Sailkapenaren ulermena eta baliagarritasuna

Sailkapenaren ulermena: fase honetan sailkapenaren definizioan parte hartu ez zuen irakasle batek 10. mailako 4 ikaslek idatzitako 16 testu (3705 hitz) landu zituen. Proba hartan 4 langai jarrita ikasle bakoitzak 4 idazlan idatzi zituen. Fase honetan irakasleak ez zituen testu gordinak aztertu, testuetan agertutako erregelak baizik. VII.2 irudian agertzen da irakasleari azaltzen zitzaizkion erregelaren fitxak. Hauetan erregelaren deskribapena, hizkuntz ezaugarriak, egituraren ezaugarriak eta hainbat adibide ematen zaizkio irakasleari.

Lan esperimental honek hiru helburu izan zituen:

- Kodeketaren ulermena aztertzea.
- Kodeketaren baliagarritasuna berrestea.
- IRAKAZI sistemaren interfazearen diseinua lantzea.

Proba honakoa izan zen:

Testuak automatikoki aztertu genituen hitz-mailako erregela linguistikoak detektatzeko, bai erregela estandarrak bai desbideratzeak, gero irakasleak lan zitzan. Prozesuan VII.2 irudian azaltzen denaren moduko fitxak automatikoki sortu ziren. Fitxa horiek gero irakasleak eskuz aztertuko zituen (*letra etzanaz idatzitakoa*). Lehen esan bezala, fitxa bakoitzean erregelaren deskribapena, adibideak, hizkuntz ezaugarriak eta egituraren ezaugarriak lortzen dira automatikoki. Aurretik esan dugun bezala, fitxa hauek erregela linguistikoaren kodeketaren ulermena zein baliagarritasuna lantzeko eta IRAKAZI sistemaren interfazea diseinatzeko aztertu ziren. Ulermenari dagokionez, irakasleak erregela ulertu izanez gero, bere azalpena ematen zuen; gainera, bere ustetan 10. mailan erregela ematen

den ala ez esaten zuen (desbideratzea den ala ez markatuz); baiezko kasuan adibidea bilatzen zuen testu berri batzuetan (10. mailako 41 testu (10311 hitz) 8 ikasle desberdinek idatzita); eta, azkenik, 10. mailan agertutako erregela berria ote den edo beheko mailetan sortutakoa ote den zehazten zuen.

LEGBE	
• <u>Deskribapena:</u> E letra gehitu du hitzaren lemaren bukaeran	
• <u>Adibideak:</u> 1995.eko hamabost honetan on*	
• <u>Hizkuntz ezaugarriak:</u>	• <u>Egituraren ezaugarriak:</u>
- Leman: EZ DA DETEKTATU	- Kontsonantez/Bokalez: <i>Bokalez</i>
- Atzizkian: ONDOKO KATEGORIETAN DETEKTATU DA:	- Hitzaren luzera: <i>Luzea eta Motza</i>
DET <i>Horretan</i>	- Ariketa-mota:
ADJ <i>Politei</i>	
IZE <i>Neskei</i>	
ADL <i>Direnak</i>	
ADT <i>Direnak</i>	
ADI	
SIG	
IOR	
- Aurrizkian: EZ DA DETEKTATU	
• <u>Erregelaren ulermena (BAI/EZ/EZIN):</u>	- Adibidea:
- Azalpena: <i>BAI: A + E = E</i>	- Testua: <i>il10jma</i>
	- Hitza: ... <i>zer gertatuko den ...</i>
• <u>Noiz sortua (10/<10/EZIN):</u> <10	
• <u>Ebaluazioaren oharrak:</u>	
IRAKASLEAK EZ DU ERREGELA ZEHARO ULERTU. BERAK JARRITAKO ADIBIDEAK ETA DEFINITUTAKO EGITURAREN EZAUGARRIEK ADIERAZTEN DUTE A + E = AE FENOMENOA ULERTU DUELA. LEGBE ERREGELAK, ORDEA, -E LETRA GEHITZEA FENOMENOA BESTERIK EZ DU ADIERAZTEN.	

VII.2 irudia. Automatikoki sortutako fitxa irakasleak aztertua (letra etzanaz) eta ebaluatua (letra larriz).

Kodeketaren baliagarritasuna: § IV.5-n analizatzaile morfologikoaren egokitzapenari buruz aritu ginen. Egokitzapen haren helburua euskara ikasten ari direnen hainbat desbideratze zein erregela estandar identifikatzea zen, konkretuki 12. mailako eta hurrengo mailako (10-11. maila) ezagutza landu genuen. Kodeketa haren baliagarritasuna aztertzeko 16 testu hartu eta hasierako analizatzailearen 61 erregela haien agerpena landu genuen. Emaitza izan zen hamabi erregela ez zirela agertu. Goitik beherako metodologia aplikatuz suposatzen da ez detektatutako erregelak hamabigarren mailari dagozkiela, eta ez hamargarrenari. Dena den, ez detektatuak ikusi ondoren lau erregela 10. mailakoak ere baziren irakaslearen ustez. Kasu batean arrazoa izan zen testuetan ez zela suertatu honelako erregularik, beste hiruretan, ordea, erregela-ereduen inplementazioan zegoen arazoa. Inplementazio-arazoa behin konponduta emaitza nahiko onak dirudite. Hala ere azken ñabardura bat egin beharko da. Emaitzak zehatzago landuz ikusi genuen detektatutako 61

erregelen artean hamar kasutan¹ hitzaren interpretazioa 10. mailakoa ez zen lema bati zegokiola. Hau gertatzen zen 10. mailako hiztegi zehatza ez izateagatik, edo hitzaren interpretazioen desanbiguaziorako erregelak zeharo zehaztu gabe egoteagatik. Horrek adierazten du abiapuntu gisa hartutako kodeketa baliagarria zela, baina ikaslearen tartehizkuntza gehiago findu behar dela; irakaslearen laguntzaz esaterako.

Kodeketaren hobekuntza-prozesu honetan probaren bukaeran azken saio bat egin genuen irakasleak bilatutako adibide berriak automatikoki tratatuz. Kasu batzuetan irakasleak jarritako adibideek erregela linguistikoen testuinguru hobeki zehaztera behartzen gintuzten. Honela, interfazean irakasleari eskatutako adibideak analizatuko lirateke, eta ondo egokituko ez balira, erregelen testuinguruari buruzko zehaztasunak eskatuko litzaizkioke irakasleari, modu honetan ikasleen tartehizkuntza gero eta gehiago fintzen delarik.

Kodeketaren ulermena eta Interfazearen diseinua: kasu honetan fitxaren formatua lantzean hainbat ohar jaso genuen irakaslearengandik. Esaterako, hasiera batean, lehen probetan adibideak agertzen zirenean hitza bakarrik agertzen zen, eta ez, berari zegokion lema, VII.2 irudian azaltzen den moduan. Irakaslea ez zen linguista, eta, horregatik, kasu batzuetan emandako adibideetan ez zuen erregelaren aplikazioa ulertzen hitzaren lema zein zen esateko gai ez zelako. Edonola ere, adibideekin edukitako arazoez gain irakaslea ez zen gai izan hamasei erregela zuzenki interpretatzeko, edo interpretatu zituenean beste erregela linguistiko bat ulertu zuen. Guztira 61 erregela izanik proportzio handia da. Arrazoiak ondokoak izan ziren:

- Adibide ez egokiak, mailako ezagutza linguistikoaren zehaztasun faltagatik (hiztegia eta desanbiguazio-erregelak).
- Erregelaren deskribapen linguistiko zabal baten beharra (emandakoa azaleko mailakoa zen).
- Kodeketak ematen duen informazioa nahikoa ez izatea (ikus VII.2 irudia).

Honen guztiaren ondorio gisa, IRAKAZI sistemaren interfazea egokitzearen beharra sortu zen. Konkretuki, irakaslearentzako inguruneak tartehizkuntzaren erregela linguistikoei buruzko informazioa eskuragarri izan behar du, edozein unetan ematen zaiona baino informazio gehiago eskatu ahal izateko, adibidez, erregela baten aplikazio-eremuak, adibideak beste corpus batzuetan, eta abar.

¹(LEKHH, LEOEXJ, LEOESZ, LEOEXZ, LEOEUO, LEOEOU, LEOEGJ, LEOEBS, LEOETD2, LEOEDT_GK2)

3. Sailkapenaren azterketa informatikoa

Azken fase batean, eskuz lortutako emaitzetatik abiatuta hitz-mailako fenomenoaren azterketa automatikoari ekin genion. Honetan goitik beherako metodologiaren aplikazioa burutu genuen. Hasiera batean ikaskuntza-maila goreneko (12. maila) erregelen definizioa lortu genuen eskuz egindako azterketetatik. Horretatik abiatuta beheko mailen azterketari ekin genion. Honi buruz § VII.3-n zehatzago arituko gara.

VII.2.2 Esaldi-mailako azterketa ikaskuntza-mailaren arabera.

Esaldi-mailako azterketaren helburua erabilitako erregela linguistikoen sailkapena eta hauek erabiltzeko arrazoen sailkapena lortzea izan zen. Horretaz gain, lortutako sailkapenen baliagarritasuna landu genuen ikasle zein irakasle konkretuei proba batzuk eginez. Esaldi-mailan egindako lan enpirikoan ebaluazioaren agente guztiek hartu dute parte. § III.4.2-n azaldutako esperientziaren barruan esaldi-mailako desbideratzeen azterketa zabalena egin zen. Aipatutako atalean, beraz, esperimentuaren helburu orokorrak, jarraitutako metodologia eta egitura linguistikoak zein ikasleak aukeratzeko irizpideak dagoeneko azaldu genituen. Orain dagokigun atalean, berriz, ebaluazioaren ikuspuntutik azalduko dugu esaldi-mailako desbideratzeen azterketa, erabilitako datuak eta lortutako emaitzak batez ere azpimarratuz.

Esaldi-mailako azterketan bi fase desberdin bereiztu genituen:

1. Desbideratze linguistiko zein sakoneko arrazoen sailkapenen definizioa.
2. Desbideratze linguistiko zein sakoneko arrazoen sailkapenen berrespena.

1. Desbideratze linguistiko zein sakoneko arrazoen sailkapenen definizioa

Azterketaren lehenengo fase batean, irakasleek informatikariarekin batera hasierako corpusa zehaztu zuten. Corpusa 9., 10. eta 12. mailako ikasleek idatzitako testuek osatzen zuten: 9. mailako 3 ikasleek idatzitako 17 testu (3737 hitz eta 241 esaldi); 10. mailako 8 ikasle desberdinek idatzitako 30 testu (6907 hitz eta 436 esaldi); eta 12. mailako 4 ikasleek idatzitako 17 testu (5569 hitz eta 364 esaldi). Ondoren, irakasleek psikolinguista eta linguistekin batera egitura linguistikoak identifikatu zituzten. Eta, azkenik, egiturak erabiltzearen sakoneko arrazoen sailkapena zehaztu zuten. Kontuan izan behar dugu irakasleek ikasleen uneko ikaste-prozesua ezagutzen zutela zuzen zuzenean. Ikuspegi informatiko batetik abiatuta irakasle desberdinen artean arrazoen sailkapen bakarra lortzea zen helburua hasiera batean. Esperimentuek adierazi ziguten, ordea, irakasle bakoitzak bere ikuspegia jarraitzea aberatsagoa izango zela. Beraz, azkenean, ikasleak ezagutzen zituzten bi irakasle desberdinek aztertu zituzten testuak, bakoitzak bere ikuspegia mantenduz arrazoiak sailkatzean. Irakasle bakoitzaren ikuspegia honela defini daiteke laburbilduz: batek desbideratzeen arrazoiak esaldien barruko egituraketan bilatzen zituen (adib. esaldi luze

batean ikasleri subjektuan -k markatzea ahaztu zaio aditza subjektutik urruti egoteagatik), eta besteak ikaslearen ezaugarri pertsonaletan aurkitzen zituen sakoneko arrazoiak (adib. idazketan ikasleak arreta eskasa jarri du esaldiaren zuzentasunean).

Azterketa-lan honetan irakasle bakoitzak ikasle bakoitzaren perfila ere landu zuen (Andueza et al., 96), hau da, desbideratzeak aztertu zituen ikaslearen ezaugarri pertsonalak kontuan izanik: hizkuntza ikasteko arrazoi pertsonalak, maila lortzeko emandako denbora hizkuntza ikasten hasi zenetik, zuzentasunari jartzen dion arreta, erabiltzen dituen ikaste-estrategiak, eta abar. Azterketa honen ondorioz ikasle-ereduaren ezaugarriak definitu genituen. Ezaugarri hauek HITES eta IDAZKIDE sistemek erabiltzen badituzte ere, IRAKAZIren bidez (ikus VI.13 eta VI.14 irudiak) jasotzen dira irakasleak emanda. Horretaz gain talde bakoitzari buruzko perfila landu zen, eta ondorio nagusia izan zen 9. eta 10. mailen arteko diferentzia nabarmenik ez zegoen bitartean, 10. mailatik 12. mailara salto handia zegoela; kontuan izan behar dugu 10. mailako ikasleek ikasturte bukaeran 11. maila ere gainditzea lortu behar zutela. Hortaz, IV. kapituluan azaldutako goitik beherako metodologian tartehizkuntza mailaka lantzea diogunean gertuko maila desberdinez osatutako multzoek osatzen dute berez aztertze-maila bakoitza. Hemen aztertzen ari garen corpusean goiko maila baxua (9. eta 10-11. maila) eta goiko maila altua (12. maila) bereizten ditugu.

Goiko maila baxuan desbideratze gehienak aditzean, deklinabidean eta menpeko esaldien marketan agertu ziren.

- Aditzari dagokionez ikasleek singularrerako joera nagusia dute, batzuetan aditzaren singularra forma plural baten antzekoa izateagatik (adib. *zeuden zeunden*), beste batzuetan esaldian aditza komunztatu behar duen osagaitik urrun egoteagatik, eta hainbatetan aditz laguntzaileetan arreta gutxiago jarri izanagatik. Horretaz gain, multzoa adierazten duten osagaietan aditzarekiko komunztaduran pluralerako joera izaten dute, nahiz eta singularrean jokatu behar.
- Deklinabidea dela eta, ergatiboaren ordeztzatibo egiteko joera izaten dute, eta mugagabearen erabilera saihesten dute osagaiak mugatuz. Joera hau izaten dute batez ere *zenbait* eta *hainbat* bezalako partikulak agertzen direnean, edo *zertan zehar* eta *noren bidez* bezalako eraikuntza berezietan. Maila honetan azpimarratzekoa da mugagaberako joera izaten dutela askotan sintagmaren esanahiak erreferentzia abstraktu edo orokor bati egiten dionean, gehienetan ikaslearen ama-hizkuntzan mugatua ez dena (adib. *bi buelta dago lehendakari aukeratzeko*).
- Menpeko esaldien kasuan ikasleek menpekoa ez dute markatzen esaldi luzeetan eta zehar galderetan; beste batzuetan (ikus B eranskina), ordea, behar ez denean markatzen dute. Azkenik, ezaugarri pertsonalei dagokienez maila honetan arinegi idazten saiatzen dira batzuk. Erreferente euskaldunak izatea, hizkuntza behartzea eta

lanbideak behartua ikastea ere izan dira desbideratzeetan eragina izan duten faktoreak.

Goiko maila altuan desbideratze-kopurua txikiagoa da goiko maila baxuan baino, ikasleek hizkuntz arazoak non dituzten errazago identifikatzen dutelako.

- Aditzari dagokionez ikasleek singularrerako joera mantentzen dute esaldian aditza komunztatu behar duen osagaitik urrun egoteagatik, eta hainbatetan aditz laguntzaileetan arreta gutxiago jarri izanagatik.
- Deklinabideko desbideratzeen portzentajea oraindik garrantzizkoa da maila honetan. Mugatasuna eta mugagabearen arteko bereizketan arazo ugari izaten dituzte oraindik ere. Maila honetan maila baxuan dagoen joera mantentzen dute, hau da, mugatzerako joera behar ez denean, batez ere *zenbait* eta *hainbat* bezalako partikulak agertzen direnean, edo *zertan zehar* eta *noren bidez* bezalako eraikuntza bereizietan. Behar ez denean mugagabea erabiltzea ez da gertatzen maila baxuko maiztasunarekin, dirudienez, ama-hizkuntzaren eragina kontzeptu mugagabeei dagokienez desagertzen da neurri handi batean goiko maila altuan.
- Menpekoekin azpimarratzekoa da goiko maila baxuan duten arazoa nabarmenki desagertzen dela ikasle gehienengan.

2. Desbideratze linguistiko zein sakoneko arazoien sailkapenen berrespena

Bigarren fase honetan definitutako desbideratzeen sailkapenak berresteko asmoz teknika desberdinak erabili genituen. Alde batetik, irakasleek corpus berria landu zuten, beste aldetik, irakasleek egindako interpretazioak oinarri hartuta ikasleei galderak egin genizkien beraien interpretazioei buruz, eta, azkenik, irakasle berri bati desbideratzeen zuzenketak eskatu genizkion zuzenketen azpian ikusten zen diagnosiaren bidez sakoneko desbideratzeak berresteko asmoz.

2.1 Berrespena corpus berria erabiliz

Bigarren fase batean irakasle bakoitzak hainbat testu berraztertu zituen beste irakaslearen ikuspegia oinarri hartuta. Esperientzia honek ikuspegi berri bat eskaini zion irakasle bakoitzari. Bai kasu batean bai bestean ikuspegi berrira moldatu ziren bi irakasleak. Honelako esperientziak adierazten digu IRAKAZI sisteman irakasleen arteko komunikazioa sortu beharko litzatekeela egiten diren diagnosiak hobetzeko asmoz. Hala ere, hau saiakera bat besterik ez zen izan; sistemen garapenean ez gara sartu zehaztasun maila horretan.

2.2 Berrespena ikasleei galdetuz

Hirugarren fasean ikasleak bihurtu ziren agente nagusiak. Ikasle konkretuekin egindako azterketan ondokoak landu genituen:

3.1 Egitura desbideratuen zergatiak.

3.2 Ikasleen jabetzea egitura desbideratuen zergatiez.

Lehenengo proban, irakasleek idazlanetatik ateratako ondorioak berresteko edo baztertzeko ikasleei proba berezia jarri genien. Proba honekiko joera positiboa zuten lau ikasle izan ziren proba burutu zutenak, konkretuki maiz ikasgelara joaten zirenak. Probaren helburu nagusia zen hainbat egitura desbideraturen zergatiak aztertzea; horren barne ariketa-motak eta ikasleen ezaugarri pertsonalek desbideratzeetan duten eragina ere landu genituen.

Bigarren proban, ikasle batzuei zuzenean galdetu genien egitura linguistiko konkretu batzuen erabilerari buruzko arrazoi sakonak zeintzuk ote ziren. Proba honen helburua zen aztertzea zenbateraino jabetzen ziren ikasleak egindako desbideratzeez, eta hauen zergatiez.

Ondoren, aipatutako probak zehatzago azaltzeari ekingo diogu.

2.2.1 Egitura desbideratuen zergatien azterketa

Irakasleen ustezko ondorioak aztertzeko goiko mailan zeuden 4 ikasleri proba bat egin genien. Proba ondoko pausoetan banatu zen:

- a) Irakasleen hipotesien zehaztapena.
- b) Ikasleentzako azterketaren prestaketa.
- c) Hasierako hipotesien egiaztapena azterketaren emaitzak direla medio.
- d) Ondorioak.

a) Irakasleen hipotesien zehaztapena

Froga honen bidez sakoneko arrazoiaren azterketan irakasleek identifikatutako zenbait hipotesi aztertu nahi genuen. Ikasleek erabiltzen dituzten egitura desbideratuen arrazoiak asko badira ere, hauetako batzuk aukeratu, eta horiek egitura linguistikoen osaketan duten eragina aztertu dugu. Landu ditugun faktoreak ondokoak dira:

- Ariketa-mota.
- Esaldiaren hizkuntz ezaugarriak (hots, esaldiaren luzera, esaldiaren beste osagai batzuekin interferentzia, menpeko esaldien erabilera...)
- Ikaslearen ama-hizkuntza.
- Bigarren hizkuntzaren egiturak edo hitzak ama-hizkuntzan tartekatuta erabiltzea.

- Egitura linguistikoaren sinplifikazioa.
- Ama-hizkuntzaren idazketa-maila.

Aipatutako faktoreen eragina lantzeko lau ariketako (hots, itzulpena, okerrak zuzen, esaldiak berridatzi eta idazlana ama-hizkuntzan) froga-azterketa (ikus D eranskina) egitea erabaki genuen. Irakasleek definitu zuten hipotesi-zerrenda ere aipatutako eranskinean aurki daiteke.

b) Ikasleentzako azterketaren prestaketa

Aurreko atalean irakasleek beren hipotesiak definitu zituzten. Oraingoan, berriz, hipotesi horiek frogatzeko azterketa prestatu genuen ikasleei aurkezteko (ikus D eranskina).

c) Hasierako hipotesien egiaztapena azterketaren emaitzak direla medio

Aurretik aipatutako froga-azterketaren emaitzek irakasleen hipotesi gehienak berretsi badituzte ere, batzuk ez dira egiaztatu ezta kasu batean ere. D eranskinean inoiz ez eta kasuren batean bete diren hipotesiak komentatzen ditugu. Adibide gisa, bi ikasleren kasua komentatzen dugu zehatz-mehatz.

d) Ondorioak

Laburbilduz, kasu gehienetan hipotesiak bete dira, ama-hizkuntzaz idatzitako idazlanaren mailari dagokiona izan ezik. Lau ikaslek egindako idazlanak direla medio, ez da ikusten ama-hizkuntzan ikasleak duen mailak eragina duenik. Edozein kasutan, ikasle gehiagoren arteko azterketa sakonagoa egin behar da ondorioren bat atera ahal izateko, eta, agian, eragina topa genezake.

Ikasleen ezaugarriak kontuan izanik, bi ikasle desberdinen arteko konparaketa egitean ematen du bata itzultzaile hobea dela bestea baino. Euskaratze zein berridazketa ariketetan lehenengoa trebeagoa den bitartean zuzenketan bigarrena hobea da. Irakasleen ustetan lehenengoaren kasuan erreferentzia euskaldunak ez izateak eragina eduki dezake. Bigarrenak, berriz, erreferentzia euskaldun asko izan ditu bizitzan zehar.

Ikaslearen ezaugarriez gain ikaslearen uneko egoerak ere eragina du erabilitako hizkuntz egituretan. Hemen azaltzen ari garen frogan esaterako hirugarren ikasle batek, desbideratze tipiko ugari egiteaz gain ongi dagoen eta ez dagoenaren artean bereizteko zailtasunak ditu. Froga egin zuen momentuan euskara nahiko alderatua zuen, eta eguneroko bizitzan ia ez zuen erabiltzen.

Egitura desbideratuen zergatien azterketa honetan tesian zehar garatu ditugun sistema informatikoetarako hainbat ondorio atera dugu. Aipagarrienak komentatuko ditugu:

- * Irakasleentzat, lehenengo fasean egitura linguistikoak lantzean aztertutako corpusean oinarrituta identifikatu ziren testuinguruak (arrazoi-sailkapena irakasle baten ikuspuntutik) ez ziren nahikoak suertatu 2. eta 3. fasean egindako corpus kodeketan. Horregatik, HITES eta IRAKAZI sistemetan bai psikolinguistari bai irakasle bakoitzari bere sailkapena izateko aukera ematea erabaki genuen.
- * Egitura linguistikoaren erabilera zenbait egitura linguistiko maila bereko ikasleentzat orokorrak izan arren, badira ikasle konkretuei dagozkien erabilera-moduak. Ondorioz, HITESen tartehizkuntzaren adierazpidea antolatzean ikaskuntza-mailaren tartehizkuntza eta ikaslearen tartehizkuntza propioa bereiztu ditugu.
- * Ariketa-motak ikaslearen output-ean duen eragina funtsezkoa da.
- * Proba hasieran ez genuen kontuan izan proba egitean ikasleak erabilitako denbora garrantzitsua zela emaitzen interpretazioan. Ikasleak zuzentasunari eskainitako arretak, esaterako, eragina du emaitzetan. Ondorioz, hurrengo proba batean (ikus § VII.2.3.2) faktore hau landu genuen. Honek IDAZKIDE sisteman denbora-kontrola izatearen beharra ekarri zuen.
- * Badira beren esperientzian eta lehenengo faseko corpusean oinarrituta irakasleek diagnostikatu zituzten arrazoi batzuk 2. eta 3. faseetan burututako corpus-kodeketan gutxitan agertu direnak, edo azaldu ere ez direnak.

2.2.2 Ikasleentzako egitura desbideratuaren zergatiaz

Lehen esan bezala, bigarren proba batean ikasle batzuei zuzenean galdetu genien egitura linguistiko desbideratu konkretu batzuen erabilera buruzko arrazoi sakonak. Proba honen helburua zen aztertzea zenbateraino jabetzen ziren ikasleak egindako desbideratzeaz, eta hauen zergatiaz. Proba lau ikasleentzat burutu zuten. Langai bera abiapuntu zutela ikasle bakoitzak idazlan bat idatzi zuen. Ondoren irakasleak idazlanetik adibide desbideratuak atera zituen, eta ikasleei adibideak zuzentzeko eta egindako desbideratzearen arrazoia emateko eskatu zien.

Azpitarratzekoa da lau ikasleentzako kasu desbideratzeetatik erdia baino gehiago ondo zuzentzeko gai izan zirela: ikasle batek 13tik 11 zuzendu zituen, eta besteek 7tik 5, 14tik 9 eta 8tik 6. Arrazoiak dagokienez bi ez ziren gai izan zergatiak azaltzeko; beste biek, berriz, honelako arrazoiak eman zituzten: idazlana ez gainbegiratzea, azkar idaztea, esaldi konplikatuak egitea oinarritzeko egiturei ez erreparatuz, egiturak ekiditea, gaztelatik itzulpen literala egitea, zatika idaztea eta egitura ez menperatzea.

Proba honen ondorioz IDAZKIDE sistema elkarrengandik hasierako diseinuan kontuan izan ez genuen identifikatze-fasea gehitu genuen, hau da, ikasleak duen gaitasuna desbideratzearen zergatia azaltzeko. Egindako probaren arabera esan dezakegu desbideratzearen

zergatia azaltzeko gaitasuna ez dela zuzenki lotzen ikasleen ezaugarri pertsonalekin. Esaterako, ikaslearen abstrakzio-maila autozuzenketan laguntzen duen ezaugarri bat izan liteke, baina, gure frogan ezetz frogatu da. Konkretuki, lau ikasleetatik hiru abstraditzeko gaitasuna izan arren, bat ez zen gai izan arrazoiak esplizituki emateko. Beraz, ikaslearekin lan egitean uneren batean sistemak egitura desbideraturen baten arrazoiak galdetzen badizkio, erantzunaren arabera sistemak erabaki beharko luke arrazoiak lantzeko galdeketa-estrategia erabili ala ez.

2.3 Berrespena irakasleen zuzenketak direla medio

Aurreko hiru faseen emaitza gisa desbideratze linguistikoak zein hauen arrazoiak zehaztuta zeuzkaten testu kodetuak lortu genituen. Laugarren fase batean irakasle berri batek, testuak idatzitako ikasleak ezagutzen ez zituenak, desbideratzeen zuzenketak egin zituen. Zuzenketaren bidez arrazoiak lantzea izan zen esperimendu honen helburua: egindako zuzenketek hasierako irakaslearen diagnostia berretsi zuten.

Aztertutako corpora lehenengo fasean erabilitako bera izan zen. Lehen fasean eskuz aztertutako egitura desbideratuak erabili zituen irakasleak. Kontuan izan egitura desbideratu bakoitzak lehen fasean egokitutako kode linguistikoa alboan zuela. Irakasleak azterketa egin aurretik, adibideen tratamendu informatikoa egin genuen, adibide bakoitzerako egokitzapen bat zehaztuz. Automatikoki aztertzerik ez zuten adibideak bigarren irakasleak ez zituen landu, eta aztertutako egituretarako ondoko informazioa eman zuten: egituraren zuzenketa eta tratamendu automatikorako egindako moldapena egokia zen ala ez. Azkenik, bigarren irakaslearen emaitzak ikusita, zuzenketen arrazoiei buruzko elkarrizketa izan zuten irakasleak eta informatikariek. Ondorio garrantzitsuena izan zen egiturari alde aurretik jarritako kode linguistikoak eragina zuela zuzenketan. Adibidez, komuntadura-motako desbideratzeetan irakaslearen jokabidea honakoa izan zen:

- *ukan* aditza singularrean egonez gero subjektua zuzendu behar da.
- subjektua ergatibo singularrean badago aditza zuzendu behar da eta singularrean jarri, *ukan* aditza bada, bestela, ez bada *ukan* aditza, subjektua zuzendu behar da.
- beste kasu guztietan aditza da zuzentzen dena.

Beraz, desbideratzeen diagnostia egitean egiturari egokitzen zaizkion sakoneko arrazoiek irakaslearen zuzenketa bideratzen dute. Honetaz gain, elkarrizketetan konturatu ginen ikasgelan une bakoitzean lantzen ari diren egitura linguistikoek ere eragina dutela. Honegatik, IRAKAZI sisteman langai bakoitzaren objektuan langai bakoitza proposatzen den unean ikasgelan lantzen ari diren egitura linguistikoak agertzen dira.

VII.2.3 Ezagutza linguistikoaren egonkortasunaren azterketa.

VII.2.3.1 Goiko ikaskuntza-maila altutik goiko maila baxura.

Atal honetan hainbat fenomeno linguistikoren egonkortasunaren garapena landuko dugu goiko ikaskuntza-maila altuaren eta goiko maila baxuaren corpus bana konparatuz. Maila bietako corpusak 16 testuk osatzen dute hurrenez hurren. Maila altuan 4 ikaslek idatzitako testuak ditugu (5356 hitz), maila baxuan berriz 5 ikaslek idatzitakoak dira (2174 hitz). Azterketa hau automatikoki egin zen, hitz-mailako erregelak detektatu eta hauen agerpenkopuruak kalkulatuaz.

Tartehizkuntzaren ezagutza aldakorra eta finkoaren garapena aztertu dugu ebaluazio honetan, konkretuki ondoko puntuak landu ditugularik:

- a) Maila baxuan agertzen den ezagutza aldakorra, baina maila altuan inondik ere azaltzen ez dena, hau da, maila altuko ikasleengan desagertutakoa.
- b) Bi mailetan agertzen den ezagutza aldakorra.
- c) Goiko maila altuan bai, baina maila baxuan finkatuta ez dauden fenomenoak.
- d) Egonkortasunaren garapenean elkarren arteko interferentzia izaten duten fenomenoak. Zehazki, maila baxuan agertzen den ezagutza aldakorra, maila altuan inondik ere azaltzen ez dena aztertu dugu. Oraingo kasua a. kasuaren azpimultzoa da: maila baxuan ezagutza aldakor barruan kokatzen ziren eta maila altuan ezagutza finkora pasa diren fenomeno linguistikoekin erlaziorik duen ezagutza.

a) Kasu honetan *c* eta *v* letren agerpena detektatu dugu gaztelerazko maileguren erabileran. Desbideratze-mota hau maila baxuan era aldakorrean agertu zaigu, maila altuan azaldu ez den bitartean.

b) Gaztelerako mailegua *de* eta *g* eta *j* letren erabileran kasu batzuetan jatorrizko hizkia mantentzen da, eta beste batzuetan aldatzen da. Fenomeno honekin erlazionatutako desbideratzeak aldakorki agertu zaizkigu bai maila baxuan bai maila altuan.

c) Kasu honen adibidea erlatiboazko esaldiak aztertzean agertu zaigu. Aditza *-t* edo *-n* letrez bukatzen denean, singularreko lehen eta bigarren pertsonak, *-n* atzizkia gehitzean *-t/-da* eta *-n/-na* aldaketak gertatzen dira. Fenomeno estandar hau maila baxuan ez da ezagutza finkoan kokatzen, maila altuan, ordea, bai.

d) Azkenik, fenomenoaren arteko interferentzia *de* eta *g*, hau da d. kasua, hiru erregela morfologikoen arteko interferentzia aipatuko dugu adibide gisa:

- *LEGBE* erregela estandarra: -e hizki epentetikoa gehitu hitzaren lemaeren bukaeran lemaeren azken letra kontsonante denean, eta lotu beharreko atzizkiaren lehen letra -a denean (adib. oihan + an = oihanean).
- *e+e=ee* erregela estandarra: hitz bat deklinatzean lemaeren azken letra -e denean, eta lotu beharreko atzizkiaren lehen letra -e denean, ez dago ezabaketarik (adib. *etxe* + *ekin* = *etxeekin*).
- **LEGTE* desbideratzea: -e letra epentetikoa gehitu hitzaren lemaeren bukaeran lemaeren azken letra kontsonante denean, eta lotu beharreko atzizkiaren lehen letra -e denean (adib. *euskaldun* + *ekin* = *euskalduneekin*).

Goiko maila altuko eta baxuko ezagutzaren garapenean aipatutako fenomenoen interferentziari dagokionez ondokoa detektatu dugu:

- Maila baxuan ikasleek bai *oihanean* bai **oihanan* bezalako adibideak erabiltzen dituzte. Maila horretako ikasleek *etxeekin*, **etxeekin*, **euskalduneekin* eta *euskaldunekin* bezalako adibideak ere idazten dituzte. Ondorioz, aipatutako hiru erregelak ezagutza aldakorrean kokatzen dira.
- Maila altua aztertzean, ordea, *oihanean*, *etxeekin* eta *euskaldunekin* bezalako adibideak bakarrik agertzen zaizkigu. Maila altuan **LEGTE* erregela desagertu da aurreko biak ezagutza aldakorretik finkora pasa direnean.

VII.2.3.2 Ariketa-motaren eragina.

Irakaslearen iritzia beti ez datorrela bat errealitatearekin § VII.2.2-n ikusi dugu, errealitatearen froga gisa corpusen azterketa hartzen badugu. Ikasleen idatzizko jardueran batzuetan agertzen da irakasleak aipatzen ez duen hainbat fenomeno, eta, beste batzuetan, aipatzen dituenak ez dira agertzen. Ondorioz, bigarren hizkuntza ikasteko ingurunean IRAKAZI eta HITES sistemen arteko elkarrekiko komunikazioa beharrezkotzat jo dugu. Irakasleen iritzia eta corpusen azterketen emaitzen konparaketa eginez psikolinguistik ikaslearen tartehizkuntza fintzen du. Baina nahikoa al da aipatutako bi sistemen arteko elkarrekintza? Ezin al da egon fenomeno fosilizaturen bat ikaslearen tartehizkuntzan aspaldian ikasia, irakasleak ez detektatua, eta corpusetan oso gutxitan agertzen dena?. Galdera honetatik abiatuta 10. mailako corpus batean oso gutxitan agertzen zen hainbat fenomeno planteatu genion irakasle bati. Horietako batzuk aukeratu genituen, bere ustetan ere gertatzen ez zirenak, eta fenomeno horiek aztertzeko froga-azterketa berezia prestatu genien 10. mailako ikasle batzuei.

10. mailako 41 testuz osatutako corpora (10311 hitz eta 866 esaldi) aztertuta corpusetan gutxitan agertzen zen, eta irakaslearen ustez 10. mailan ematen ez den zenbait desbideratze aukeratu genuen. Alde batetik, irakaslearen iritziz beste ikaskuntza-maila batzuetan elkarrekiko interferentziak izaten dituzten fenomeno batzuk hautatu genituen (hots,

*LEGTE, *LEKTE eta *LEGBA desbideratzeak, eta LEOEAE eta LEGBE erregela estandarrik), eta, beste alde batetik, hitz bakar baten barruan aldi berean suertatzen diren batzuk ere aztergai izan genituen (*LEOESZ eta *LEOEZS).

Ikasleen ezaugarriei dagokienez, froga honetan ondoko ezaugarri pertsonalei erreparatu genien gehienbat:

- Ikasleak idaztean zuzentasunari begiratzen dio?
- Desbideratzeak sahiesteko ziurtasunik ez duen kasuetan arriskurik ez hartzen saiatzen al da?
- Idazlanak maiz egiten al ditu? luzeak, motzak?

Ezaugarriak direla eta, froga egin zuten sei ikasleren artean hiru aukeratu eta beste hiru baztertu egin genituen. Azken hauek ondoko arrazoiengatik: bi, zuten hizkuntz maila ez zelako taldearen adierazgarri, batek maila baxuegia zuen, eta bestea azkarregia zen; eta hirugarren ikaslea ez genuen aztertu urtean zehar eginak zituen idazlanak ez genituelako eskuratu.

Beraz, corpusetan agertzen ez ziren eta irakaslearen ustez gertatzen ez ziren hainbat fenomenoren egonkortasuna landu nahi izan genuen proba honen bidez, fenomeno sistematiko edo finkoak, aldakorrak edota lapsusak ote diren identifikatuz. Horretaz gain, fenomenoak ikaslearentzat zuzengarriak edo zuzentzezinak diren ere detektatu nahi izan genuen fenomeno bera aztertzeko ariketa-mota desberdinak proposatuz (hots, hutsuneak bete, okerrak zuzen eta itzulpena). Eta, azkenik, ikaslearen ezaugarrien eragina ere ondorioztatu genuen.

Proba prestatu aurretik aukeratutako erregelak erabiltzen dituen hainbat hitz hautatu genuen, hitzaren aldaera batzuekin batera, eta hitz zein aldaera horien erabilera landu genuen. Adibide batzuk itzulpenean bakarrik aztertu genituen, idazlanen eta itzulpenen arteko konparaketa egiteko asmoz. Beste adibide batzuk itzulpenean eta okerrak zuzen edo hutsuneak bete izeneko ariketetan kokatu genituen, itzulpenaren eta ariketa gidatuen (okerrak zuzen edo hutsuneak bete) arteko diferentzia aztertzeko. Aipatu behar da ere proban ariketa bakoitzerako denbora mugatua eman zitzaiela ikasleei. Gainera, ariketaren helburuaz ez jabetzea nahi genuen. Horretarako, ariketei ordena jarri genien: lehenengo, itzulpen ariketa egin zuten; ondoren, okerrak zuzen ariketa, eta, azkenik, hutsuneak bete. Froga egin ondoren, ariketak egitean zuzentasun gramatikalean arreta berezia jarri ote zuten galdetu genien; horretaz gain, itzulpenean esanahian arreta jarri ote zuten ala ez; eta, azkenik, beren ustetan zeintzuk ziren ariketak prestatzean guk geuk aurreikusitako desbideratze-motak.

Frogaren ondorioak.

Froga honetatik hainbat ondorio atera genuen ariketa-motaren eraginari dagokionez. Horretaz gain, irakasleen aurreikuspenen erlatibotasuna ere ondorioztatu zen proban. Ondorioak aipatuko ditugu banan banan:

- Idazlanez osatutako corpora gehi irakaslearen iritzia ez dira nahikoak tartehizkuntza lantzeko. Emaitzak ikusitakoan irakasleak erakutsi zuen harridura azpimarratzekoa da. Honek frogatzen du irakaskuntza-programek askotan ikaskuntza-maila batean hainbat fenomeno barneratu dela suposatzen duten arren, ezagutza hori errealitatean oraindik ikaslearen ezagutza aldakorreen mantentzen dela askotan.
- Ariketa egiteko ikasleak duen jokaerak (zuzentasunari begiratzea edo idaztean arriskuak hartzea) eragina handia du egindako desbideratzeetan.
- Ikaslearen partehartze esplizitu zein implizitua (ariketen bidez) ezinbestekoa da tartehizkuntzaren detekzio-prozesuan.
- Fenomeno konkretuak lantzeko ariketa bereziek idazlanetan agertzen ez den informazioa ematen dute. Hau dela eta, ikaslearekin lan egingo duen IDAZKIDE sisteman ariketa-moten diseinu zehatza beharrezkoa da.
- Corpusetan agertzen diren fenomenoaren aldakortasuna aztertzeke okerrak zuzen eta hutsuneak bete bezalako ariketak automatikoki sor daitezke, aldeztatik corpusaren adibide konkretu batentzako aldaera desberdinen sorkuntza egiten bada.

VII.3 Ingurune informatikoaren emaitzak.

Wolfe-Quintero et al.ek (1998) komunikazio idatzian hizkuntzen ikaskuntza arloan hizkuntza desberdinetarako egin diren ehunetik gorako neurketa-azterketak lantzen dituzte barne-txosten batean. Azterketa horiek komunean dute aztertutako testuak komunikazio-egoera bat adierazten dutela, hau da, testu libreak dira eta ez dira sortu egitura linguistiko konkretuak lantzeko ariketa batzuetatik. Ikerketa horietan ondoko neurriak aztertzen dira: ikasleen jario-tasuna, zuzentasuna eta konplexutasuna. *Jario-tasunean* testuen aberastasuna lantzen da denbora kontuan izanik: honi buruzko datu batzuk lortu ditugu § VII.3.3.2-n hiztegiaren aberastasuna lantzean. *Zuzentasunean* errore-kopurua neurtzen dute: tesi honen IV. kapituluan azaldu genuenez errorean (gureztat desbideratzean) detekzioa lortzeko xedero-korreko tresna linguistiko-konputazionalak aurkeztu ditugu. Azkenik, *konplexutasunean* hitzen zein esaldien konplexutasun gramatikala aztertzen dute: honi buruzdatu batzuk ematen ditugu § VII.3.3.3-n eta § VII.3.3.4-n. Txostenaren bukaeran, Wolfe-Quintero et al.ek (1998) aipatutako neurketak hobetzeko ratioak definitzen dituzte. Ratio horien baliagarritasuna lantzea izan daiteke lan honetan proposatu dugun MUGARRI izeneko ingurune informatikoaren erabileretako bat.

Gure tesian, alde batetik, ikasleen tartehizkuntza errepresentatzeko oinarri diren corpusak aipatu ditugun ikerketetan erabilitakoen modukoak dira, hau da, egoera komunikatibo bat islatzen dute; eta, beste aldetik, ikerketa horietan egindako neurketak erdiautomatikoki egiteko ingurune informatikoa proposatzen dugu. Horregatik, gure ingurune informatikoa, hau da, MUGARRI baliagarria dela frogatzeko asmoz, saiakera bat egin dugu komentatutako neurketak lantzeko. Dena den, gure esperimenduak jariotasuna eta konplexutasunera murriztu ditugu.

Atal honetan, § VII.3.1-n erabilitako corpusak aipatuko ditugu, § VII.3.2-n corpus horien azterketa aurkeztuko dugu. Gero, § VII.3.3-n, hainbat ikasleren tartehizkuntzari buruzko informazioa automatikoki lortuko dugu inplementatutako ingurune informatikoa erabiliz, ondoko alderdiak ikasleengan aztertzeko: jariotasunari dagokionez, hiztegiaren aberastasuna eta hainbat kategoria gramatikalen erabilera; eta, konplexutasunari dagokionez, hitz-mailako fenomeno morfosintaktikoak eta esaldien konplexutasuna.

VII.3.1 Erabilitako corpusak.

III. kapituluaren zein kapitulu honen aurreko ataletan azaldutako lan enpirikoak oinarri izanik bigarren hizkuntza ikasten duten ikasleen tartehizkuntza lantzeko MUGARRI izeneko ingurune informatikoa garatu dugu. Ingurunea IV. kapituluaren proposatutako goitik beherako metodologian oinarrituta dago. Goiko hizkuntz mailetatik hasi, eredu kolektiboak atera, eta hauetatik abiatuta beheko hizkuntz maileri dagozkien eredu kolektiboak lantzen dira ingurunean.

Gure hurbilpenean ikasle baten output-a bere tartehizkuntzaren adierazle da. Honen arabera, adibidez, hitz baten kasuan hitzaren interpretazio posible guztiak baliagarriak kontsideratzen dira. Hortaz, hitz baten interpretazioen artean estandarra egoteak ez du esan nahi estandarrak ez direnak baztertu behar direnik – zuzentzaile ortografikoetan kontsideratzen den moduan–.

Oinarri hau berresteko, eta HITES sistema baliagarria dela frogatzeko asmoz, saiakera bat egin dugu tesi honetan garatu diren sistemak nola erabil litezkeen erakusteko. Saiakera honetan erabili ditugun corpusak honako hauek dira:

Corpusaren izena	Ikaskuntza-maila/k	Testu-kopurua	Hitz-kopurua
-------------------------	---------------------------	----------------------	---------------------

Goiko_maila	9.etik 12.era	137	35018
Erdiko_maila	6.etik 8.era	151	34402
Beheko_maila	1.tik 5.era	231	36502
Ariketa_corpusa	10.a	30	7785
Erd-B/Goi-B	8.a eta 9.a	16	3733
Beh_b-A/Beh_g-A	2.a eta 5.a	16	2562
Hasierako_corpusa	10.a	16	4528
Corpus_berritua	10.a	24	6355
Ikas_proz_corpusa	8.a	4	1097

VII.1 taula. Inguruneko probetan erabilitako corpusak.

VII.1 taulan azaldutako corpusak ondoko bi ataletan agertuko dira ikasleen tartehizkuntzaren alderdi desberdinak aztertzerakoan. Taulan corpusei eman zaizkien izenak agertuko dira, bai irudietan bai diskurtsoan zehar.

VII.3.2 Aplikatutako metodologiaren berrespena.

Ikaskuntza-maila baten fenomeno linguistikoen identifikazio erdiautomatikoan HITES sistemaren prototipoa erabili dugu probak egitean.

Goitik beherako metodologiaren arabera ikaskuntza-maila gorena abiapuntu izanik, hurrengo ikaskuntza-mailaren tartehizkuntza lantzen dugu. Horretaz gain, output-a interpretatzean desbideratzeak egitea zilegitzat jotzen da, beraz, adibidez, hitz baten interpretazioen artean interpretazio estandarra egoteak ez du esan nahi estandarrek ez direnak baztertu behar direnik. MORFEUS (Agirre et al., 92) analizatzaile morfologikoaren kasuan, berriz, interpretazio estandarra dagoen kasuetan desbideratzeak diren interpretazioak baztertu egiten dira. Ondorioz, alde batetik, analisi-kopurua hitzeko HITESen baino txikiagoa da, baina, beste aldetik, zuzenak izan daitezkeen interpretazio desbideratedunak ez ditu kontuan hartzen. HITESi egin dakioken kritika gogorrena gainsorkuntza dugu, interpretazio estandarrek dituzten hitzen kasuan desbideratzeak ere onartzen direlako. Gainsorkuntza hau, ordea, ez da oso altua ondoren frogatuko dugun bezala.

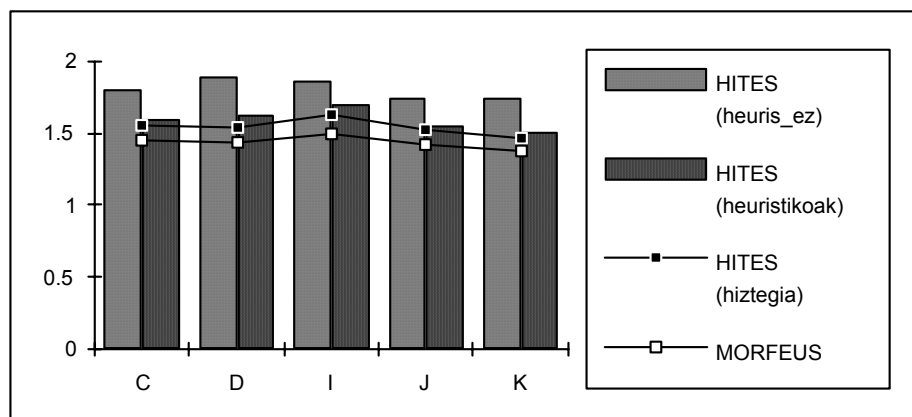
Gainsorkuntza dela eta, hainbat corpus hartu eta hitzen analisiak aztertu ditugu bai HITES sistema erabiliz bai MORFEUS erabiliz. Konparaketa egitean fenomeno berdinak tratatzen dituzten bertsioak landu ditugu, hau da, batek zein besteak tratatzen dituen desbideratze-motak berdinak dira.

HITES sisteman hainbat corpusekin probak egin ondoren, hitz bat analizatzean lortzen den interpretazio-kopurua jaisten duten heuristiko batzuk definitu ditugu. Aplikatutako heuristikoak ondokoak dira:

1. Hiru desbideratze edo gehiago batera detektatzen direnean interpretazioa baztertuko da.
2. Hitzaren luzera bost hizki baino gutxiagokoa denean, eta gainera bi desbideratze batera aplikatzen badira interpretazioa baztertuko da.
3. Hitzaren luzera bost hizki baino handiagokoa denean badira desbideratze konkretu batzuk ezingo direnak batera gertatu. Hala denean interpretazioa baztertuko da.

Heuristikoak finkatzeaz gain hizkuntz mailei dagokien hiztegia ere landu dugu. Honen ondorioz ere interpretazio-kopurua jaistea lortu dugu. Behin heuristikoak aplikatuta eta ikaskuntza-mailaren hiztegia finduta VII.3 irudian azaltzen diren emaitzak lortu ditugu goiko ikaskuntza-mailari dagozkion bost corpus (bakoitza 1500 hitz ingurukoa) landu ondoren.

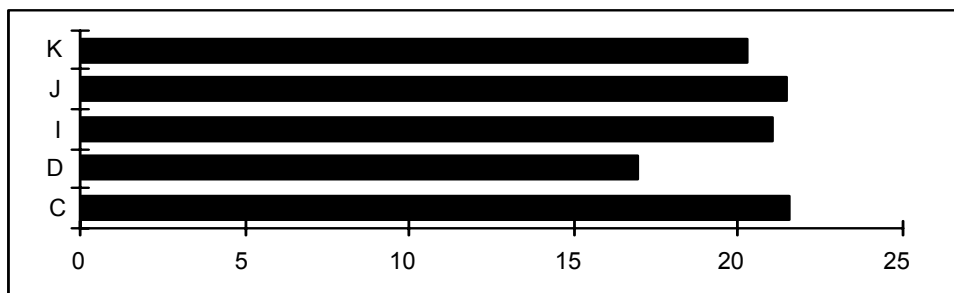
Irudian ikusten denez bost ariketa-mota (C, D, I, J, K) landu dira *ariketa_corpusa* izeneko proba honetan. Seina testu hartu eta kasu bakoitzeko hitz baten analisisen tasa landu dugu. Lehen zutabeen HITESen tasa gordina azaltzen da. Bigarren zutabeen, berriz, heuristikoak aplikatuta lortu diren emaitzak ikusten dira. Horretaz gain, corpusaren hamar testu aztertu ondoren findutako hiztegia hartuta hirugarren emaitza dugu. Azkenik, MORFEUSEkin lortutako emaitzak ematen ditugu.



VII.3 irudia. Hitzen analisisen tasa goi-mailako ikasleen corpusetan.

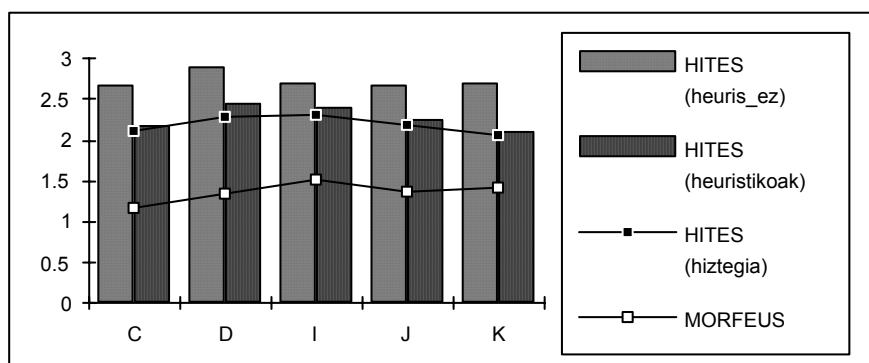
Analisien emaitzak konparatzen baditugu MORFEUS analizatzailea (murrizpen gramatikaren desanbiguazio morfosintaktikoa ere aplikatuta) eta HITES (desanbiguazio-erregelak aplikatuta eta hiztegia findu ondoren) sistemaren arteko diferentzia 0.1ekoa da, diferentzia onargarria eta tratagarria. HITES sisteman psikolinguistikak desanbiguatu beharko

lituzkeen hitzen kopurua ez da altua (ikus VII.4 irudia) – irudian agertzen diren datuak ehunekotan ematen dira corpusen hitzen kopurua kontuan izanik –. Analizatutako testuak aztertu ondoren, fenomeno linguistiko asko automatikoki ondorioztatuko ditu sistemak, eta gutxi batzuk landu beharko ditu psikolinguistak, bai hiztegia finduz, bai erregela morfosintaktikoak ikaskuntza-mailarako egokituz.



VII.4 irudia. Corpuseko hitz forma desberdinak, analisisian anbiguo direnak (%).

VII.5 irudian ikusiko dugunez, analisisetan aplikatutako erregelak direla eta, HITES sistemak MORFEUS analizatzaileak baino erregela morfofonologiko gehiago esplizituki detektatzen dituenaz, emaitzetan erregela horiei dagozkien analisiak hitzeko nabarmenki igotzen dira. MORFEUSEk desbideratzeak besterik ez du detektatzen erregelaren artean, eta HITESek desbideratzeaz gain erregela estandarrek ere detektatu eta esplizituki azaltzen ditu.



VII.5 irudia. Hitzen analisisien tasa erregeladun analisisetan.

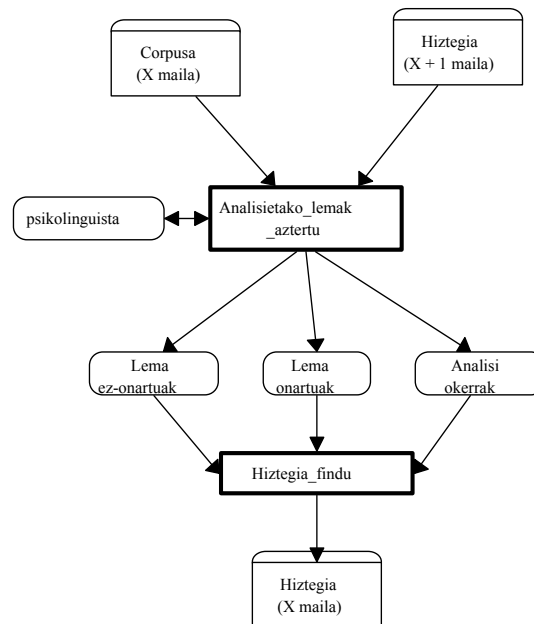
Egindako konparaketa osatua izan dadin analisi oker eta zuzenen arteko konparaketa zehatza egin beharko genuke; hau da, psikolinguista batek eskuz analizatu eta ikusi ea zein kasutan interpretazio zuzenik ez den asmatu. Eskuz kodetutako corpusen falta dela eta, VII.5 irudian agertzen diren ariketei zegozkien testu batzuekin psikolinguistak proba egin eta analisi okerrak aztertu zituen. Kasu gehienetan errorearen jatorria hiztegia finduta ez izatea zen, beste kasu batzuetan, ordea, erregela morfosintaktikoa egokitu beharra zegoen aztertzen ari ginen ikaskuntza-mailarako. Ondoren, hiztegiaren finketari buruz aritzean, konparaketa hau zehatzago azalduko dugu.

Azkenik, aipatu behar da, HITES sisteman psikolinguistak analisiak desanbiguatzeko dituenen, ez dituela hitz anbiguo guztiak aztertzen, baizik eta anbiguoak diren hitz forma desberdinak. Konparatzen baditugu hitz anbiguoak eta anbiguoak diren hitz forma desberdinak kasu askotan diferentzia nabarmena da; esaterako, corpusaren D ariketan 1890 hitzetatik 572 hitz anbiguo egon arren hitz forma desberdinak 321 dira, hauek dira psikolinguistak aztertu behar dituenak.

Goitik beherako metodologiaren aplikazioa hiztegiaren finketan.

Goitik beherako metodologiaren arabera ikaskuntza-maila gorenaren hiztegia gure abiapuntu izanik, aurreko (X) ikaskuntza-mailaren hiztegia zehazten dugu goiko (X+1) ikaskuntza-mailaren hiztegiaren azpimultzoa mugatuz. VII.6 irudian ikusten denez X ikaskuntza-mailaren testuak emanda, eta hurrengo maila (X+1) goikoaren hiztegia abiapuntu hartuta, testuak analizatu egiten dira. Ondoren, hitz bakoitzeko lortutako analisi guztiak kontuan izanik X ikaskuntza-mailaren hiztegia zehazten da ondoko urratsak emanez:

- 1.- Hitzari dagozkion lemak ezagunak edo ezezagunak diren esatea
- 2.- Hitzaren interpretazioen artean interpretazio zuzena falta den (analisi okerrak) markatzea.



VII.6 irudia. Hiztegiaren finketa goitik beherako metodologiaren aplikazioan.

Egindako probetan goiko mailaren hiztegia findu dugu, horretarako corpusaren hitzen analisiak (10 testu eta 2868 hitz) aztertu ditugu, ariketa-mota irizpide izanik. Ariketa-mota diogunean testuaren gaia eta testu-egitura kontsideratzen ari gara. Bost ariketa desberdinetan

idatzitako testuak analizatu eta aztertu egin ditugu testu luzeenetatik hasita. Ariketa-mota bakoitzeko bina testu landu dugu.

Ikaskuntza-mailari zegokion lehen proba hiztegia findurik izan gabe egin dugu. Ariketa bakoitzeko testu bana aztertu ondoren analisi okerren portzentajea % 2.27koa atera da batez beste. Ondoren hiztegia findu, hau da, ikasleentzat ikaskuntza-maila horretan ezezagunak ziren hitzak hiztegitik kendu eta berriro testu bana analizatu dugu ariketa bakoitzeko. Bigarren proban aztertutako testuak ikasle berri batzuek idatziak ziren. Emaitzak eskuz aztertu ondoren analisi okerren kopurua nabarmenki jaitsi da: % 0.62koa kasu honetan. Beraz, argi dago hiztegia fintzeak eragin zuzena duela emaitzetan. Emaitza hauetatik pentsa liteke ariketa-motak eragina duela hiztegia fintzen dugunean. Neurri batean hau egia da hainbat hitz gaiarekin lotuta egoten direlako; baina ez nahitaez. Hurrengo proba batean suertatu zaigu I¹ ariketa-motarentzat testu berria hautatzean bigarren proban analizatutako testuan baino analisi oker gehiago detektatu direla, nahiz eta bigarren probako testuaren hiztegia findu ondoren analizatu.

Ondorioa: hortaz, ariketa-motaren arabera egindako finketaren ondorioz topatu dugu analisi okerren kopurua ez dela beti jaisten hiztegiaren finketa aurrera joan ahala proba egin aurretik suposatu genuen bezala. Arrazoiatariko bat da testu desberdinek testu-gaia bera izan arren ikasleek gaiarekin zerikusirik ez duen hiztegi zabala ere erabiltzen dutela. Beraz, HITES sisteman psikolinguistak ikaslearen hiztegia fintzen duenean ariketa-mota ez da irizpide izango prozesua bideratzeko.

Ariketa-mota irizpide ez izan arren, analisi okerrek bai bideratzen dute HITES eta psikolinguistaren arteko elkarrekintza. Hau dela eta, elkarrekintza horretan ondoko irizpide ondorioztatu dira:

- Analisi oker baten lema guztiak ezagunak direnean HITESek erregela morfosintaktiko konkretu batzuk eguneratzeko eskatuko dio psikolinguistari.
- Oker analizatutako hitz baten interpretazioen artean baten bat ezezaguna denean, ezezagun gisa markatu eta hitza berriro analizatuko da. Honelako kasuetan egindako proban, ia kasu gehienetan analisi zuzena agertu da interpretazioen artean finketaren ondorengo analisisian.

Horren ondorioz, esan dezakegu hiztegiaren finketa oso garrantzitsua dela, hiztegiak eragin handia duelako analisisien emaitzetan. Beraz, lehenengo hiztegia findu, eta, ondoren, erregela morfosintaktikoak eguneratuko dira goitik beherako metodologiaren aplikazioan.

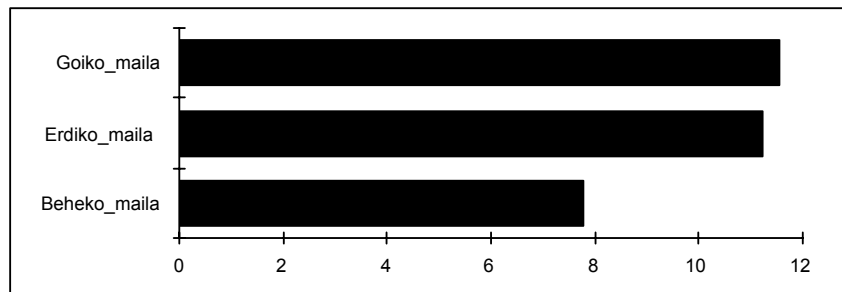
¹ Liburu bateko zati baten laburpena

VII.3.3 Tartehizkuntzaren azterketa automatikoa.

Tartehizkuntz egituren azterketari dagokionez egindako probetan ikaskuntza-mailak ez du eragin handirik erakutsi hitz-mailako analisisien kopuruetan. Corpus handiak lantzeko asmoz hiru ikaskuntza-maila nagusi desberdindu ditugu:

- Goiko maila (9. mailatik 12. mailara).
- Erdiko maila (6. mailatik 8. mailara).
- Beheko maila (1. mailatik 5. mailara).

VII.7 irudian hiru ikaskuntza-maila nagusien corpusetan anbiguo diren hitz forma desberdinen portzentajea ematen da: goiko ikaskuntza-mailaren 137 testu (35018 hitz), erdiko mailaren 151 testu (34402 hitz) eta beheko mailaren 231 testu (36502 hitz). Corpus hauen hitzeko analisi-tasa – 1.6, 1.58 eta 1.58 hurrenez hurren – antzekoa da maila guztietan; corpusetan anbiguo diren hitz forma desberdinen portzentajea, berriz, beheko mailan diferente da erdiko zein goiko mailekin konparatuz gero: beheko mailan aztertutako corpusean anbiguo diren % 7.79 hitz forma desberdin topatu dira, erdiko mailan %11.26, eta goiko mailan %11.57. Diferentzia honen azterketa interesgarria gertatu zaigu tartehizkuntzaren detekzio-prozesuan, ondoren ikusiko dugun bezala.



VII.7 irudia. Hiru ikaskuntza-maila nagusietako corpusen anbiguotasuna (%).

Tartehizkuntza dela eta hainbat faktore landu dugu testuetan:

- Fenomeno morfosintaktikoen detekzio erdiautomatikoa.
- Erabilitako hiztegiaren aberastasuna.
- Testuetako hitzen kategoria gramatikalak.
- Esaldien konplexutasuna.

VII.3.3.1 Fenomeno morfosintaktikoen identifikazio erdiautomatikoa.

Esan bezala, gure metodologian ikaskuntza-maila baten tartehizkuntza lantzean goiko mailaren ezagutza hartzen dugu abiapuntu. Horretaz gain, ikaskuntza-maila gorena aztertzekeo hitzun adituen ezagutza erabiltzen da oinarri bezala, hau da, gure kasuan

MORFEUSen ezagutza. Beraz, fenomeno morfosintaktikoak automatikoki detektatzeko MORFEUS eta HITESen arteko konparaketa egiten dugu hasiera batean, – gogoratu HITESek ez dituela interpretazio ez-estandarrek baztertzen interpretazio estandarra dagoenean –. Konparaketa honetatik abiatuta sistemak analisi desberdinak aurkezten dizkio psikolinguistari.

Corpus batzuetan konparaketak egin ondoren hainbat analisi desberdin topatu dugu, kasu batzuetan MORFEUSen zeudenak eta HITESek baztertuak, eta beste batzuetan alderantziz. Behin analisi desberdinak aztertuta, sistemak fenomeno morfosintaktikoak osatzen ditu automatikoki analisisietan aktibatutako erregelen konbinaketak eginez. Ondoren, psikolinguistak fenomeno horien baliagarritasuna lantzen du.

Sistemak automatikoki detektatzen dituen fenomeno morfosintaktikoak ondokoak dira:

- a) Erregela morfologikoak.
- b) Erregela morfosintaktikoak interpretazio-mailan.
- c) Erregela morfologikoak eta interpretazio-mailako erregela morfosintaktikoak konbinatuta.
- d) Erregela morfosintaktikoak hitz-mailan.

Identifikatutako fenomeno gehienek azalpen zehatza dute psikolinguistak dioenez, d. atalekoak kenduta. Azken hauek murriztapen gramatikaren *Select*¹ edo *Remove*² izeneko erregelak dira, eta hauen artean *Select* moduko erregelai bakarrik aurkitu diegu esanahia. Honek esan nahi du jatorrizko gramatikan *Remove* erregelak erabiliz adierazitako ezagutza ez dela baliagarria ikasleen ezagutza linguistikoa adierazteko.

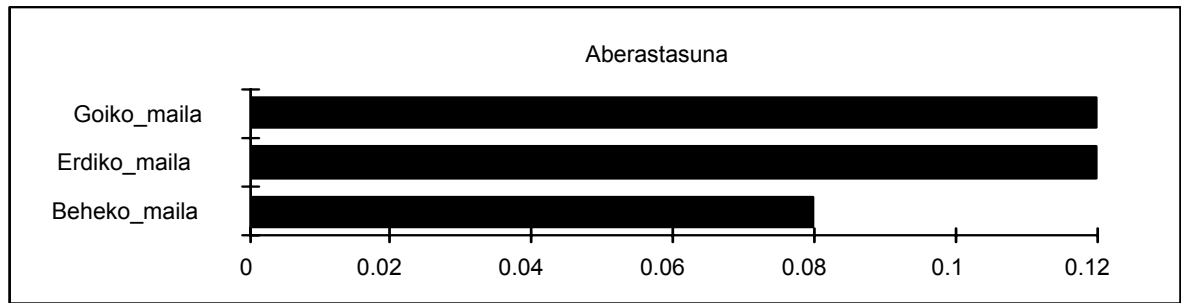
VII.3.3.2 Hiztegiaren aberastasuna.

Hiztegiaren aberastasuna testu batean lema desberdinen kopurua zati testuaren hitz-kopurua dela definitu dugu. Beraz, hitz bakoitzeko dagokion lema lortu dugu. Dena den, hainbat kasutan hitz baten analisisien artean lema bat baino gehiago topatzen dugu; kasu hauetan ausazko heuristikoa erabili dugu: *hitz batek lema bat baino gehiago dituenean lehen lema hautatuko dugu*.

VII.7 irudian aipatu ditugun hiru ikaskuntza-maila nagusien corpusak erabiliz, hauen hiztegiaren aberastasuna landu, eta ondoko emaitzak lortu ditugu:

¹SELECT (GAIA) if (BAL): murrizpen gramatikaren erregela; hitz baten analisisaren interpretazioen artean GAIA ezaugarria dutenak aukeratu egiten ditu BAL baldintzak betetzen badira.

²REMOVE (GAIA) if (BAL): murrizpen gramatikaren erregela; hitz baten analisisaren interpretazioen artean GAIA ezaugarria dutenak ezabatu egiten ditu BAL baldintzak betetzen badira.



VII.8 irudia. Hiru ikaskuntza-maila nagusietako corpusetan erabilitako hiztegiaren aberastasuna.

VII.8 irudiak bi datu garrantzitsu ematen dizkigu:

- Beheko mailan aberastasuna txikiagoa da beste bietan baino, espero zen bezala.
- Goiko eta erdiko mailaren arteko aberastasuna antzekoa da, espero ez zen bezala.

Beheko mailan aberastasuna txikiagoa izatearen arrazoia, ez dago ikasleek menperatzen duten hiztegi murritzean bakarrik, baizik eta sintaxian ere. Beheko mailan erabiltzen diren esaldietan ez dira menpekoak agertzen, koordinatuak, ordea, bai. Hau da, *eta*, *baina* bezalako juntagailuak askotan erabiltzen dituzte. Honelakoek corpusaren hitz anbiguen portzentajea zein aberastasuna jaisten dute.

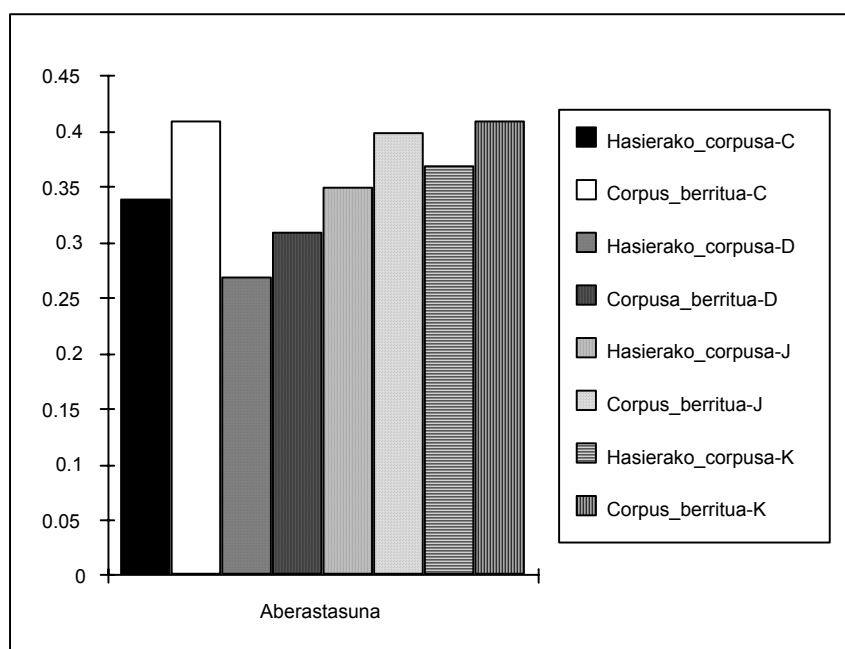
Esan bezala, erdiko mailan dagoen aberastasuna eta goiko mailakoa antzekoa atera da. Hau dela eta, ariketa-mota bera hartuta goiko mailako eta erdiko mailako zortzina ikaslek idatzitako testu bana aukeratu genituen. Emaitzek erdiko mailan aberastasun altuagoa adierazi zuten: erdiko mailako zortzi ikaslek idatzitako zortzi testuen (1271 hitz) aberastasuna %33koa izan zen; goiko mailako zortzi ikaslek idatzitako zortzi testuen (1913 hitz) aberastasuna, berriz, %20koa izan zen. Beraz, honen arabera erdiko mailan dauden ikasleen hiztegia aberatsagoa dela esan daiteke: honek datu interesgarria ematen digu: VII.8 irudian erdiko mailan zein goiko mailan dauden ikasleen hiztegiaren aberastasuna antzekoa dela esan beharrean emaitza beste modu batean ere irakur liteke: printzipioz goiko mailan dauden ikasleen ezagutza kolektiboa hiztegiari dagokionez erdikoena baino zabalagoa izan beharko lukeenez, emaitzetatik ondorioztatzen da erdiko mailan dauden ikasleen ezagutza partikularra zabalagoa dela.

Ondorioa: honek adierazten digu corpus handien emaitzetan askotan datu interesgarriak ezkututzen direla.

Aberastasunaren azterketan corpusak osatzean erabilitako populazioa eta testuinguruaren eragina ere sumatu da; hau da, zenbat ikasle desberdinek idatzi dituzten testuak, testuetan azaltzen diren gaiak eta testu-kopuruak eragina eduki dezake honelako azterketa batean. Faktore hauek lantzeko bi corpus desberdinen aberastasuna konparatu dugu (ikus VII.9 irudia):

- *Hasierako_corpusa* (16 testu): goiko mailan dauden lau ikasle desberdinek idatzitako launa testu (lau ariketa-mota).
- *Corpus_berritua* (24 testu): *Hasierako_corpusa* corpusean ariketa-mota bakoitzari bina testu gehitu zaio; C eta D ariketetan *i* eta *s* ikasleek idatzitako testuak gehitu ditugu; eta J eta K ariketetan *a* eta *s* ikasleek idatzitako testuak erantsi ditugu.

Ondorioa: bi corpusen aberastasuna konparatzen badugu argi ikusten da ariketa guztietan bi testu gehitzeak nabarmenki igotzen duela aberastasuna. Testuaren gaiak, ordea, ez dirudi eragin handirik duenik, bestela aberastasuna proportzio oso desberdinetan igo beharko zukeelako. Ikasle-motari dagokionez irudi honek ez du eragin nabarmenik erakusten. Hala ere, ondoren azalduko dugun beste corpus-konparaketa batean ikasle-motak aberastasunean eduki dezakeen eragina frogatuko dugu.



VII.9 irudia. Goiko mailako bi corpusen hiztegi-aberastasuna.

Ikaslearen ezaugarrien eragina aberastasunaren azterketan.

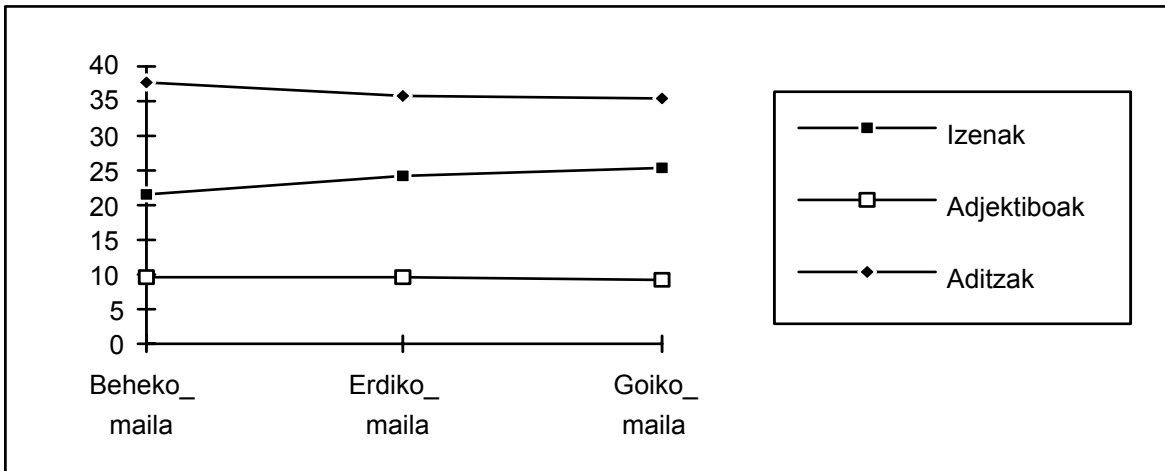
Aurreko adibidean ikusi dugunez corpusaren tamainak eragina izaten du aberastasunaren azterketan: ariketa bakoitzeko bi testu gehituta corpusean erabilitako lema desberdinen proportzioa igo da kasu guztietan. Hau dela eta, ikasle-motak eduki dezakeen eragina aztertzeko, hitz-kopuru desberdina duten bi corpus txiki hartu ditugu, suposatuz hitz gehien dituen corpusaren kasuan aberastasuna altuagoa izan beharko lukeela. Emaitzek, ordea, aurkakoa adierazten dute, testuak idatzi dituzten ikasleen ezaugarri batek eragin zuzena izan duelako. Ondoren egindako proba azalduko dugu:

Ariketa-mota bera izanik, erdiko maila altuan dauden lau ikaslek testu bana idatzi dute, guk azter ditzagun: *ikas_proz_corpusa* izeneko corpusa dugu hau. Lau ikasleetako bik ikaste-prozesu oso azkarra eraman dute gaur egun duten maila lortzeko, konkretuki hiru urtetan maila horretara heldu dira. Beste biek, berriz, hamar urte lehenago hasi ziren hizkuntza ikasten. Lehenengo kasuan idatzitako bi testuen (412 hitz) aberastasuna aztertuta corpusaren lema desberdinen portzentajea % 44koa detektatu da. Bigarren kasuan, hau da, motelago ikasi dutenen bi testuak (685 hitz) aztertzean aberastasuna %34koa izan da. Bigarren kasu honetan hitz-kopurua altuagoa izanik aberastasun handiagoa espero genuen arren, probak adierazten du urte gutxitan ikasi dutenek hiztegi aberatsagoa dutela: agian, ikasitakoa freskoagoa dutelako. Kontuan izan behar dugu maila lortzeko hamar urte pasa dituzten bi ikaslek eten handiak izan dituztela ikaste-prozesuan zehar. Hiru urtetan erdiko maila altua lortu dutenek, berriz, ikaste-prozesu jarraitua eta trinkoa izan dute.

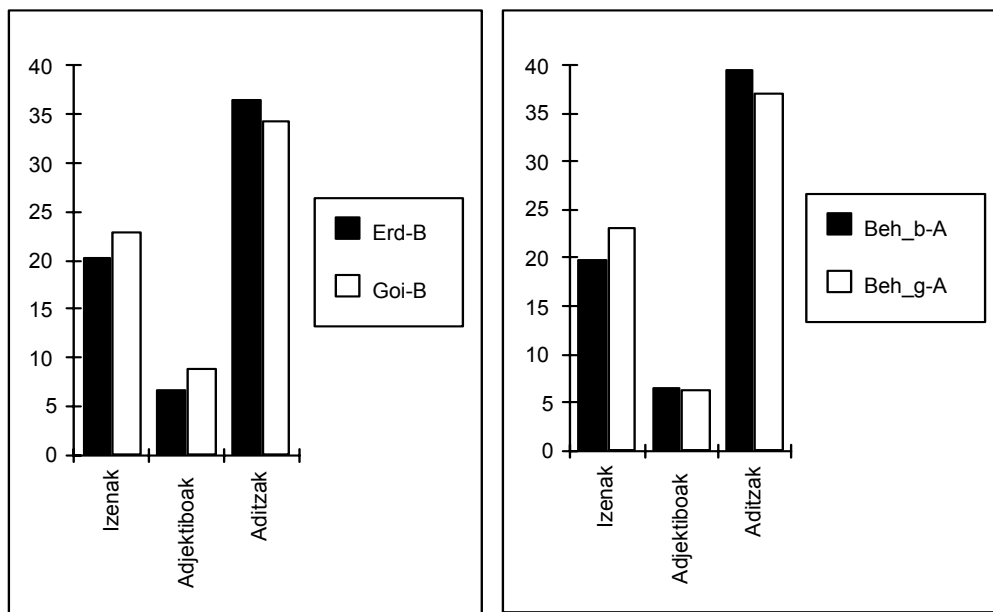
Ondorioa: esan dezakegu ikasle baten ezaugarri batzuek eragina dutela testuen aberastasunean. Azaldutako probak adierazten du ikasleak ikaste-prozesuan izan duen garapenak eragina duela hiztegiaren aberastasunean. Beste batzuek, ordea, ez dute inolako eraginik erakutsi.

VII.3.3.3 **Kategoria gramatikalen erabilera.**

Kategoria gramatikalei dagokienez ondokoak aztertu ditugu: izenak, adjektiboak eta aditzak. VII.10 irudian kategoria hauen erabilera aztertzen da VII.7 irudian azaldutako corpus handietan. VII.11 irudian, berriz, bi ariketa-mota hartuta ikaskuntza-maila desberdineko ikasleek idatzitako testuen emaitzak ematen dira. Alde batetik, erdiko (Erd-B) eta goiko (Goi-B) mailaren zortzina ikaslek idatzitako bina testu (16 testu), eta, beste alde batetik, beheko maila baxuko (Beh_b-A) eta beheko maila altuko (Beh_g-A) zortzina ikaslek idatzitako bina testu (16 testu).



VII.10 irudia. Kategoria gramatikalen erabilera (%) goiko, erdiko eta beheko hizkuntz mailetan.



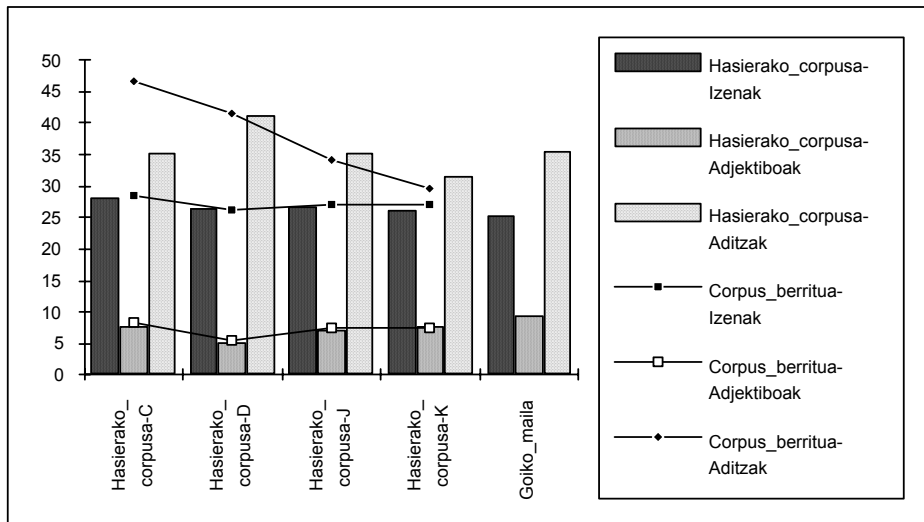
VII.11 irudia. Kategoria gramatikalen erabilera (%) ikaskuntza-maila igo ahala.

Aipatutako irudietan ikusten da nola ikaskuntza-maila igo ahala aditzen erabilera jaisten den, izenen erabilera, berriz, igotzen da, eta adjektiboak igo edo mantendu egiten dira. Honen arrazoi nagusia esaldien konplexutasunean dago: ikaskuntza-maila igo ahala esaldi konplexuagoak erabiltzen dira. Beraz, dirudienez, ikasleek esaldi luzeagoetan aditz gutxiago erabiltzen dituzte, dena den, honi buruz zabalago arituko gara § VII.3.3.4-n.

Ikaslearen ezaugarrien eragina kategoria gramatikalen azterketan.

Har ditzagun VII.9 irudian agertzen diren *Hasierako_corpusa* eta *Corpus_berritua* izeneko corpusak. *Hasierako_corpusa* corpusean lau ariketa-mota emanda lau ikasleek launa testu idatzi dute. *Corpus_berritua* corpusa osatzeko ikasle berri batzuek idatzitako bina testu

gehitu ditugu ariketa-mota bakoitzeko: C eta D ariketetan *i* eta *s* ikasleek idatzitako testuak, eta J eta K ariketetan *a* eta *s* ikasleek idatzitakoak. Katgoria gramatikalak aztertu ondoren (ikus VII.12 irudia) C ariketan nabarmenki igotzen da aditzen kopurua.



VII.12 irudia. Ikasle baten eragina katgoria gramatikalen azterketan.

Corpus_berritua corpusean C ariketan erantsitako bi testuak aztertuta, *i* ikasleak idatzitakoan ikusten da nabarmenki aditzen kopurua altua dela. Honen ondoren, aipatu ikasleak idatzitako 7 testu (1538 hitz) analizatu eta katgoria gramatikalen erabilera landu dugu. Emaitzak ondokoak izan dira: izenak (%27), adjektiboak (%8) eta aditzak (%35). Portzentaje hauek eta ikaskuntza-mailari dagozkionak (ikus VII.12 irudia) antzekoak dira. Beraz, C ariketan aditzen portzentajea ez da igo ikaslearen ezaugarri pertsonalengatik, baizik eta testua idatzi duen unean azkarrago idatzi duelako, edo une horretan bere helburua ekintzak azpimarratzea zelako.

Ondorioa: orokorki, hau bezalako probetan ondorioa ondokoa izan da: ikaslearen ezaugarri pertsonalek baino testua idazten duen unean ematen den egoerak izaten du eragina emaitzetan.

VII.3.3.4 Esaldien konplexutasunaren azterketa.

Atal honetan ikasleek idatzitako testuetan agertzen diren esaldien konplexutasunari buruzko datuak emango ditugu. Emaitzak erdiautomatikoki lortu dira euskararentzako analizatzaile sintaktikoa (Gojenola, 99) dela medio.

Analizatzaile sintaktikoak egoera finitutako gramatika testuei aplikatu eta esaldi bakoitzeko VII.2 taulan agertzen den bezalako emaitzak ematen ditu. Emaitza hauek

ateratzeko osagai sintaktikoen anbiguotasuna desanbiguatzeko du, honetarako hainbat heuristiko aplikatzen duelarik: izen sintagma eta menpeko esaldia bezala interpretatu daitezkeen kasuetan izen sintagma baztertzen da, sintagmaren mugak definitzean sintagma luzeena hartzen du anbiguotasun kasuetan, eta, azkenik, gelditzen diren interpretazio guztietatik informazio orokorra iragazten da patroio orokorrak lortuz, adibidez, *esaldi sinplea + eta + esaldi sinplea*.

Esaldia	Analisisa
Horregatik, susmatzekoa da honekin batera, garraioak ere igo egingo direla, bidaiatzea gero eta garestiagoa izan dadila zeren eta dibisak erosteko pezeta gehiago ordaindu beharko da erebai	Menpekota (-tzeko) Esaldi sinplea Menpekota (-ela) Menpekota (-tzea) Menpekota (-ela) Menpekota (-tzeko) Esaldi sinplea

VII.2 taula. Esaldi baten analisi sintaktikoa.

Esaldi-mailako emaitzak izanik, hitz-mailako azterketan lortutako hainbat datuekin erlazionatu dugu, hain zuzen ere hiztegiaren aberastasunarekin eta aditzen erabilerarekin. Honetaz gain, esaldietan erabilitako menpeko esaldiei buruzko informazioa ere landu dugu.

Hiztegiaren aberastasuna

Goiko eta erdiko hizkuntz mailetan idatzitako testuen hiztegia aberatsagoa da behekoetan baino, ez bakarrik ikasleek duten hiztegia zabalagoa dutelako, baizik eta beheko mailetan maiz erabiltzen dituztelako *eta*, *baina* bezalako juntagailuak.

Proba gisa ariketa-mota bera hartuta, beheko maila eta erdiko mailaren arteko konparaketa egin dugu. Bi ikaskuntza-maila desberdineko testuak hartu ditugu, konkretuki, beheko maila baxua eta beheko maila altua (erdiko mailatik gertu) (ikus VII.11 irudia). Maila bakoitzeko zortzi ikaslek idatzitako bina testu aztertuta garbi ikusten da maila altuenean hiztegia aberatsagoa dela. Arrazoi anitz topatzen ditugu honetarako:

- Testu-kopurua berdina izanik maila altuan esaldi gehiago osatzen dira, eta hitz-kopurua ere altuagoa da: 117 esaldi (1322 hitz) eta baxuan 99 esaldi (864 hitz).

- *eta, baina, edo* bezalako juntagailuak aztertuta, eta esaldien patroiak hartuta, maila altuan patroia horietako %58,82ak esaldi koordinaturen bat dauka eta maila baxuan %57,58ak. Beraz, patroiei dagokienez portzentajea antzekoa da. Dena den, corpusetan erabilitako patroia desberdinen portzentajea kalkulatu badugu ikusten dugu idatzitako esaldi-kopuruaren arabera %33,3 patroia desberdin daudela maila baxuan eta %58,12a maila altuan. Honek adieraz dezake, hipotesia besterik ez da, erabilitako esaldi koordinatuen kopurua handiagoa dela maila baxuan. Honi erantzuteko testuetako esaldien patroiak banan banan aztertu ditugu eta maila baxuan gehien errepikatzen diren patroiak benetan esaldi koordinatuei dagozkiela ziurtatu dugu. Konkretuki, *esaldi sinplea + eta + esaldi sinplea* patroia esaldi guztien %25,25a da maila baxuan eta %5,13a maila altuan. Aipatutako patroiaz gain, maila baxuan esaldi koordinaturen bat duen patroia errepikaturik ez dago, eta maila altuan, errepikatzen den beste bi daude; hauek suposatzen dute esaldi guztien %3,42a.

Ondorioa: laburbilduz, esan dezakegu testu bat hartuta maila altuan esaldi gehiago idazten dituztela maila baxuan baino, eta esaldi horiek luzeagoak. Maila baxuan, berriz, esaldi gutxiago, motzagoak eta esaldi koordinatu asko erabiltzen dituzte. Hiru faktore hauek bildurik hiztegia murriztagoa egiten dute.

Aditzen erabilera

Ariketa-mota bera hartuta, bi ikaskuntza-maila desberdineko ikasleek idatzitako testuetan ikusten da aditzen erabilera jaisten dela maila altuan. Honen arrazoia esaldien konplexutasunean dago. Adibide gisa, VII.11 irudian hartutako corpusa (beheko ikaskuntza-maila baxua eta altua) landu ondoren ondokoa ondorioztatu dugu:

Esaldi luzeagoak aurkitzen ditugu maila altuan, eta hauen konplexutasuna handiagoa, besteak beste, menpekoen erabilera aberatsagoa delako. Esaldietan erabilitako patroia desberdinak hartuta menpekorik gabeko patroien portzentajea txikiagoa da maila altuan (%32,35 versus %69,7a maila baxuan), bi menpekodun patroiak, berriz, ugariagoak dira (%29,41 versus %0a) eta hiru menpekodunak ere agertzen dira (%4,41 versus %0a). Honek esan nahi du aditzen erabilera jaitsi egingo dela.

Ondorioa: maila altuak eta baxuak konparatzean, ariketa-mota bera hartzen badugu maila altuetan idazten diren testuen luzera handiagoa da, esaldi-kopurua ere altuagoa izaten da, esaldien patroia-kopurua ere altuagoa, proportzioan juntagailuen erabilera murriztagoa da esaldien luzera kontuan izanik, eta, azkenik, menpeko esaldi gehiago erabiltzen dira. Guzti honegatik maila altuan aditzen erabilera jaitsi egiten da.

VIII. Ondorioak eta zabaldutako ikerlerroak.

Azken kapitulu honetan, tesian zehar aurkeztu dugun lanaren emaitzak eta ekarpenak aipatu ondoren azaldutako lanaren ondorioak deskribatuko ditugu. Alde batetik, metodologia dela eta diziplinarteko lanaren ondorioz atera ditugun ideiak azalduko ditugu, eta, beste aldetik, bigarren hizkuntzaren ikaste-prozesua ikertzeko garatu dugun MUGARRI izeneko inguruneak eskaintzen dizkigun ikerketarako irizpideak laburbilduko ditugu.

Azkenik, etorkizuneko lanari buruz arituko gara. Dena den, horretaz hainbat zehaztasun eman da dagoeneko IV, V eta VI kapituluetakoa azken ataletan. Haietan tesian garatu ditugun tresna informatikoetan landu beharrekoak aipatu genituen batez ere. Kapitulu honetan, berriz, zehaztasun horiek laburtzeaz gain, tesiaren ondorioz zabaldu diren ikerlerro berri nagusiak zerrendatuko ditugu; honela, lan honi bukaera emango diogularik.

VIII.1 Emaitzak eta ekarpenak.

Bigarren hizkuntzaren ikaste-prozesua ikertzeko MUGARRI ingurunea aurkeztu dugu tesi honetan. Aipatzekoa da ikasleek idatzitako testu errealen tratamendu automatikoa dela tesiaren helburua, laborategi-adibideak baztertu ditugularik. Gure proposamenean bi ekarpen-mota bereizten ditugu: oinarrizko planteamendu metodologikoak eta diseinatutako arkitektura.

Oinarrizko planteamendu metodologikoak

Oinarrizko planteamendu metodologikoei dagokienez hiru ekarpen nagusi egiten ditugu, bibliografiako sistemekin konparatuz gero:

- Lehenengo, errore kontzeptua desagertzen da gure hurbilpenean, hau da, hizkuntz estandarra ez da erreferentzia nagusia ikaslearen ezagutza linguistikoa ulertzean, baizik eta ikasleak eta bere ikaskuntza-mailako ikaskideek duten ikaste-egoera. Azken hau kontuan izanik errorearen parekoa izan daitekeen **desbideratze kontzeptua** proposatu dugu.
- Bigarren, ikaskuntza-maila definitzen dugunean hizkuntz eskoletan diseinatzen diren hizkuntz programak baztertzen ditugu ikaslearen ezagutza interpretatzeko. Hauen ordez, ikasleen benetako output-a aukeratzen dugu, eta honengan oinarrituta **ikaskuntza-mailen tartehizkuntz ereduak** definitzen ditugu ikasleen ezagutza linguistikoaz arrazontzeko.

- Hirugarren, **ebaluazio inkrementala** egiteko aukera eskaintzen digun MUGARRI ingurunea lortu dugu. Sistema asko eta asko, behin inplementatuta, aukera berriak ezin dituzte integratu. Guk, berriz, ingurune informatikoaren arkitekturan lortutako ezaugarriak direla medio arkitekturaren zabalkuntza inkrementala gainditu dugu.

Diseinatutako arkitektura

Proposatutako MUGARRI izeneko ingurunearen garapen informatikoari dagokionez, sistema anitzeko arkitektura diseinatu dugu modulartasunean oinarriturik. Ingurunean garatutako tresna informatikoak, alde batetik, autonomoak dira eta, beste aldetik, berrerabilgarriak dira testuinguru desberdinetan; esaterako, ingurunearen hainbat tresna linguistiko inguruneke agente autonomo batean baino gehiagotan erabili da zeregin desberdinetarako. Honetaz gain, proposatutako arkitekturan ezagutza-baseen maila bereiztu dugu ezagutzaren sendotasuna eta trinkotasuna lortzeko asmoz. Hortaz, MUGARRI ingurunearen arkitekturak ondoko ezaugarriak betetzen ditu: autonomia, berrerabilgarritasuna, modulartasuna, sendotasuna eta trinkotasuna. Hauei buruz zabalago arituko gara § VIII.2.2-n ondorioak azaltzean.

MUGARRIren arkitekturako osagaiak hiru mailatan bildu ditugu: sistemak, tresna linguistikoak eta ezagutza-baseak. Ondoko lerrootan bakoitzean egindako ekarpenak komentatuko ditugu.

Sistemen maila

Sistemen mailan hiru agente autonomo sortu ditugu: HITES, IRAKAZI eta IDAZKIDE sistemak, azken biak agente lagungarri gisa kokatzen direlarik. Hiruren ezaugarri komuna da giza agente baten partehartzea behar dutela, lehenengo kasuan psikolinguistarena, bigarrenean irakaslearena eta, hirugarrenean, ikaslearena. Ondoren, hauengan egindako ekarpen nagusiak aipatuko ditugu.

HITES sistema adimenduna da. Bere arkitekturan helburuetan eta atazen egoeretan oinarritutako arrazonatze- eta erabakitze-modulua agertzen zaigu. Honetaz gain, ikaskuntza-elementua izeneko modulua bidez ikaslearen zein tresna linguistikoaren eguneraketa erdiautomatiko burutzen da psikolinguistaren laguntzaz. Horretarako, saioari buruzko informazioa gordetzen da lan-memorian.

IRAKAZI irakaslearekin elkarreragiten du, bai ikasleari buruzko ezaugarriak jasotzeko, bai ikasleak erabiltzen dituen egitura desbideratuen diagnosia lantzeko. Prozesu honetan irakaslea da diagnosiari buruzko informazioa ematen duena. Egitura linguistikoak HITESek sortutako tartehizkuntz ereduetatik jasotzen ditu IRAKAZI.

IDAZKIDE sistema lagungarri adimenduna da, eta ez du sistema tutoreen filosofia jarraitzen. Bere helburua ikaslearen ikaste-prozesuan zehar laguntza ematea da, beraz, ez dago ikasleari irakasteko alde zurririk erakutsi beharreko ezagutza linguistiko antolaturik. Diseinu-mailan sistemak hiru ikaskuntza-metodo nagusi eskaintzen ditu: esplorazio askea, ikasketa gidatua eta ikasketa aske-gidatua. Modu gidatuan, sistemak ikaslearen ikaskuntza-maila, ikaste-estrategiak eta hainbat ezaugarri izaten ditu kontuan. Modu askean sistemak zenbait aukera eskuragarri jartzen dio ikasleari idazte-prozesuan zehar: hitz baten analisi morfologikoa egitea, hitz baten definizioa erakustea, hitz desbideratu baten proposamenak eskatzea, e.a. Azkenik, esango dugu IDAZKIDE sistema elkarreragileak jasotzen duela, alde batetik, HITES sistemak ikaskuntza-maila bakoitzeko sortutako ereduak, eta, beste alde batetik, IRAKAZIren bidez lortutako informazioa dela medio diagnosirako laguntza.

Ezagutza-baseen maila

Ezagutza-baseen mailan egindako ekarpen nagusiak bi dira: ikasle-eredu komuna eta tartehizkuntza definitzeko eredu kontzeptuala.

Ikasle-eredu komuna, sistema informatiko desberdinek erabiliko dutena. Hau dela eta, ereduaren ezaugarriek iragazki bina dute sistema bakoitzeko, bat adierazteko ezaugarria ikusgarria den sisteman eta bestea islatzeko ezaugarriaren portaera sistemarekiko (dinamikoa edo estatikoa).

Tartehizkuntza pentsamendu-objekturako eredu kontzeptualaren deskribapena. Tartehizkuntza hainbatetan definitua izan arren deskribapen-eredurik ez zaio aurkitzen literaturan. Guk ereduaren ezaugarri nagusiak definitu ditugu eta objektuei zuzendutako programazioa erabiliz implementatu. Definitutako ereduan tartehizkuntz egitura bat baldintza berdinetan suertatzen diren fenomeno linguistikoez osatuta dago. Eta, fenomeno linguistiko hauei ikuspegi desberdinak egokitzen zaizkie askotan, bai azal mailan bai maila linguistikoan.

Tresna linguistikoaren maila

Aipatutako sistema informatikoetan erabiltzeko hainbat tresna linguistiko egokitu edota sortu dugu.

Egokitzapenari dagokionez, ekarpen garrantzitsuena erregela estandar eta desbideratuen identifikazio automatikoa izan da. Lanaren oinarria den analizatzaile eta desanbiguatzaile morfosintaktikoak hasiera batean erregelak detektatu arren ez zituzten identifikatzen. Aipatutako bi tresna horiek egokitu ditugu aplikatutako erregelak identifikatzeko.

Sorkuntza dela eta, alde batetik, ikasleentzako tresna lagungarria lortu dugu hitzen deklinabidea automatikoki sortzeko; hau posible izan da Bi Mailatako Morfologia bai analisirako bai sorkuntzarako erabilgarria delako. Beste alde batetik, goitik beherako metodologia aplikatu ahal izateko ikaskuntza-maila desberdinetan egindako analisiak konparatzeko aplikazioa ere sortu dugu; hori psikolinguistarekiko elkarrekintzan erabiltzen da.

VIII.2 Ondorioak.

Tesi hau diziplinarteko proiektutzat jo dugu helburu gisa. Hori dela eta, irizpide metodologikoak definitu behar izan genituen hasiera batetik. Ildo berean lanean garatutako landa-lana MUGARRIren oinarria izan da, besteak beste. Honegatik, ondorioak bi multzo nagusitan banatu ditugu:

- Diziplinarteko lanaren ondorio metodologikoak.
- Bigarren hizkuntzaren ikaste-prozesua ikertzeko MUGARRI ingurunearen garapenerako irizpideak.

Azal dezagun, bada, bakoitzean ondorioztatutakoa.

VIII.2.1 Diziplinarteko lanaren ondorio metodologikoak.

Ondoren, tesi honetan egindako lanaren testuinguru diziplinartekoan lortu diren ondorioak aurrera nola eraman ditugun komentatuko dugu labur.

Diziplinarteko koordinazioa

Tresna informatikoen eraikuntza-prozesuan askotan informatikaren mugek arazo larriak planteatzen dituzte. Tresnak eraiki baino lehen, beharrak non dauden aztertzeko irakasle eta psikolinguisten iritzia beharrezkoa da. Aipatutako tresnak linguistikoak direnean, hauek diseinatzeko linguisten laguntza ere ezin da ahaztu, batez ere, hizkuntza normalizazio-prozesuan dagoen herri honetan. Hau guztia dela eta, lan honetan aurkeztu dugun prozesu metodologikoan ikusten denez, diziplinarteko proiektuen garapena egingarria dela frogatu dugu. Eta ez bakarrik egingarria baizik eta beharrezkoa. Adibide gisa, § III.4.2-n – diziplinarteko lan baten esperientzia azaltzen da han –, MUGARRI ingurune informatikoaren definizioarekin hasi ginenetik irakasleekin lan egitea nahitaezkoa dela ikusten da. Lan hartan ikusi genuen egitura desbideratuen sakoneko arrazoien azterketa egitean irakasle bakoitzak ikuspegi desberdina jarraitzen zuela: batek arrazoiak ikaslearen ezaugarri pertsonaletan identifikatzen zituen bitartean, besteak, testua eta esaldien egituraren baitan topatzen zituen. Honek adierazi zigun ingurune informatikoak irekia utzi behar ziola

irakasleari nahi zuen sailkapena erabiltzeko; informatikariok, berriz, hasiera batean helburu bezala izan genuen sailkapen bakarra eskaintzea irakasleei. Hortaz, diziplinarteko koordinazioaren beharra detektatu genuen hasieratik.

Ebaluazio inkrementala

Aurretik komentatu dugunez, informatika eta hizkuntzen ikaskuntza munduen diziplinarteko kultura sortzea beharrezkoa dela ikusten dugu. II. kapituluan aztertutako sistema informatikoen hainbat kasutan sumatzen da aipatutako bi arloen arteko koordinaziorik ez dela gehienetan egon: sistema informatiko batzuk garatu bai, baina gero ez dira erabiltzen. Zergatia, gure ustetan, helburu koordinatuen ezaren baitan izaten da gehienetan. Beste batzuetan, berriz, helburu espezifikoetarako tresna informatikoak direnez, zaharkituta gelditzen dira tresna horiek, epe luzerako helbururik ez dituztelako izaten. Sistema asko eta asko, behin inplementatuta, aukera berriak ezin dituzte integratu. Guk aurkeztutako ingurune informatikoan berriz, arkitekturaren zabalkuntza inkrementala gainditu dugu. IV. kapituluan azaldutako tresna linguistikoen egokitzapenak hauen isla dira. Proiektu osoan zehar irakasleekin eta ikasleekin egindako probetarako aurretik garatutako tresnak egokitu eta berriak integratu ditugu. Honek aukera eman digu ebaluazio inkrementala egiteko, hau da, I. kapituluan proposatutako definizio-diseinu-inplementazio-ebaluazio zikloa aplikatzeko.

VIII.2.2 Bigarren hizkuntzaren ikaste-prozesua ikertzeko MUGARRI ingurunearen garapenerako irizpideak.

Ondoko lerrootan tesian proposatu dugun MUGARRI ingurunearen sistema informatiko desberdinek betetzen dituzten ezaugarri nagusiak azalduko ditugu ondorio gisa.

Autonomia

Esan bezala, sistema askoren kasuan, behin inplementatuta, aukera berriak integratzea oso zaila izaten da. Tresna informatikoen berrerabilpena eta integrazioa, berriz, funtsezkoa da bigarren hizkuntzaren ikaste-prozesuaren azterketan. Arlo honetan zehazten diren helburuak epe luzera begira izaten dira; informatikak, berriz, aldaketa handiak izaten ditu epe motzetan. Horregatik, garrantzitsua da tresna berrien integrazioa eskaintzen dituzten sistemak garatzea. Eta hau izan da tesi honetan aurkeztu den ingurunearen oinarria: tresna informatikoak agente autonomoak (Müller, 98) dira, modulu gisa bere aldetik funtzionatzen dutenak, eta ikaskuntza-ingurune zabalago desberdinetan integra daitezkeenak.

Berrerabilgarritasuna

Lengoaia naturalaren prozesamenduan garatutako tekniken egokitzapena bigarren hizkuntzaren ikaste-prozesuaren azterketan bideragarria dela frogatu dugu honetan. Konkretuki, gaur egun, euskara hizkuntza lantzeko dauden tresna linguistiko-konputazionalak egokitu ditugu gure zereginetarako. MUGARRIko tresna linguistikoen berrerabilgarritasuna, beraz, frogatu dugu tresna bereko erabilpen desberdinak aurkeztu dugun momentutik. Adibide gisa analizatzaile morfosintaktikoa dugu: alde batetik analisi batean aplikatutako erregelak detektatzeko erabili dugu HITESen, eta, beste alde batetik, hitz baten osagai esanguratsuak ikasleari aurkezteko erabili dugu IDAZKIDEn.

Modulartasuna

Kontrola eta ezagutza ondo berezituak egotea ezinbestekotzat jotzen dugu xede orokorreko tresnak garatu nahi ditugunean. Ildo honetatik, aurretik aipatutako berrerabilgarritasuna posible izan da erabilitako tresnen modulartasuna dela medio, zehazki ondo bereiztuta daudelako kontrola eta ezagutza linguistikoaren arteko mugak. Hau egingarria izan da tresnetan erabilitako formalismoei esker, hau da, morfologian Bi Mailatako Morfologia (Koskenniemi, 83), eta morfosintaxian Murriztapen Gramatika (Karlsson, 92). MUGARRIn dauden HITES, IRAKAZI eta IDAZKIDE sistemen berrerabilgarritasuna ere bideragarria antzematen da bi arrazoi nagusiengatik: batetik, erabiltzen duten ezagutza ezagutza-baseen mailan kokatzen da, ezagutza linguistikoaren kasuan gertatzen den bezala; eta, bestetik, sistemen arkitekturei erreparatzen badiegu kontrola eta sistemen barneko ereduak ondo berezituak daude.

Sendotasuna

Hizkuntza baten ikaste-prozesuan zehar ikasleek batzuetan hainbat egitura linguistikoren erabilera ekiditen dute beren testuetan. Beraz, egitura horiek corpusetan esplizituki agertzen ez direnez, corpusetan soilik oinarritutako ikerketak ez dira sendoak, ikaslea behartu behar baita aipatu egiturak erabiltzen. Lan honetan hau bera frogatu dugu azaldutako landa-lanean. Honegatik, gure ingurunean egitura hauei buruzko ezagutza detektatzeko, IDAZKIDE bezalako ikaste-sistema lagungarria proposatu dugu: honek ikasleari testu libreen idazketan laguntzeaz gain egitura konkretuak lantzeko ariketak aurkezten dizkio.

Trinkotasuna

Xede-orokorreko ingurune informatikoetan ezagutzaren sendotasuna arazo handia izaten da. Honelako inguruneetan sistema desberdinek ezagutza bera erabiltzen dute askotan. Gure proposamena, beraz, ondokoa da: ezagutza-baseak agente autonomoak izatea, nahiz eta

ingurunearen beste osagaiekiko pasiboak izan. Modu honetan ezagutzaren trinkotasuna babesten dugu, eta ingurune HITES, IRAKAZI eta IDAZKIDE sistemen artean zeharkako komunikazioa lortzen dugu.

Zeharkako komunikazioa dela eta, honakoak dira lan honetan landu ditugunak:

- Tartehizkuntz egituren ezagutza-basea eguneratzeko metodoa inplementatuta dago.
- Tresna linguistikoaren ezagutza linguistikoa eguneratzeko erregelen sorrera automatikoa ere inplementatuta dago.
- Ikasle-ereduei dagokienez IKASKIDE sistemaren ikasle-ereduaren prototipo bat egin dago, ikasle-ereduaren gunea, berriz, hau da, ezagutza-baseen mailakoa ez da oraingoz modulu autonomo gisa garatu. Une honetan modulu hau IRAKAZI barruan inplementatu dugu probak egin ahal izateko.
- Azkenik, diagnosi-ezagutzaren kasuan diseinuaren ideia orokor batzuk eman ditugu.

Komunikazioa

Azkenik, ondorio garrantzitsua izan da gure lanean HITES, IRAKAZI eta IDAZKIDE bezalako sistemen arteko komunikazioa garatzea beharrezkoa dela bigarren hizkuntzaren ikaste-prozesua automatikoki lantzeko.

Zuzeneko komunikazioari dagokionez bi aspektu azpimarratu behar ditugu:

- Tresna linguistikoekiko komunikazioa: kasu honetan bai HITES bai IDAZKIDE sistemarekiko komunikazioa inplementatu dugu.
- HITES, IDAZKIDE eta IRAKAZI sistemen arteko komunikazioa: arlo honetan elkarren arteko komunikazioaren lerro nagusiak definitu ditugu.

VIII.3 Zabaldutako ikerlerroak.

Etorkizuneko lana dela eta, hainbat zehaztasun aipatu dugu dagoeneko IV., V. eta VI. kapituluaren azken ataletan. Hauetan tesian garatu ditugun tresna informatikoetan landu beharrekoak aipatu genituen. Atal honetan, zehaztasun haiek aipatu besterik ez ditugu egingo. Dena den, horiek gogora ekarri aurretik tesi honek zabaldu dituen ikerlerro berri nagusiak komentatuko ditugu.

VIII.3.1 Zabaldutako ikerlerro nagusiak.

Esan bezala, garatu ditugun sistemak autonomoak dira eta ingurune berrietan integratzeko erraztasunak eskaintzen dituzte. Honetaz gain, hainbatetan aipatu dugu garatutako moduluen

arteko integrazioa eta komunikazioa ere landu behar dela etorkizun hurbilean. Hau guztia dela eta, hiru ikerlerro nagusi definitzen ditugu etorkizunean garatzeko:

- Agenteanitz elkarreragileez osatutako ingurunea.
- Komunikazio-teknika berrien integrazioa MUGARRIn.
- MUGARRIren erabilpena hizkuntzen ikaskuntzarako metodologiaren ikerketan.

Azter ditzagun banan banan:

Agenteanitz elkarreragileez osatutako ingurunea

Gure lehen proposamena ikerlerro nagusi bezala ondokoa da: HITES, IRAKAZI eta IDAZKIDE sistema autonomoak, eta hauek konpartitzen dituzten baliabideak agente autonomo (Müller, 98) gisa planteatu ingurune erreal batean integratuz, ingurune hiru sistemen arteko komunikazio osatua lortzeko asmoz. Aurreko atalean esan dugun bezala konpartitutako baliabideen bidezko komunikazioa batez ere garatu dugu tesi honetan. Zuzeneko komunikazioari dagokionez, berriz, lerro nagusien definizioari mugatu dugu gure zeregina.

Müller-ek (1998) gaur arte agente adimendunen arloan garatu diren aplikazioak eta arkitekturak aztertu, eta azterketa horretatik abiatuta agenteen taxonomia proposatzen du bi ezaugarri desberdinetan oinarrituta: agentearen egoera materiala eta agentea eta ingurunearen arteko elkarrekintza-mota.

Taxonomiatik abiatuta ondoko agente-motak proposatzen ditu: HW agente autonomoak, SW agente autonomoak, HW agenteanitzak, SW agenteanitzak, HW agente lagungarriak eta SW agente lagungarriak.

Sailkapen hau kontuan izanik, tesi honetan garatu ditugun sistemak honela definituko genituzke: HITES sistema SW agente autonomoen artean kokatzen da, bere jardura autonomoa delako eta ingurune batean erabakiak hartzea bere helburua delako; IDAZKIDE, (Johnson & Rickel, 97)-n azaltzen diren agente hezitzaileen moduan, SW agente lagungarria da bere exekuzioa erabiltzailearen nahien arabera egokitzen delako; eta. IRAKAZI SW agente lagungarrien artean kokatzen da, irakasleari laguntzen diolako ikasleak egindako desbideratzeen diagnosi-prozesuan.

Dena den, lan honetan proposatu dugun ingurune informatikoaren azken helburua aipatutako sistemen arteko kooperazioa dugunez, agenteanitzeko ingurunea garatu beharko genuke, eta hau da etorkizunean ikertzeko proposatzen dugun lerroa. Ingurune horretan aipatutako hiru sistemek elkarrekiko komunikazio zuzena izango lukete portaera sozial-kooperatiboa oinarri izanik. Ezin dugu ahaztu HITESek ikasleen corpusetatik eskura dezakeen informazioa osatua ez denez, ikaslearengandik informazioa jaso beharra duela

(IDAZKIDE), eta irakaslearengandik diagnostikoari buruzko iritzia ezagutu behar duela (IRAKAZI). Horregatik, etorkizuneko lanean SW agenteanitz elkarreragileak garatzea proposatzen dugu. Horretarako, hiru sistemen arteko zuzeneko komunikaziorako protokoloak aztertu beharko dira.

Komunikazio-teknika berrien integrazioa MUGARRIn

Bigarren proposamena dugu MUGARRI zabaltzea gaitasun komunikatiboa aztertzeke asmoz. Tesi honetan ikaste-prozesuaren arlo konkretu bat aztertu dugu bereziki, tartehizkuntzaren konpetentzia gramatikala hain zuzen. Honek, berriz, ikaslearen ezagutzaren ikuspegi partziala eskaintzen digu. Gaitasun komunikatiboak, bestalde, ikaslearen ezagutzaren beste ikuspegia islatzen du. Gaur egun, multimedia aplikazioak eta ahots tratamendurako teknikak direla medio gaitasun komunikatiboa lantzeko bidea ematen digute. Horregatik, teknika hauen bidez tesian aurkeztu dugun ingurunea zabaltzea etorkizunerako lan interesgarritzat jotzen dugu.

MUGARRIren erabilpena hizkuntzen ikaskuntzarako metodologiaren ikerketan

Hirugarren ikerlerro nagusia honela definitzen dugu: bigarren hizkuntzaren ikaste-prozesua ikertzeko ingurunea amaraunean integratu euskaltegi desberdinetan jarraitzen diren metodologiaren emaitzak aztertzeke.

Adibide gisa ondoko galderak planteatu ditzakegu metodologia baten emaitzak lantzean: Zergatik batzuetan, 2. mailan ikasitako kontzeptu batzuk 10. mailan, esaterako, ez dira oraindik barneratu?, edo zergatik ikasle batzuek lortzen duten hizkuntz aberastasuna handiagoa da nabarmenki beste batzuek lortzen dutenarekin konparatuz? Galdera asko planteatu dira arlo honetan. Hauei erantzuteko proiektu diziplinartekoa planteatuz gero, corpus handiak aztertu eta datu nahiko zehatzak lor daitezke.

VIII.3.2 HITES, IRAKAZI eta IDAZKIDE sistemetan zein tresna linguistikoetan landu beharrekoak.

Tresna linguistikoaren kasuan ondokoak dira landu beharreko puntuak:

- Corpusetan agertzen diren egitura linguistikoaren **detekzioa zabaldu sintaxi mailara**. Arlo honetan lehen esperimentuak egin dira § VII.3.3.4-n ikusi dugunez. Honetarako, Gojenolaren (1999) tesian garatutako analizatzaile sintaktikoa erabiltzen ari da.

- Corpusetan agertzen diren egitura linguistikoaren **detekzioa zabaldu semantika-mailara** ikasleak bere testuetan erabiltzen dituen kontzeptuen mapa lortzeko asmoz.
- Tresna linguistikoaren **azkartasuna eta eraginkortasuna aztertu eta hobetu**. Bide honetan, gaur egun dagoen hainbat tresnaren integrazioa landu behar dugu, besteak beste, Xeroxen *Twolc* eta *XFST (Finite State Tools)* tresnak (Karttunen et al., 97).

HITES sisteman landu beharreko arloak ondokoak dira:

- Tresna linguistikoaren eta tartehizkuntz egiturak sortzen dituen moduluaren arteko **komunikazioa gauzatu**.
- HITES sistemaren atazan **kontrol-egitura integratu ataza objektuaren ezagutza prozedurala bezala**.

IRAKAZI sisteman garatu beharreko puntuak ondokoak dira:

- Diagnosi-prozesuan **irakaslearen diagnostika sistemak berak gida dezala** planteatu, laguntza-gakoak emanez, eta kontuan izanik, beste irakasle batzuekin izandako saioak.
- IDAZKIDEk bidalitako saioen laburpenak tratatuz **diagnosi-ezagutza eguneratu**.

IDAZKIDE ikaste-sistema lagungarria landu beharrek dira:

- **Ikasle-ereduaren modulua interfazean integratu**.
- **Tresna linguistiko berriak egokitu**, besteak beste, sinonimo-hiztegiak, analizatzaile sintaktikoa.
- **Ariketak gehitu** idazte-prozesuaren laguntza moduan, tartehizkuntz egituren zehaztasun hobea lortzeko asmoz.
- Ariketen **sorrera automatikoa bideratu**.

BIBLIOGRAFIA

- Adjemian, C.. On the nature of interlanguage systems. *Language Learning*, 26, 297-320. 1976.
- Aduriz, I.. *Errore Ortografikoen Azterketa eta Zuzenketa Bi Mailatako Morfologiaren Arabera*. UPV/EHU Euskal filologia. Barne-txostena-doktoregoa. 1994.
- Aduriz I., Aldezabal I., Alegria I., Artola X., Ezeiza N., Urizar R.. EUSLEM: A lemmatiser/Tagger for Basque in *Proc. of EURALEX'96*. Göteborg, Sweden. 1996.
- Aduriz I., Alegria I., Artola X., Ezeiza N., Sarasola K., Urkia M.. A spelling corrector for Basque based on morphology. *Literary & Linguistic Computing*, 12 (1). Oxford University Press, Oxford, U.K.. 1997.
- Aduriz I., Aldezabal I., Ansa O., Artola X., Díaz de Ilarraza A., Insausti J.M.. EDBL: a Multi-Purposed Lexical Support for Treatment of Basque, in *Proc. of First International Conference on Language Resources and Evaluation*. Granada, Spain. 1998.
- Agirre, E., Alegria, I., Arregi, X., Artola, X., Díaz de Ilarraza, A., Maritxalar, M., Sarasola, K. & Urkia, M.. XUXEN: A spelling checker/corrector for Basque based on Two-Level Morphology, in *Proc. of the Third Conference ANLP (ACL)*, 119-125. Trento, Italy. 1992.
- Agirre E., Arregi X., Arriola J.M., Artola X., Díaz De Ilarraza A., Insausti J.M., Sarasola K.. Different issues in the design of a general-purpose Lexical Database for Basque in *Proc. of First Workshop on Applications of Natural Language to Databases*. Versailles, Francia. 1995.
- Aldezabal I., Alegria I., Artola X., Díaz de Ilarraza A., Ezeiza N., Gojenola K. Aduriz I., Urkia M.. EUSLEM: Un lematizador/etiquetador de textos en euskara, in *Proc. of the X. Congreso de SEPLN*. Córdoba : SEPLN. 1994.

- Alegria, I. *Euskal morfologiaren tratamendu automatikorako tresnak*. Euskalerriko Unibertsitatea. UPV-EHU. Euskalerrria. Doktoretza-tesia. 1995.
- Alegria I., Arriola J.M., Artola X., Díaz de Ilarraza A., Gojenola K., Maritxalar M., Aduriz I.. A Corpus-Based Morphological Disambiguation Tool for Basque, *in Proc. of the XII. Congreso de SEPLN*. Sevilla : SEPLN. 1996.
- Andueza A., Díaz de Ilarraza A., Maritxalar M., Martiarena J., Pikabea, I. *Hizkuntza baten Ikaskuntza Prozesuari buruzko landa lana. Sistema Informatiko adimendun baten oinarria*. Barne-txostena, UPV/EHU/LSI/TR 8-96. Informatika fakultatea. UPV/EHU. 1996.
- Arregi, X.. *ANHITZ: Itzulpenean laguntzeko hiztegi-sistema eleanitza*. Euskal herriko Unibertsitatea. UPV-EHU. Euskalerrria. Doktoretza-tesia. 1995
- Arruarte, A.. *Fundamentos y diseño de IRIS: Un entorno para la generación de Sistemas de Enseñanza Inteligentes*. Euskalerriko Unibertsitatea. UPV-EHU. Euskalerrria. Doktoretza-tesia. 1998.
- Barchan, J., Woodmansee, B., & Yazdani, M.. A Prolog-based tool for French grammar analysis. *Instructional Science*, 15, 21-48. 1986
- Beauvois, M. H.. High-Tech, High-Touch: From Discussion to Composition in the Networked Classroom. *Computer Assisted Language Learning*, 10 (1), 57-69. 1997.
- Black, A., van de Plassche, J., Williams, B.. Analysis of Unkown Words through Morphological Decomposition, *in Proc. of 5th Conference of the EACL*, 1, 101-106. 1991.
- Brent, B. W.. *Practical Programming in Tcl and TK*. Prentice Hall. New Jersey. ISBN 013-182007-9. 1995.
- Brent, M. R.. Advances in the computational study of language acquisition, *Cognition* , 61, 1-38. 1996.
- Bull, S.. Promoting effective learning strategy use in CALL. *Computer Assisted Language Learning*, 10 (1), 3-39. 1997.
- Burmeister, B. & Sundermeyer, K.. Cooperative problem-solving guided by intentions and perception in Demazeau, Y. & Werner, E (eds). *Decentralized A.I.*, 3, North-Holland. 1992.
- Busquets, F.. Clic: un entorno para el desarrollo de actividades educativas multimedia. *Revista Novática*, 117. 1995.
- Carr, B & Goldstein, I.. Overlays: A Theory of Modelling for Computer-Aided Instruction. *International Journal of Man-Machine Studies*, 5, 215-236. 1997.
- Chanier, T.. Language Learning. *International Journal of Artificial Intelligence in Education*, 5(4), 417-428. 1994.

- Chomsky, N.. *Aspects of the theory of syntax*. MIT Press, Cambridge, Mass. 1965.
- Cook, V.. *Linguistics and Second Language Acquisition*. England: The Macmillan Press LTD. 1993.
- Corder, S.. The significance of learner's errors. *International Review of Applied Linguistics*, 5, 161-70. 1967.
- Corder, S.. *Describing the Language Learner's Language*. Center for Information on Language Teaching (CILT), Reports and Papers 6, 57-64. 1971a.
- Corder, S.. Idiosyncratic dialects and error analysis. *International Review of Applied Linguistics*, 9, 147-59. 1971b.
- Corder, S.. *Introducing Applied Linguistics*. Harmondsworth, Middx: Penguin Books. 1973.
- Corder, S.. Error Analysis. In Allen, J. & Corder, S., (eds.), *The Edinburgh Course in Applied Linguistics*. Oxford: Oxford University Press. 1974.
- Cuena, J.. Los sistemas multiagente basados en el conocimiento: ¿una posible alternativa para la moderna ingeniería del software?. *Inteligencia Artificial*, 5, 85-95. 1998.
- Culley, G., Mulford, G. & Millbury-Steen, J.. A foreign language adventure game: Progress report on an application of AI to language instruction. *CALICO Journal*, 4, 69-94. 1986.
- Díaz de Ilarraza, A.. *Diseño de un Módulo de Interacción Tutor-alumno para un Sistema Inteligente de Enseñanza de la Programación*. Euskalerriko Unibertsitatea. UPV-EHU. Euskalerrria. Doktoretza-tesia. 1990.
- Díaz de Ilarraza, A., Maritxalar, M. & Oronoz, M.. Reusability of NLP Tools for Detecting Rules and Contexts when Modelling Language Learners' Knowledge in *Proc. of Recent Advances in Natural Language Processing Conference*, 342-347. Tzigov Chark. Bulgaria. 1997.
- Díaz de Ilarraza A., Maritxalar M., Oronoz M.. An Implemented Interlanguage Model for Learners of Basque. In *Language Teaching and Language Technology*. Swets and Zeitlinger (Publisher). Sake Jager, John Nerbonne & Arthur van Essen (eds.), 149-166. Lisse. 1998a.
- Díaz de Ilarraza A., Maritxalar A., Maritxalar M., Oronoz M.. Integration of NLP Tools in an Intelligent Computer Assisted Language Learning Environment for Basque: IDAZKIDE, in *Proc. of Natural Language Processing and Industrial Applications*, 25-32. Moncton, Canada. 1998b.
- Díaz de Ilarraza A., Maritxalar A., Maritxalar M., Oronoz M.. IDAZKIDE: an intelligent computer-assisted language learning environment for second language acquisition. *ReCALL Journal of the CTI Centre for Modern Languages*, 11 special issue, 12-19. University of Hull. 1999.

- Dimitrova, V. & Dicheva, D.. Learning terminology in a Foreign Language in *Proc. of Recent Advances in Natural Language Processing Conference*. Tzigov Chark. Bulgaria. 348-353. 1997.
- Dulay, H. C., Burt, M, Krashen, S.. *Language Two*. Rowley, Mass.: Newbury House. U.K. 1982.
- Edge, J.. *Mistakes and Correction*. Longman: London. 1989.
- Ellis, R.. *Understanding second language acquisition*. Oxford University Press, Oxford. 1985.
- Elorriaga, J. A., Ferrero, B., Fernández-Castro, I.. GBC: una herramienta para la construcción de bases de conocimiento, en *Actas de CAEPIA-95*. Alicante, 367-376. 1995.
- Elorriaga, J. A.. *Planificación de la Instrucción en Sistemas Tutores Inteligentes desde un Enfoque de Razonamiento basado en Casos*. Euskalerriko Unibertsitatea. UPV-EHU. Euskalerrria. Doktoretza-tesia. 1998.
- Fernández de Castro, I.. *Estrategias de Enseñanza en un Sistema Inteligente de Enseñanza asistida por ordenador*. Euskalerriko Unibertsitatea. UPV-EHU. Euskalerrria. Doktoretza-tesia. 1988.
- Frasson, C. & Gouardères, G.. Workshop on Pedagogical Agents (ITS'98). Texas, U.S.A.. 1998.
- Frederiksen, C., Donin, J., Décary, M., & Hooper, M.. Semantic discourse processing and tutoring systems for second language learning. In M.L. Swartz & M. Yazdani (eds.), *Intelligent Tutoring Systems for Foreign Language Learning* 103-121. Berlin: Springer-Verlag. 1992
- Gereka, J.. Software kurrikularraren garrantzia. *Informatikari Euskaldunen II. Topaketak*. UEU. Donostia. 1998.
- Gojenola, K. & Sarasola, K.. Aplicacion de la Relajación Gradual de Restricciones para la Detección y Corrección de Errores Sintácticos, en *Actas del Congreso S.E.P.L.N.*. Córdoba. Spain. 1994.
- Gojenola, K.. *Euskararen tratamendu sintaktikoa*. Euskalerriko Unibertsitatea. UPV-EHU. Euskalerrria. Doktoretza-tesia (burutzear). 1999.
- Gregg, K.. Krashen's Monitor and Occam's Razor. *Applied Linguistics*, 5, 79-100. 1984.
- Haddadi, A. *Communication and Cooperation in Agent Systems: A Pragmatic Theory*. Springer-Verlag. Berlin. 1996.
- Hamel, M. & Vandeventer, A.. FIPSGram: Un analyseur syntaxique dans un contexte d'ELAO. in *Proc. of Natural Language Processing and Industrial Applications*, 18-24. Moncton, Canada. 1998.

- Holland, M., Kaplan, J. D. & Sams, M. R.. *Intelligent Language Tutors: Theory Shaping Technology*. Lawrence Erlbaum Associates, Mahwah, New Jersey. 1995.
- Holt, P., Dubs, S., Jones, M., Greer, J.. The State of Student Modelling. *In Student Models: The Key to Individualized Educational Systems*. J. Greer and G. McCalla (eds.). Springer Verlag, 3-35. 1994.
- Johnson, WL & Rickel, J.. Integrating pedagogical capabilities in a virtual environment agent, *in Proc. First International Conference on Autonomous Agents*, ACM Press, 30-38. 1997.
- Karlsson, F., Voutilainen, A., Heikkilä, J., Anttila, A.. *Constraint Grammar: a language-independent system for parsing unrestricted text*. Mouton de Gruyter. 1995.
- Karttunen, L., Beesley, K. R.. *Two-level Rule Compiler*. Technical Report. ISTL-92-2. Xerox PARC. Palo Alto. California. 1992.
- Karttunen, L., Chanod, J. P., Grefenstette, G., Schiller, A.. Regular Expressions for Language Engineering. *Natural Language Engineering*, 2 (4) 305-328. Cambridge University Press. 1997.
- Koskenniemi, K. *Two-level Morphology: A general Computational Model for Word-Form Recognition and Production*. University of Helsinki, Department of General Linguistics. Publications no. 11, 1983.
- Kramsch, C., Morgenstern, D., & Murrey, J.. An overview of the MIT Athena language learning project. *CALICO Journal*, 2, 31-34. 1985.
- Krashen, S.. Some issues relating to the Monitor Model. In Brown, H, Yorio, C, & Crymes, R (eds.) *in Proc. of TESOL '77*, 144-58. TESOL, Washington, D.C. 1977.
- Krashen, S.. Individual variation in the use of the monitor. In Ritchie, W (eds.) *Second Language Acquisition Research*, 175-83. Academic Press, New York. 1978.
- Lessard, G., Maher, D., Tomek, I.. Modelling Second Language Learner Creativity, *Journal of Artificial Intelligence in Education*, 5(4). 455-480. 1994.
- Long, M.. Input and Second Language Acquisition Theory, in Gass, S. & Madden, C. (eds.). *Input in Second Language Acquisition*, 377-93. Newbury House, Rowley, Mass. 1985.
- Maritxalar, M., & Díaz de Ilarraza, A.. *Integration of Natural Language Techniques in the ICALL Systems Field: The treatment of incorrect knowledge*. Barne-txostena. UPV/EHU/LSI/TR 9-93. Informatika Fakultatea (UPV/EHU). 1993.
- Maritxalar, M., & Díaz de Ilarraza, A.. An ICALL System for Studying the Learning Process. *Computers in Applied Linguistics Conference*. Iowa State University. 1994.
- Maritxalar, M., & Díaz de Ilarraza, A.. Informatika eta Hizkuntzen ikaskuntza. *ELE-16*, 59-73. Euskalerrria. 1995.

- Maritxalar, M., & Díaz de Ilarraza, A.. *Hizkuntza baten Ikaskuntza-Prozesuan zeharreko Tarte-hizkuntz Osaketa: Sistema Informatiko baten Diseinurako Azterketa Psikolinguistikoa*. Barne-txostena. UPV/EHU/LSI/TR 7-96. Informatika Fakultatea (UPV/EHU). 1996a.
- Maritxalar M., Díaz de Ilarraza A., Alegria, I., Ezeiza, N.. Modelización de la Competencia Gramatical en la Interlengua basada en el Análisis de Corpus, in *Proc. of the XII. Conference SEPLN*. Sevilla, Spain. 1996b.
- Maritxalar M., Díaz de Ilarraza A., Oronoz M.. From Psycholinguistic Modelling of Interlanguage in SLA to a Computational Model in *Proc. of CONLL97 Workshop (ACL Conference)*. Madrid. Spain. 1997.
- McLaughlin, B.. The monitor model: some methodological considerations. *Language Learning*, 28, 309-32. 1978.
- Menzel, W.. Error diagnosis and selection in a training system for second language learning, in *Proc. of COLING-88*, 414-419. 1988.
- Morante, R. & Martí, T.. Learning Spanish and Catalan Verbs through EuroWordnet, in *Proc. of Natural Language Processing and Industrial Applications*, 58-64. Moncton, Canada. 1998.
- Mulford, G.. Semantic processing for communicative exercises in foreign-language learning. *Computers and the Humanities*, 23, 31-44. 1989.
- Müller, J. P.. Architectures and applications of intelligent agents: A survey. *The Knowledge Engineering Review*, 13(4), 353-380. 1998.
- Nemser, W.. Approximate Systems of Foreign Language Learners. *International Review of Applied Linguistics*, 9, 115-23; 1971.
- Nerbonne, J. & Smit, P.. GLOSSER-RuG: in support of reading, in *Proc. of COLING'96*, 830-835. Copenhagen. 1996.
- Norrish, J.. *Language learners and their errors*. Macmillan. 1981.
- O'Malley, J. M., Chamot, A. U., Stewner-Manzanares, G., Kupper, L., Russo, R. P.. Learning strategies Used by Beginning and Intermediate ESL Students. *Language Learning*, 35, 21-46. 1985.
- O'Malley, J. M., & Chamot, A. U.. *Learning strategies in second language acquisition*. Cambridge: Cambridge University Press. 1990.
- Oxford, R. & Cohen, A.. Language learning strategies: Crucial issues of concept and classification. *Applied Language Learning*, 3, 1-35. 1992.
- Pérez, T. A.; Gabiola, K.; Gutiérrez, J.; López, R.; González, A. & Carro, J.. Hezinet: Interactive (Adaptive) Education Through Activities. *Educational Multimedia and Hypermedia, ED-MEDIA'99*. AACE: Seattle, Washington, U.S.A.. 1999.

- Pulman, S.. Limited domain systems for language teaching, in *Proc. of COLING-84*, 84-87. 1984.
- Randal, L., Schwartz, L. W.. *Programming Perl. UNIX programming*. O' Reilly & Associates, Inc. ISBN 0-937175-64-1. 1991.
- Richards, J.. Error Analysis and second language strategies. *Language Sciences*, 17, 12-22. 1971.
- Schachter, J. & Celce-Murcia, M.. Some reservations concerning error analysis. *TESOL Quarterly*, 11, 441-51. 1977.
- Schneider, D. A. & McCoy, K. F.. Recognizing Syntactic Errors in the Writing of Second Language Learners, in *Proc. of the 17th International Conference on Computational Linguistics (COLING-ACL '98)*. 1998.
- Schoelles, M. & Hamburger, H.. The NLP Role in Animated Conversation for CALL, in *Proc. of Recent Advances in Natural Language Processing Conference*, 127-134. Tzigov Chark. Bulgaria. 1997.
- Schreiber G., Wielinga B. & Breuker J.. (eds.). KADS: A Principled Approach to Knowledge-Based System Development. *Knowledge-Based Systems*, 11. London: Academic Press. 1993.
- Schuster, E.. The role of native grammars in correcting errors in second language learning. *Computational Intelligence*, 2, 93-99. 1986.
- Schwind, C.. An intelligent language tutoring system. *International Journal of Man-Machine Studies*, 33, 557-579. 1990.
- Self, J.. A Framework for Learner Modelling. *Interactive Learning Environments*, 2(2), 111-137. 1992.
- Selinker, L.. Interlanguage. *International Review of Applied Linguistics*, 10 (3), 209-231. 1972.
- Selinker, L.. *Rediscovering Interlanguage*. London: Longman. 1992.
- Selinker, L. & Gass, S.. *Second Language Acquisition*. London: Lawrence Erlbaum Associates. 1994.
- Selinker, L., & Douglas, D.. *Interlanguage-draft*. 1994.
- Stevick, E.. *Memory, meaning and method*. Newbury House, Rowley, Mass. 1976.
- Swartz, M., & Yazdani M.. *Intelligent Tutoring Systems for Foreign Language Learning: The Bridge to International Communication*. Berlin: Springer-Verlag. 1992.
- Tarone, E.. Interlanguage as chameleon. *Language Learning*, 29, 181-91. 1979.
- Taylor, B.. Toward a theory of language acquisition. *Language Learning*, 24, 23-36. 1974.

- Urkia, M.. *Euskal morfologiaren analisi automatikorantz*. Euskalerriko Unibertsitatea. UPV-EHU. Euskalerrria. Doktoretza-tesia. 1997.
- Verdejo, M. F.. User Modelling in Knowledge-Based Systems. *Cognition, Semantics and Philosophy*, J. Ezquerro and J. M. Larrazabal (eds.), Kluwer Academic Publishers, 23-46. 1992.
- Weber, G. & Specht, M.. User modeling and adaptive navigation support in WWW-based tutoring systems, in *Proc. of User Modelling (UM' 97)*. Cagliari, Italy. 1997.
- Weischedel, R., Voge, W., & James, M.. An artificial intelligent approach to language instruction. *Artificial Intelligence*, 10, 225-240. 1978.
- Werner, E.. Towards an Expert System for Upper Sorbian, in *Proc. of Computational Lexicography Complex'94*, Budapest, 245-252. 1994.
- Wolfe-Quintero, K., Inagaki, S. & Kim, H. Y. *Second language development in writing: Measures of fluency, accuracy and complexity*. Technical-report, 7. Honolulu: University of Hawai. 1998.
- Woolf B. P., Murray T.. Using Machine Learning to Advise Student Model. *Journal of Artificial Intelligence in Education* , 3(4). 1993.
- Yang, J. & Akahori, K.. Development of Computer Assisted Language Learning System for Japanese Writing Using Natural Language Processing Techniques: A Study on Passive Voice in B. Boulay & R. Mizoguchi (eds.), *Artificial Intelligence in Education*. IOS press, 263-270. 1997.
- Yazdani, M. *An English Tutor: Project Report (1987-89)*. Sheffield: Training Agency. 1990.
- Zock, M.. SWIM or sink: The problem of communicating thought. In M.L. Swartz & M. Yazdani (eds.). *Intelligent Tutoring Systems for Foreign Language Learning*, 235-247. Berlin: Springer-Verlag. 1992.
- Zubiri I. Ikas-Txip, euskara ordenadore bidez. *Zutabe*, 24, 9-55. Donostia: H.A.B.E.. 1990.
- Zubiri, I.. *Gramática Didáctica del Euskara*. H.A.B.E. 1991.
- Zubiri, I.. *Gramática Didáctica del Euskara*. H.A.B.E. 1994.

IRUDIEN ZERRENDA

- I.1 irudia.** MUGARRI ingurunearen giza agenteak.
- I.2 irudia.** Ingurunearen garapenaren eskema: ebaluazioa hobekuntzaren eragile gisa.
- I.3 irudia.** Hiru sistemen arteko komunikazioa.
- I.4 irudia.** MUGARRI ingurunearen arkitektura.
- IV.1 irudia.** Ezagutza aldakorra eta ezagutza finkoa.
- IV.2 irudia.** *Hizkuntz_ezagutza_mailarena* eta *Tartehizkuntz_eredu_komuna* definizioak.
- IV.3 irudia.** *Baldintza* egituraren definizioa.
- IV.4 irudia.** Baldintza-mota kontzeptuen definizioa.
- IV.5 irudia.** *Fenomeno_linguistikoa* egituraren definizioa.
- IV.6 irudia.** *institutoko* hitzaren ikuspegi azalekoak eta linguistikoak.
- IV.7 irudia.** *Fenomeno_aldaera* egituraren definizioa.
- IV.8 irudia.** *Erregela* egituraren definizioa.
- IV.9 irudia.** *Aldakortasuna* multzoaren definizioa.
- IV.10 irudia.** *Ikasle_multzoa* multzoaren definizioa.
- IV.11 irudia.** X ikaskuntza-mailaren tartehizkuntz ereduaren modelizazioa.
- IV.12 irudia.** Goitik beherako metodologia.
- IV.13 irudia.** *hark* hitzaren bi analisi.
- IV.14 irudia.** *MOR_LEKTE3*: -e letra kendu du morfeman.
- IV.15 irudia.** *MOR_LEGBE*: -e letra gehitu du morfeman.
- IV.16 irudia.** *arriskurik* hitzaren analisiak desanbiguatzailea pasa aurretik (formato laburtua).
- IV.17 irudia.** *arriskurik* hitzaren analisiak desanbiguatzailea pasa ondoren (formato laburtua).
- IV.18 irudia.** Testuinguruetan oinarritutako desanbiguatzailearen adibidea.
- IV.19 irudia.** *analisi* hitzaren analisi morfosintaktikoa.
- IV.20 irudia.** Ikaskuntza-mailaren araberako analisi morfosintaktikoa.
- IV.21 irudia.** *ereduak_lortu* tresnaren irteera. LEOEAE erregelaren eredia.

- IV.22 irudia.** *ereduak_lortu* tresnaren irteera. LEKTE1 erregelaren eredia.
- IV.23 irudia.** *ereduak_lortu* tresnaren irteera. MS_EZ erregelaren eredia.
- IV.24 irudia.** *arraio* hitzaren analisia.
- IV.25 irudia.** h letraren ezabatzea tratatzeko erregelak hasierako analizatzailean.
- IV.26 irudia.** Erregelak egokitutako analizatzailean.
- IV.27 irudia.** LEKHH eta LEKEH erregelen automatak.
- IV.28 irudia.** & diakritikoa duten hitzei dagozkien erregela morfofonologikoak.
- IV.29 irudia.** LEKBA3 erregelaren automata.
- V.1 irudia.** Hiru sistemen arteko komunikazioa.
- V.2 irudia.** MUGARRI ingurunearen arkitektura.
- V.3 irudia.** HITESen arkitektura.
- V.4 irudia.** Lan-memoriaren errepresentazioa.
- V.5 irudia.** Tartehizkuntzaren aldakortasuna.
- V.6 irudia.** *fenomeno* kontzeptuaren definizioa.
- V.7 irudia.** *helburu* kontzeptuaren definizioa.
- V.8 irudia.** Helburuen adibideak.
- V.9 irudia.** Ataza baten egitura.
- V.10 irudia.** *Ikasle_baten_tartehizkuntz_egituren_sorkuntza*.
- V.11 irudia.** Atazen antolaketa.
- V.12 irudia.** Fenomeno_hipotesiak_diskriminatzea.
- V.13 irudia.** Arrazonatze- eta erabakitze-modulua.
- V.14 irudia.** Agendaren egitura.
- V.15 irudia.** Ikaskuntza-elementua.
- V.16 irudia.** Fenomeno-ezezagunak-aztertzea.
- V.17 irudia.** Ikaskuntza-inguruneko objektuen hierarkia.
- V.18 irudia.** HITES sistemaren objektu-hierarkia.
- V.19 irudia.** Saio baten osagaien objektu-hierarkia.
- V.20 irudia.** Tartehizkuntz egituraren osagaien objektu-hierarkia.
- V.21 irudia.** Menu nagusia.
- V.22 irudia.** Tartehizkuntzaren eraketa.
- V.23 irudia.** Ataza eta helburuen prozesua elkarrekintzan zehar (I).
- V.24 irudia.** Corpusaren aukeraketa.
- V.25 irudia.** Ataza eta helburuen prozesua elkarrekintzan zehar (II).
- V.26 irudia.** Ataza eta helburuen prozesua elkarrekintzan zehar (III).
- V.27 irudia.** Ezagutza linguistikoaren azterketa.
- V.28 irudia.** Erregela linguistikoaren edizioa.
- V.29 irudia.** Fenomeno linguistikoaren diskriminazioa.
- VI.1 irudia.** MUGARRI ingurunearen ezagutza-baseak
- VI.2 irudia.** Ezaugarria kontzeptuaren definizioa.

- VI.3 irudia.** Ikasle-ereduaren definizioa.
- VI.4 irudia.** Diagnosi-ezagutza ezagutza-basea.
- VI.5 irudia.** IDAZKIDEn arkitektura.
- VI.6 irudia.** Saioaren ikasle-eredu iraunkorraren definizioa CLIPSen.
- VI.7 irudia.** Saioaren ikasle-eredu tenporalaren definizioa.
- VI.8 irudia.** Ikaste-estrategiaren definizioa CLIPSen.
- VI.9 irudia.** Taldekatze-estrategia IDAZKIDEn.
- VI.10 irudia.** IDAZKIDEn leiho nagusia.
- VI.11 irudia.** *amona* hitzaren deklinabidea.
- VI.12 irudia.** **sur* desbideratzerako proposamenak.
- VI.13 irudia.** IRAKAZI sistemaren objektuen arteko erlazioak.
- VI.14 irudia.** Ikasle baten ikaste-historiaren bilketa.
- VI.15 irudia.** Ikasle baten desbideratzeen diagnosi-azterketa.
- VI.16 irudia.** Datuen esplotazioa.
- VII.1 irudia.** Ingurunearen garapenaren eskema: ebaluazioa hobekuntzaren eragile gisa.
- VII.2 irudia.** Automatikoki sortutako fitxa irakasleak aztertua (letra etzanaz) eta ebaluatua (letra larriz).
- VII.3 irudia.** Hitzen analisisien tasa goi-mailako ikasleen corpusetan.
- VII.4 irudia.** Corpuseko hitz forma desberdinak, analisisian anbiguo direnak (%).
- VII.5 irudia.** Hitzen analisisien tasa erregeladun analisisetan.
- VII.6 irudia.** Hiztegiaren finketa goitik beherako metodologiaren aplikazioan.
- VII.7 irudia.** Hiru ikaskuntza-maila nagusietako corpusen anbiguotasuna (%).
- VII.8 irudia.** Hiru ikaskuntza-maila nagusietako corpusetan erabilitako hiztegiaren aberastasuna.
- VII.9 irudia.** Goiko mailako bi corpusen hiztegi-aberastasuna.
- VII.10 irudia.** Kategoría gramatikalen erabilera (%) goiko, erdiko eta beheko hizkuntz mailetan.
- VII.11 irudia.** Kategoría gramatikalen erabilera (%) ikaskuntza-maila igo ahala.
- VII.12 irudia.** Ikasle baten eragina kategoría gramatikalen azterketan.

TAULEN ZERRENDA

III.1 taula. Desbideratze-moten tratamendua.

VII.1 taula. Inguruneko probetan erabilitako corpusak.

VII.2 taula. Esaldi baten analisi sintaktikoa.

A ERANSKINA

Azaleko maila hutsezko desbideratzeen sailkapena.

A. Azaleko maila hutsezko desbideratzeen sailkapena.

Eranskin honen edukia ondokoa da:

- Azaleko maila hutsezko desbideratze konkretuen sailkapena, IRALE eskolatik hartutako 50 bat testuetatik (1991/92 ikasturtea) aterata.
- Aurreko desbideratze konkretuen kodeak § III. kapituluan azaldutako azaleko maila hutsezko desbideratzeen sailkapenean kokatuta.
- Desbideratze bakoitzaren fosilizazioa fenomenoarekiko kokapena (sistematikoak diren ala ez markatuz): sistematikoak direnak **(S)**; asistematikoak edo kasu partikularrak direnak **(A)**; eta sistematikoak izan litezkeenak **(S*)**. Honetaz gain, ikasleen testuetatik ateratako adibide konkretuak ematen ditugu.

A1 Omisioa edo kenketa

a) Letra bat kendu.

Kode-adibideak: LEKBA / LEKHH / LEKTH / LEKBR

LEKBA (*heriotzø*) **(S)** LEKHH (*øor*) **(S)**

LEKTH (*biøotza*) **(S)** LEKBR (*zeø*) **(S)**

LEKTE (*oihanøan*) **(S)**

b) Morfema bat kendu.

Kode-adibideak: MOK-EN /MOK-RE /MOK-TA

MOK-EN (*direnøtzat*) **(S*)** MOK-RE (*soldataøkin*) **(S*)**

MOK-TA (*nagusieøtik*) **(S*)**

A2 Gehikuntza

* Marka bikoitza

Hitz mailan honelako desbideratzeak ez dira topatu 50 testuetan zehar.

* Erregularizazioa

Hitz mailan honelako desbideratzeak ez dira topatu 50 testuetan zehar.

* Gehikuntza sinpleak

a) Letra bat gehitu.

Kode-adibideak: LEGBA / LEGTE / LEGHH / LEGTR

LEGBA (*basauntzaei*) (S) LEGTE (*eskeleak; euskalduneeen; neskeen*) (S*)

LEGGH (*harrisku*) (S*) LEGTR (*zorritzarreko*) (A)

b) Morfema bat gehitu.

Kode-adibideak: MOG-A /MOG-TU

MOG-A (*institutoako*) (S*) MOG-TU (*gordetu*) (S*)

(Oharra: kontuan izan LEGBA eta MOG-A gauza bera direla, baina ikuspegi desberdina jarraitzen dute)

A3 Erakuntza

* Erregularizazioa

Hitz mailan honelako desbideratzeak ez dira topatu 50 testuetan zehar.

* Archiformak

Hitz mailan honelako desbideratzeak ez dira topatu 50 testuetan zehar.

* Forma trukaketa

a) Letra bat ordezkatu.

Kode-adibideak: LEOEAE / LEOEEA / LEOEDR / LEOEEI / LEOEHJ / LEOEJG / LEOEKG / LEOENM / LEOEOU / LEOEUO / LEOEXJ / LEOESX / LEOEXS / LEOESZ / LEOEZX / LEOEXZ / LEOEZTX / LEOEZTZ / LEOETSS / LEOETZZ

LEOEAE (*erreztasuna*) (S)

LEOEEA (*ospatsuak*) (S)

LEOEDR (*inbiria*) (S*)

LEOEEI (*erlien*) (S)

LEOEHJ (*zijoala*) (S*)

LEOEJG (*imaginazioan*) (S*)

LEOEKG (*gonbaraziorik*) (S*)

LEOENM (*trampa*) (S*)

LEOEOU (*freskua*) (S*)

LEOEUO (*partido*) (S)

LEOEXJ (*axolarik*) (S*)

LEOESX (*experientzia*) (S)

LEOEXS (*piskanaka*) (S)

LEOESZ (*garrazia*) (S)

LEOEZX (*arasoak*) (S)

LEOEXZ (*gaizotasunez*) (S)

LEOEZTX (*baxter*) (S)

LEOEZTZ (*txorakeriak*) (S*)

LEOEZTZ (*zehatzki*) (S*)

LEOETSS (*pensatzen*) (S)

LEOETZZ (*proporzioak*) (S*)

b) Morfema bat ordezkatu.

Kode-adibideak: MOO-TEN-TZEN / MOO-TZEN-TEN

MOO-TEN-TZEN (*jasatzen*) (S*) MOO-TZEN-TEN (*irteten*) (S*)

***Asmaketa**

a) *Letra bat asmatu.* (Asmatutako letrak ama-hizkuntzatik datoz, dena den euskararen ikuspegitik asmatutako letrak direla kontsideratuko ditugu.)

Kode-adibideak: LEAV / LEAC

LEAV (video) (S) LEAC (magnetoscopia) (S*)

b) *Morfema bat asmatu.*

Kode-adibideak: Ez dira bereizi kode desberdinak. Denei MOA kodea jarri zaie.

Adibideak:

MOA (*ekipontzat*) (S*)

c) *Lema bat asmatu.*

Kode-adibideak: Ez dira desberdinak kode desberdinak. Denei LMA jarri zaie. Dena den, desbideratzeen sakoneko arrazoietan sartu arren, interesgarri dirudi aipatzea atal honetan topatutako desbideratzeak hiru multzo nagusitan bana daitezkeela:

c1) *Gaztelerak eragindako transferentziak direnak (Erderakadak).* LMA-E

c2) *Normalizazio prozesuaren ondorioz sortutakoak.* LMA-N

c3) *Banatuta idatzi beharko liratekeenak, hitz elkarketa izateagatik; eta ikasleak, ordea, hitz bakarria balitz idatzitakoak.* LMA-B

LMA-E (*idea* <- *ideia*) (S) LMA-N (*haundi* <- *handi*) (S)

LMA-B (*hitzegiten* <- *hitz egiten*) (S)

A4 Ordena aldaketa

Kode-adibideak: LETH

LETH (*lehiøoa*) (S*)

Orain arte azaldutako kode-adibideak IRALE eskolatik hartutako 50 bat testuetatik (1991/92 ikasturtea) atera ditugu, beraz, honek ez du kentzen kode-adibide berriak azal litezkeenik beste testu batzuen azterketan. Adibidez, kodeketa hau egin ondoren ILAZKI hizkuntz eskolako zenbait ikaslek (1992/93 ikasturtea) idatzitako 20 idazlan desberdin lantzean ondoko kode-adibide berriak topatu genituen hitz mailako desbideratzeen artean:

Gehikuntza sinpleetan: LEGTH (*iha*)

Erakuntzan: MOO-TA-TAK (*aspertutak*) MOO-ELAKO-LAKO (*dakiølako*)

(Oharra: kontuan izan MOO-ELAKO-LAKO LEKTE bezala ikus daitezkeela)

Aipatutako 20 idazlanetan hitz mailako desbideratzeen azterketaren bigarren fase batean "zuzenak" izan litezkeen hitzak, baina beren testuingurua kontuan izanik "okerrak" izan litezkeenak ere landu genituen. Azterketa honetatik honako desbideratze berriak topatu genituen:

Omisioa edo kenketan: MOK-A (*hamaiketako*)

Gehikuntza sinpleetan: MOG-KO (*mendebaldeko*)

Erakuntzan:

MOO-EZ-AGATIK (*egunagatik*) MOO-EZ-ETIK (*gauetik*)

MOO-RA-A (*egitea*) MOO-A-RIK (*arazorik*)

MOO-TZEKO-TZERA (*bukatzera*) MOO-REN-KO (*argiako*)

B ERANSKINA

Desbideratze linguistikoen sailkapena.

B.1 Hitz-mailako desbideratze linguistikoen sailkapena.

a) Erregelen bidez tratatu diren fenomeno orokorrak:

— Zenbait arazo orokor tratatzen dutenak:

- 1) *H*-aren arazoa. Bi erregela daude hemen, ezabaketa eta gehikuntza tratatzen dituztenak. (*ikus LEKHH; LEKTH; LEGHH; LEGTH; LETH azaleko maila hutsezko sailkapenean*)

Adibideak: zihur, hesparru, mee, au...

- 2) *-a* organikoaren arazoa. Normalean ezabaketa (*gauz*), eta konkretuki zenbait leku-izen berezitan gertatzen dena (*Azpeitia* baina *Azpeitira*). Lau erregela daude hemen guztira. (*LEKBA; LEGBA*)

Adibideak: gauz, idei, Azpeitiara...

- 3) Txistukari frikari eta afrikariak. Bi daude: bata frikaria sortzeko eta bestea, afrikaria. (*LEOEZTX; LEOEZTZ; LEOETSS; LEOETZZ*)

Adibideak: aditzlagun, persiana, alabanza...

- 4) Leherkari gor/ozen alternantzia. Sei erregela daude. (*LEOEKG*)

Adibideak: pake, ahosdun, hozgailu, boskarren...

- 5) *f/p/b* alternantzia. Ez da horren fenomeno orokorra. Hiru erregela daude.

Adibideak: apari, fago, farre...

- 6) Txistukarien arteko aldaketak. Sei erregela daude hemen ere. (*LEOESX; LEOEXS; LEOESZ; LEOEZX; LEOEXZ; LEOEZTX*)

Adibideak: goisean, arratzaldean...

- 7) Palatalizazioaren fenomenoak. Bi erregela: *-in/-ñ-* eta *-il/-ll-*

Adibideak: ille...

— Maileguen egokitzapenetan sortzen direnak:

- 8) *x/j* alternantzia. (*LEOEXJ*)

Adibideak: keja...

- 9) *g/j* alternantzia *-e, -i* bokalen aurrean. (*LEOEJG*)

Adibideak: filologia, erligio...

10) -o/-u alternantzia hitz-bukaeran. (LEOEOU; LEOEUO)

Adibideak: exenplo, alfabetu...

11) Zenbait hitz-bukaeratan, -isis, -is dutenetan, -s horren desagerpena gertatzen da.

Adibideak: analisis...

12) -e epentetikoa. (LEKTE)

Adibideak: roilu, semena...

Erregela berriak egin behar izan dira: hizkuntz ikasleen testuetan azaleko maila hutsean egindako azterketa eta XUXENen errore tipikoen maila linguistikoaren artean egindako erlazioa ikusi ondoren, ikasleen testuetan "errore" tipiko berriak, hobe esanda desbideratzeak, identifikatzen ditugu. Hauetako batzuek aldaketak eskatzen dituzte dagoeneko existitzen diren erregeletan; beste batzuek, ordea, fenomeno zeharo berriak ditugu, azken hauek ondokoak dira: *LEGTE / LEGTR / LEOEAE / LEOEEA / LEOEEI / LEOEHJ / LETH / LEOENM / LEAV / LEAC*

Testuetan adibiderik topatu ez dugun arren, MOO-ELA-LA (*dakiøla*) adibideak maiz ikusten dira ikasleen testuetan. MOO-ELAKO-LAKO eta MOO-ELA-LA konparatzen baditugu LEKTE (**Omisioa edo kenketan**) kodearen adibideak bezala ere sailka genitzake. Beraz, fenomeno orokorrak (erregelak) diren edo fenomeno espezifikoak diren (morfemaka datu-base lexikoan) askotan ez dago hain garbi.

b) Morfemaka tratatu beharreko fenomeno espezifikoak.

Multzo honetan sartzen dira azaleko maila hutsezko *LMA* eta *MOA*, baita *MOK*, *MOG* eta *MOO*.

Bigarren multzo honetan XUXEN zuzentzaile ortografikoan erabilitako kodeak azaltzeari ekingo diogu. Kasuren batean, gure ustez, kodeak sakoneko sailkapenean sartu beharko lirateke.

– LD_FO: arazo fonetikoek eragindako lema desberdinak (Horrela deituak, nahiz eta jakin denak ezin direla fonetikaren bidez esplikatu).

– LD_MA: maileguen egokitzapenean sortzen diren lema desberdinak.

Badakigu, jakin, bi horien arteko bereizketa, datozen kasuetan bezala, ikuspuntuan edo irizpidean datzala. Hau da, biak izan daitezke arazo fonetikoek eraginak, baina bigarrena

osatu bada horien inguruan informazio gehiago eman nahi dugulako izan da, esana baitago zenbait fenomeno maileguen egokitzapenean bakarrik sortzen direla.

- KONP: konposizioan ematen direnak.
- ERAT: eratorpenean ematen direnak.
- DE_ER: erakusleen deklinabidean sortzen direnak.
- DE_LE: leku-izenen deklinabidean sortzen direnak.
- A_FAK: aditz faktitiboa eratzean sortzen direnak.
- ATZ: menpeko esaldiak sortzeko erabiltzen diren atzizkiekin sortzen direnak.
- DIAL: aho-hizkerak nahiz dialektalismoak eragindakoak.
- DE_DI: gehienetan dialektalismoak eragindakoak.

Azken biak, DIAL eta DE_DI sakoneko sailkapenean sartu beharreko kodeak ditugu.

Fenomeno espezifiko hauek morfemaka tratatzeko, sistema lexiko paralelo bat dugu. Sistema honetan morfema desbideratuak sartzen dira sistema lexiko orokorrean dagokion morfema zuzenarekin lotuta. Orain, ikusiko ditugu adibide batzuk. Adibideak ondoko formatuan emango ditugu:

- LX_OSAGAIA deritzon multzoa dugu lehenik: hemen, sarrera, sistema lexiko zuzenean nola azaltzen den agertuko da. Sarrera hori izan daiteke deklinabideko atzizki bat, erator-atzizki bat, erlazio-atzizkia, lema (juntagailuak, lokailuak, determinatzaileak, etab. barne), izena, adjektiboa, adberbioa, aditza...

- ET_OSAGAIA ondoren. Hau da errore tipikoa.

- ET_LX_DEITURA: gure sisteman, elementu hori zein azpilexikoitan kokatzen den (*Agirre, et. al. 93*).

- E_KOD: errore-kodea da hau (ikusiko dugu jadanik).

Hona adibideak:

<u>LX OSAGAIA</u>	<u>ET OSAGAIA</u>	<u>ET LX DEITURA</u>	<u>E KOD</u>
\$tu	\$u	et_tu	ATZ
\$araz	\$eraz	et_araz	A_FAK
\$gatik	\$gaitik	et_gan_tik	DE_DI
ahots	abots	et_izenak	LD_FO
ardo	arno	et_izenak	DIAL
independente	independiente	et_adjektiboak	LD_MA

B.2 Esaldi-mailako desbideratze linguistikoen sailkapena.

Ondorengo sailkapena 9. eta 12. mailetako hainbat ikaslek idatzitako 150 idazlanetatik atera dugu. Fossilizazio fenomenoaren ikuspegitik interesgarri diren fenomeno linguistikoak aukeratu ditugu. Fenomenoak zein testuingurutan ematen diren aztertu dugu, besteak beste.

KOM => Aditza eta beste osagai batzuen arteko komunztadura pluraltasunean (osagaia=sujetua edota objektua).

KOMAP —> Aditza pluralean.

Testuinguru-baldintzak

KOMAP_MUL Osagaiak multzoa adieraztea.

KOMAP_ERA Osagaia erakuslea izatea.

KOMAP_KIN Esaldian *noren* edota *norekin* edota *non* kasuko sintagma pluralean agertzea.

KOMAP_ABS Nominalizazioa eta *nor* edota *nork* kasuko sintagma pluralean agertzea.

KOMAS —> Aditza singularrean.

Testuinguru-baldintzak

KOMAS_KOP Osagai den sintagmaren hitz-kopurua 1 baino handiago izatea.

KOMAS_OBJ Osagaia objektu zuzena edo ez-zuzena izatea eta pluralean egotea. Eta sintagmaren hitz-kopurua 1 izatea.

KOMAS_ZEN Osagai den sintagman zenbakia agertzea.

KOMAS_ETA Osagai den sintagman *eta* partikula agertzea.

KOMAS_ERL Aditza erlatibozko esaldi baten aditz nagusia izatea.

KOMAS_BER Aditz berezia ikaslearen ikuspegitik (delakoa, dagokio...)

KOMSE —> Sujetua ergatiboan.

KOMSN —> Sujetua nominatiboan.

KOMOE —> Objektua ergatiboan.

KOMOS —> Objektua singularrean.

MUG => Sintagma baten mugatasuna.

MUGMU —> Sintagma mugatua edo sintagmaren barruko osagairen bat mugatua.

Testuinguru-baldintzak

MUGMU_ERA Sintagman erakuslea agertzea.
MUGMU_GAL Sintagman galdegilea agertzea.
MUGMU_IND Sintagman indefinitua agertzea.
MUGMU_ZEN Sintagman zenbatzailea agertzea.
MUGMU_BAT Sintagman *bat* partikula agertzea.
MUGMU_BER Eraikuntza berezi baten sintagma izatea, hots, *zertan bihurtu, zeren gisan, nola egon* .
MUGMU_ADJ Sintagman adjetiboa agertzea.
MUGMU_INA Sintagman izen nagusia agertzea.

MUGMG —> **Sintagma mugagabea.**

Testuinguru-baldintzak

MUGMG_BER Eraikuntza berezi baten sintagma izatea, hots, *zertan zehar, noren bidez, zer izan, zerbait iruditu*..
MUGMG_ABS Sintagmaren esanahia erreferentzia abstraktua edota orokorra izatea.
MUGMG_ETA Sintagman *eta* partikula agertzea.

MEN => Aditzaren bukaera ondoko menpeko esaldietan: konpletiboak, zehar galderak eta erlatibozkoak.

MENNB —> **Aditza "-n"z bukatu.**

MENNB_ZUZ Galdera zuzena izatea.

MENNB_ERL Erlatibozkoa + deklinabidea behar izatea.

MENLA —> **Aditza "-la"z bukatu.**

MENLA_BE1 Menpeko esaldiaren eraikuntza berezi baten aditza izatea, hots, "zeren eta.....-ela" "ea-ela" "ustez -ela" "-enez -ela" "-entzat -ela".

MENLA_BE2 Esaldi nagusiarekin lotuta dagoen eraikuntza berezi baten aditza izatea, hots, "-enez.....-ela" "seguraski.....-ela" ".....-la besterik ez"

MENNZ —> **Aditza "-n + deklinabidea"z bukatu.**

MENNZ_GAL Menpeko esaldian galdegilea agertzea.

MENEZ —> **Aditza bukaerarik gabe.**

MENEZ_BES Esaldian beste aditz batek "-la" bukaera izatea.

MENEZ_KOR Menpeko esaldia koordinatua izatea.

MENEZ_GAL Menpeko esaldian galdegilea agertzea, edo bestelakoak (zeren eta)

MENEZ_ERL Erlatibozkoa faltan

MENEZ_ESK Aditz nagusiak hala eskatzea (-la)

C ERANSKINA

Desbideratzeen sakoneko arrazoien sailkapena

C. Desbideratzeen sakoneko arrazoiaren sailkapena.

Eranskin honen edukia ondokoa da:

- Desbideratzeen sakoneko arrazoiaren sailkapena.
- Ikaste-estrategiak arrazoiaren identifikatuta.

Ondorengo sailkapenean arrazoiaren artean honako hauek desberdinduko ditugu: zeintzuk dauden lotuta testu osoarekin eta zeintzuk dauden lotuta aztertzen den egitura linguistikoko konkretuekin. Beraz, bi arrazoi-multzo bereiziko ditugu: testu mailakoak eta esaldi mailakoak, azken hauek egitura linguistikoko konkretuei lotuta — kontuan izan esaldi mailan gerta daitekeela esaldi berean egitura linguistikoko bat baino gehiago aztertu behar izana, ondorioz, bakoitzerako arrazoiak aztertu beharra —.

1. Testu mailako azterketa

Arrazoi posibleak: testu bakoitzerako eragina eduki dezaketen baldintzak.

Bidalitako ariketa

ARI Idazlana egiterakoan bidalitako ariketa-mota (entzundako zerbaiten laburpena, irakurritako zerbaiten inguruan idatzi, idazlan zeharo librea...).

Ikaslearekin lotutako baldintzak

IP Ikaslearen gaiarekiko inplikazio pertsonala idazlana egitean.

IP_ Ikaslearen gaiarekiko inplikazio pertsonal eza idazlana egitean.

AZ Ikasleak azkarregi idaztea, testuaren antolamendua eta zuzentasunean gutxi erreparatuz. (*SIM* —> *Autoebaluazioa/ Dedukzioa/ Planifikazio Funtzionala*).

AZ_ Ikasleak poliki idaztea, testuaren antolamendua eta zuzentasunean erreparatuz. (*SIM* —> *Autoebaluazioa/ Dedukzioa/ Planifikazio Funtzionala*).

ML Material lagungarria erabiltzea, hots, hiztegiak, gramatika, ikasgelan hartutako oharrak, e.a. (*SI* —> *Baliabideen erabilpena*).

ML_ Material lagungarria ez erabiltzea, hots, hiztegiak, gramatika, ikasgelan hartutako oharrak, e.a. (*SI* —> *Baliabideen erabilpena*).

2. Esaldi mailako azterketa

2.1. Ikaslearen ezagutza:

HG Ikaslearen hutsune gramatikala, hau da, euskaldun aditu baten bertsioa egin ahal izateko ezagutza nahikoa du.

HG_ Ez da ikaslearen hutsune gramatikala, hau da, euskaldun aditu baten bertsioa egin ahal izateko ez du ezagutza nahikorik.

X Ez dakigu.

2.2. Arrazoi posibleen azterketa: egitura linguistiko bakoitzerako eragina eduki dezaketeen baldintzak.

Testuaren egiturarekin lotutako baldintzak

TEM Idatzitako testu-mota (eskutitza, idazlana, elkarrizketa, laburpena...).

LUZ/LUZ_ Esaldi luzea (esaldiaren hitz-kopurua 10 baino handiago denean).

MEN Esaldian subordinatuak edota koordinatuak agertzea (erlatibozkoak, konpletibak, adbersatiboak...).

URR/URR_ Aztertzen ari garen egitura linguistikoaren osagaien arteko urruntasuna (tartean 5 hitz edo gehiago. *Adibidea: Zure **txakurra** zaunka egin zuen*).

HUR/HUR_ Aztertzen ari garen egitura linguistikoa eta esaldiaren beste osagai batzuen arteko hurbiltasuna (tartean 5 hitz baino gutxiago. *Adibidea: Neska **txakurrekin** agertu ziren*).

ELI Esaldiaren barruan osagaien bat eliptiko egotea (*adib. oso **gutxi** izaten **zirelako** baserriaren jabe*. Kasu honetan gazteleraren eraginarekin batera — porque muy **pocos solian** ser dueños del caserío — **jende** hitza eliptiko egotea ere arrazoa izan daiteke ikasleak honelako esaldia egiteko).

INT Bestelako interferentzia linguistikoak: esaldi barruan dauden beste osagai batzuen agerpenak egitura linguistikoan eragina dute (*adib. **ikasle** batzuk aditza pluralean jartzeko joera izaten dute aditzaren sujetoa multzoa adierazten duen hitza denean*).

KON Aztertzen ari garen egitura linguistikoa bera konplexua izatea.

Ikaslearekin lotutako baldintzak: ikaslearen barruko faktoreak

DIA H2ren euskalkien eragina (ikaslearen euskalkia). Hau euskaldun zaharren kasuan bakarrik gerta daiteke. (*SI —> Itzulpena*).

NAH H2ren bi egituren arteko nahasketa (*adib. **nik** gogoratzen **naiz***. Kasu honetan ikasleak nahastu ditu ondoko bi egitura: "nik gogoratzen dut" eta "ni gogoratzen naiz") (*SI —> Elaborazioa*).

ERD Ikasleak ezagutzen duen beste hizkuntz arrotz baten eragina (ingelesa, frantsesa...) (*SI* → *Transferentzia*).

GAZ Ama-hizkuntzeko mailegua izanik, H2an ama-hizkuntzan bezala erabiltzea (*adib. baserri propio zutenek*. Kasu honetan esan beharko luke: "baserri **propioa** zutenek", baina, gazteleran bezala erabili du — los que tenían caserío **propio** —) (*SI* → *Transferentzia; Hitz gakoa*)

(*Kasu hau kontzeptu-mailako transferentzia kontsidera genezake: ikasleak gaztelerako esaldian "caserío propio" kontzeptuaren mugagabetasuna ikusten du, hau da, ez da baserri konkretua; ondorioz, euskaraz mugagabetasunean -a atzizkirik jartzen ez denez, "baserri propio" egitura osatzen du*).

H2A H2ko egitura izanik, egitura hori ama-hizkuntzan tartekatuta erabiltzea, adibidez: *he visto a mi andereño*. Horren ondorioz, H2an ama-hizkuntzan bezala erabiltzen du (*adib. nere andereño okertu egin zen* — andereñoa esan beharrean —) (*SI* → *Transferentzia; Hitz gakoa*).

ITZ Ama-hizkuntzaren itzulpena (oharra: aurreko biak eksklusiboak dira honekiko, beraz, aurreko bietako bat arrazoa izanez gero, hau ez dugu arrazoitzat emango) (*SI* → *Itzulpena*).

AMA Ama-hizkuntzan bertan esatea (*adib. ezagutzen duzu, mucho ruido y pocas nueces*)(*SK*).

ANT H2ren bi egituren arteko antzekotasuna (*adib. lehenengo/lehenengo — ikerlari/ikerlariri*) (*SB*).

MAN Egitura horri ikasleak mania izatea (*adib. ikasle batek zenbait kasutan, esaterako, erlatibozkoetan, aditza iraganean, e.a., "n" bukaeran ezabatzen zuen, erlatibozko esaldiak ikastean mania hartu zielako azken hauei*) (*SB*).

ING Itzulinguruzko azalpena edo deskribapena (*adib. ikuilua* esan beharrean *baserrian animaliak gordetzeko tokia* esan) (*SK*).

ASM Ikaslearen asmaketa. Esan nahi duena adierazteko forma egokirik bigarren hizkuntzan aurkitzen ez duenean bere forma propioa sortzen du (*SK*).

SIN Sinplifikazioa esaldiaren mezuari lehentasuna emanaz (*adib. haurrek liburuak ekarri dute* — "haurrek liburuak ekarri **dituzte**" esan beharrean. Kasu honetan ikasleak, agian, esaldiaren mezuari lehentasuna eman dio, gramatikaren zuzentasuna baztertuz) (*SK*).

AHA Ikasleak ahaztua (*SB*).

DEN Ikaslea oso denbora laburrean bigarren hizkuntzaren maila altua hartzea.

AHO Ikaslearen ahozkoaren eragina (*SI*).

Ikaslearekin lotutako baldintzak: Ingurunearen faktoreak

EUS H2ren euskalkien eragina (euskaldun zaharregandik jasotako *input-a*) (*adib. ministro ministrari* esan beharrean → hobesten diren hitzak) (*SS*).

KAR Kartelak, iragarkiak, egunkaria, telebista, e.a.en bidez jasotakoa (*SS*).

- NOR** Normalizazio-prozesuan finkatu gabeko egitura linguistikoa (**SS**).
- MAT** Irakasleak ikaste-prozesuan erabilitako materialek sorrerazia (**SR**).
- IRA** Ikaslearen irakasleek ikasgelan erabili ohi duten materialek sorrerazia (**SR**).
- IKA** Ikasgelako ikaskideek erabili ohi duten materialek sorrerazia (**SR**).

Sailkapen honetan ikusten dugu zenbait arrazoi inola ere erlazionatu ez duguna, adibidez: Idazlana egiterakoan bidalitako ariketa-mota (ARI); Ikaslearen inplikazio pertsonala idazlana egitean (IP); Idatzitako testu-mota (TEM); Esaldi luzea (LUZ/LUZ_); e.a. Arrazoi hauek tartehizkuntzaren errepresentazioarekin lotuta daude. Esaterako, ARIekin lotutako arrazoiak Selinkerrek eta Douglas-ek (1994) definitutako ariketa-motak (*activity-types*) dira. Modu berean, IP discourse domain delakoarekin erlazionaturik dago. TEM, LUZ eta beste zenbait arrazoi tartehizkuntzaren kontestualizazioan kokatuko genituzke; TEM eta LUZ tartehizkuntzaren egituren testuinguru linguistikoak lirateke. ARI, hau da, irakasleak idazlana egiteko ikasleari bidalitako ariketa-mota, ordea, testuinguru ez-linguistiko bezala ikus genezake.

D ERANSKINA

Landa-lanean ikasleei jarritako probetako bat.

D.1 Irakasleen hipotesien zerrenda.

Itzulpen ariketan:

A.1 Esaldi konpletiboen kasuan, esaldia motza izanez gero, ikasleek *-ela* atzizkia zuzen jartzen dute.

A.2 Komuntadurari dagokionez, aditzak konuntatu behar duen objektua pluralean dagoenean aditza singularrean jartzen dute sinplifikatuz.

A.3 Bigarren hizkuntzan ikasitakoa, gero ama-hizkuntzaren diskurtsoan tartekatua erabiltzearen poderioz, azkenik bigarren hizkuntzan berrerabiltzean euskal arauak ez dituzte errespetatzen.

A.4 Deklinabidean mugatu beharrean mugagabe egiten dute ama-hizkuntzaren kontzeptu abstraktuen kasuetan, esaterako, gaztelerazko *sobre burros* itzultzean *astori buruz* edo *astoz* idazten dute.

A.5 Deklinabidean gaztelerazko *otro* euskaratzean mugatu egiten dute, ez baitute *bat* mugatzailea idazten.

A.6 Sintaxian, esaldi luzea izanik, zehargaldera bat, konpletiboa beste bat eta *eta* lokailua tartean dela, hipotesia da interferentzia eragingo ote duen *eta* lokailuak, eta ondorioz batean bakarrik markatuko duten subordinatua. Gainera, *dirudienez*, *antza denez* edo *-nez* idatziz gero konpletiboaren *-ela* markatuko lukete ikasleek, esaterako, *dirudienez ...la*.

A.7 Ama-hizkuntzaren mailegua izanik, bigarren hizkuntzan ama-hizkuntzan bezalaxe erabiltzen dute, adibidez, *kotxe propio zutenek*.

A.8 Komuntadura kasuetan esaldi motzetan ere sinplifikatzera jotzen dute.

Okerrak zuzen ariketan:

B.1 Esaldi zuzena eta zuzentzat emango dute.

B.2 Esaldiko osagaien bat pluralean egonez gero, aditzak singularra behar duenean pluralean erabiltzen dute. Adibidez, interferentzia sortuko du *zerekin plurala* kasuak, eta, ondorioz, aditzaren komuntadura subjektuarekiko ez da beteko, hau da, plurala idatziko dute ikasleek singularra izan behar badu ere.

B.3 Ama-hizkuntzetik itzultzearen ondorioz mugagabea erabiliko dute behar ez denean.

B.4 Esaldi ez-egokia eta zuzentzat emango dute.

B.5 Aposizio kasuetan sinplifikatzera joko dute.

Esaldiak berridatzi ariketan:

C.1 Bi esaldi oso labur lotzean, berridaztean *-ela* subordinazio marka jarriko dutelakoan gaude.

C.2 Bi esaldi berridaztean subordinatu beharreko esaldia luzexeagoa izanik akaso ez dute *-ela* marka gehituko.

C.3 Bi perpaus motz zehar galderaren bidez lotu behar dituztenean, motzak izanik *-en* sartzea espero dugu.

Idazlanetan:

Gaztelerazko idazlanetan darabilten hizkera eta aberastasuna ikusi nahi genituzke, ikasle batzuek gazteleraz ere egiten baitituzte desbideratzeak azkar idaztearen ondorio direnak edota idazteko ohitura ezarenak.

D.2 Hipotesi-zerrenda lantzeko ikasleei egindako azterketa.

Puntu honetan aurrekoan azaldutako hipotesi bakoitza aztertzeko erabakitako ariketa azalduko dugu. Adibidez, *Estaba claro que no estabais de acuerdo* esaldiaren itzulpenaren bidez A.1 hipotesia landuko da.

EUSKARA ITZAZU :

- A.1 Estaba claro que no estabais de acuerdo.
- A.2 Quieren ver casados a los hijos.
- A.3 Les dio los buenos dias.
- A.4 Hoy hablaremos sobre burros.
- A.5 Por favor, llamame en otro momento si vienes con otro problema.
- A.6 Al parecer se le ha olvidado que hoy le toca preparar la cena y cuando ha de presentarla.
- A.7 Los que tenian coche propio, se fueron antes.
- A.8 Han traído regalos de muchas casas pobres.

OKERRAK ZUZEN beharrezkoa denean soilik :

- B. 1 Leiho txuria irekitzeko moduan zegoen.
- B.2 Bapatean txakur bat hamalau edo hamabost ogirekin agertu ziren.
- B.3 Mantenuaren trukean egiten zuen lan.
- B.4 Bospasei asmatu zituzten.
- B.5 Etxeko mutilak, mutil zahar, ez dute ezkondu nahi.

ESALDIAK BERRIDATZI :

- C.1 Ez zeundeten ados. Garbi zegoen.
- C.2 Jonek esku-oihal urdina darabil. Guk hori ikusi dugu.
- C.3 Jarraituko du ? Ez dakigu.

REDACCION (10-150 PALABRAS - en castellano): una anécdota.

D.3 Suertatu diren eta ez diren hipotesiak azterketen emaitzetan.

Oraingoan, inoiz ez eta kasuren batean bete diren hipotesiak komentatzen ditugu. Adibide gisa, bi ikasleren kasua komentatzen dugu zehatz mehatz, hauek dira, il10s eta il10j.

Inoiz bete ez diren hipotesiak

A.6 Bi ikaslek, il10s-k eta il10j-k, bi subordinatuak markatu dituzte, espero ez genuen bezala. Dena den, il10s ikasleak *-en* jarri beharrean *-ena* erabili du (erlatibo jakin batekin nahastuta edo), eta il10j ikasleak berriz nominalizatuz sinplifikatu du (*-tea*). Bestetik *-enez* ere ongi, ez baitiote *-ela* erantsi (il10j-k), il10s-k berriz *omen* erabili du, beraz ez dute hipotesia bete, aurrerapena ote? Alegia, ikasi eta gainditu dutela.

C.2 Esaldi luzea izanik, *-ela* marka ez dute ahaztu, espero bezala.

Kasuren batean bete diren hipotesiak

Gainontzekoak bete direla esan dezakegu nahiz eta gehienetan ikasle guzti guztiek ez dute hipotesia bete. Ikusiko dugu zer egin duten kasu hauetan aztertzen ari garen bi ikasleek:

A.1 Bi ikasleok, il10s-k eta il10j-k, konpletiboa markatu dute, nahiz eta batek marka ez zuzena erabili (*-enik*).

A.2 Sinplifikazioa singularrera egin du il10s ikasleak, ez ordea il10j-k.

A.3 il10j-k hipotesia bete du aditzaren komunztaduran singularrean jartzean. il10s-k, berriz, euskal arauak errespetatu ditu, eta zuzen egin du.

A.4 Ama-hizkuntzatik itzulita mugagabe semantiko delakoa asmatu du il10s-k. il10j-k mugatu egin du, plurala egin beharrean singularra erabili duen arren.

A.5 Espero bezala il10j-k *beste* era mugatuan erabili du. il10s-k *beste ... bat* egitura erabili du.

A.7 *Kotxe propio* ongi euskaratu du il10s-k (nahiz eta menpeko aditzean objektuarekiko komunztaduran erratu: plurala egin du *beraien* interferentziagatik hedatuta). il10j-k hipotesia bete egin du.

A.8 Esaldi motza izanik ere, bi ikasleok, il10s-k eta il10j-k komunztadura sinplifikatu dute. il10s-k hiru hitzak nahastu ditu eta aldaketa semantikoaren ondorioz sinplifikatzearren akaso ondokoa idatzi du: *etxe pobretik opari asko*, eta il10j-k *etxe behartsuaren askotako opariak*.

B.1 Zuzen zegoena ontzat eman dute biek, nahiz eta il10s-k ortografian huts egin (*lehio**).

B.2 Interferentzia *norekin* kasuarekin urruntasunak eraginda: il10j-k subjektu eta aditzaren arteko komunztaduran pluralera jo du; il10s-k, berriz, ez du hipotesia bete zuzen egin duelako.

B.3 Hipotesi hau A.4rekin erlazionatuta dago. Ikusten da gainera nola il10s-k bai A.4 bai B.3 bete dituen, eta il10j-k, ordea, ez bat ez bestea.

B.4 il10s-k bete du hipotesia, il10j-k, ordea, singularrera jo du.

B.5 il10s-k aposizioa sinplifikatu du, hau da, hipotesia bete du, il10j-k, ordea, ez.

C.1 Bi ikasleok, il10s-k eta il10j-k, konpletiboa markatu dute, nahiz eta batek marka ez zuzena erabili (*-enik*).

C.3 Bi ikasleok, il10s-k eta il10j-k *-en* marka ezarri dute. il10s-ren kasuan, gainera, aberastasunez, *-enentz* ikasi baitu; il10j-k, berriz, puntuazio ez zuzena erabili du behar ez zen galdera ikurra sartuz.